Springer Proceedings in Physics 145

Burra G. Sidharth
Marisa Michelini
Lorenzo Santi *Editors*

Frontiers of Fundamental Physics and Physics Education Research

Springer

Springer Proceedings in Physics

Volume 145

For further volumes:
http://www.springer.com/series/361

Burra G. Sidharth · Marisa Michelini
Lorenzo Santi
Editors

Frontiers of Fundamental Physics and Physics Education Research

Springer

Editors
Burra G. Sidharth
Birla Science Centre
Hyderabad
India

Marisa Michelini
Lorenzo Santi
DCFA - Section of Mathematics and Physics
University of Udine
Udine
Italy

ISSN 0930-8989 ISSN 1867-4941 (electronic)
ISBN 978-3-319-00296-5 ISBN 978-3-319-00297-2 (eBook)
DOI 10.1007/978-3-319-00297-2
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013942324

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Frontiers of Fundamental Physics and Physics Education Research

Book of selected papers presented in the International Symposium Frontiers of Fundamental Physics—12th edition, held at 21–23 November 2011 in Udine, Italy

Editors of the Book

Burra G. Sidharth, *Birla Science Center Director, Hyderabad, India*
Marisa Michelini, *GIREP President, University of Udine, Italy*
Lorenzo Gianni Santi, *University of Udine, Italy*

Editorial Board

Editorial Secretariat

UNIVERSITY OF UDINE

EUROPEAN PHYSICAL SOCIETY

International Centre for Theoretical Physics

FFP12 International Symposium on
Frontiers of Fundamental Physics
21–23 November 2011 University of Udine, Italy

International Advisory Board

David Douglas Osheroff, *Hon. Chair, Stanford University, USA*
Burra G. Sidharth, *Director, B.M. Birla Science Centre, India*
Fernando Quevedo, *Director, International Centre for Theoretical Physics, Italy*
Marisa Michelini, *University of Udine, Italy*
Alessandro De Angelis, *University of Udine, Italy*
David Finkelstein, *Georgia Tech, USA*
Walter Greiner, *Director, Frankfurt Institute Advanced Studies, Germany*
Helmut Kröger, *University of Laval, Canada*
Marc Lachiez-Rey, *University of Paris, France*
Padma Kant Shukla, *Ruhr Universitat, Germany*

International Organizing Institutions

University of Udine, Italy — Department of Chemistry Physics and Environment (DCFA), Interdepartmental Center of Research in Education (CIRD), and Physics Education Research Unit—PERU with the cooperation of Faculty of Education (FASF), Faculty of Mathematical, Physical and Natural Sciences (FAMA), Scuola Superiore of the University of Udine

International Centre for Theoretical Physics (ICTP)
B.M. Birla Science Centre (BIRLASC)
International Centre for Mechanical Sciences (CISM)

Collaborating Scientific Institutions

European Physical Society (EPS)
Latin American Physics Education Network (LAPEN)
Group International de Research in Physics Education (GIREP)
European Science Education Research Association (ESERA)
International Commission on Physics Education (ICPE)
Multimedia in Physics Teaching and Learning Group (MPTL)
Multimedia Educational Resource for Learning and Online Teaching (MERLOT)
National Institute of Nuclear Physics (INFN)
Italian Society of General Relativity and Gravitational Physics (SIGRAV)
Italian Physical Society (SIF)
Association for Physics Teaching (AIF)

Local Organizing Committee

Cristiana Compagno, *Rector of the University of Udine*
Furio Honsell, *Mayor of Udine*
Carlo Tasso, *Rector's Delegate for India, University of Udine*
Gian Luca Foresti, *Dean, Faculty of Educational Sciences, University of Udine*
Franco Parlamento, *Dean, Faculty of Mathematical, Physical and Natural Sciences, University of Udine*
Lorenzo Fedrizzi, *Director of DCFA, University of Udine*
Lorenzo Santi, *Director of CIRD, DCFA, University of Udine*
Marisa Michelini DCFA, *University of Udine*
Mauro Francaviglia, *President of Italian Society of General Relativity and Gravitation SIGRAV*
Alfredo Soldati, *International Centre of Mechanical Sciences, Udine*
Giovanni Pauletta, *INFN Group & DCFA, University of Udine*
Marina Cobal, *DCFA, University of Udine*
Alberto Stefanel, *DCFA, University of Udine*
Stefano Vercellati, *DCFA, University of Udine*
Sri Rama Chandra Prasad Challapalli, *DCFA, University of Udine*

Secretariat and Technical Support of the FFP12 Symposium

Donatella Ceccolin, *CIRD Secretariat, Udine University*
Sandra Muzzin, Martina Scrignaro, *DCFA, University of Udine*
Silvia Adduca, Elena Larcher, Alessia Lofaro, *Public Relation*
Domelio Da Rù, Mario Gervasio, Filippo Pascolo, Alberto Sabatini, Mauro Sabbadini, Giorgio Salemi, *Technical Support*
Europhysics Conference Abstract Booklet—Printed by Litho Stampa, Udine, Italy, November 2011—ISBN 2-914771-61-4

Contents

Part IX Computational Physics

Part X Physics Teaching/Learning and Teachers Formation

Introduction

For nearly a decade and a half an International Symposium Series Frontiers of Fundamental Physics has attracted some of the brightest and greatest Physicists in the world.

The broad objective of the Series is to provide a platform for exchange of ideas and status reports of several frontier areas of physics like in particular Astronomy and Astrophysics, Particle Physics, Theoretical Physics, Gravitation and Cosmology, Computational.

These Symposia have been held in India, Italy (3), Spain, Canada, Australia, and France. The eminent physicists who have deliver Special Lectures over the years, sometimes more than once, have included Nobel Laureates Professors G.'t Hooft, S. Chu, C. Townes, Von Klitzing, P. Gilles De Gennes, D. D. Osheroff, H. Kroto, A. Leggett, and several other prominent scholars like Prof. Yuval Ne'eman.

A striking feature of the Series has been the involvement of physicists from different parts of the world including the Middle East, Europe, Russia, US, South America, Asia, and elsewhere.

The Symposium format has been one of Special Lectures, Invited Lectures, Contributed Papers, and Posters.

Over the years, Special sessions have been devoted to a wider range of topics which involve a larger community for people.

The International Symposium "Frontiers of Fundamental Physics—FFP12" held in Udine was organized by the University of Udine in collaboration with the International Centre for Theoretical Physics (ICPE) in Trieste, the BM Birla Science Centre in Hyderabad in India, the International Centre for Mechanical Sciences (CISM) in Udine and with the support of the Italian Society of General Relativity and Gravitational Physics (SIGRAV), the Italian National Institute of Nuclear Physics (INFN), and the Municipality of Udine. Scientific institution cooperating were the European Physical Society (EPS), the Latin American Physics Education Network (LAPEN), the Group International de Research in Physics Education (GIREP), the European Science Education Research Association (ESERA), the International Commission on Physics Education (ICPE), the Multimedia in Physics Teaching and Learning Group (MPTL), the Multimedia Educational Resource for Learning and Online Teaching (MERLOT), the Italian Physical Society (SIF), and the Association for Physics Teaching (AIF). FFP12 offered an excellent opportunity to meet high-level scholars working in various fields of research in physics to exchange expertise, ideas, results, and reflections on new prospects and for the development of research studies, to share and enhance knowledge. It addressed 12 important areas of research in Physics (Astrophysics, High Energy Physics and Particle Physics, Theoretical Physics, Gravitation and Cosmology, Condensed Matter Physics, Statistical Physics, Computational Physics, Physics, Mathematics, Physics Education, Training teachers of physics and Popularization of Physics). Among these, a significant role and importance for the dissemination of science and educational research, with particular attention to teacher training was achieved. The scientific value of FFP12 was enhanced by the participation of Prof. Douglas Osheroff of Stanford University (USA), a Nobel Prize laureate in physics for the discovery of superfluidity of helium-3. The wonderful success for this symposium is due in large part, to explicitly emphasize, of the 9 General Talks in the plenary session, the 29 invited talks in specific subject area, with 59 oral presentations, and 21 posters which formed the contents of the Symposium, presented in 30 sessions chaired by the respective heads of subject areas, spread over 3 days. They not only provided an overview on the fundamentals of physics but also on the advanced research in different thematic areas, providing everyone a great opportunity for an overview on the frontiers of research in different fields of physics. The poster session was structured and organized as adynamic forum among presenters and other participants in the Symposium. This has encouraged the discussion and exchange of ideas and perspectives. The European Physics Society (EPS) has enhanced the Symposium by funding, a poster prize for the best poster presented by a Ph.D. student. All conference participants had the opportunity to vote for, one of the best posters presented. The result of the votes was also consistent and in accordance with the Scientific Committee responsible for assessment.

The scientific level of the Symposium was provided by responsible of the topics, who are the members of the editorial board of the present book of selected papers. The topics' responsible operated with the help of anonymous referees for the selection of out of the 250 proposals received from 28 countries of 5 continents. They have taken care of the scientific activity associated with the topics of their expertise. The contributions published here were intact further selected by them contribution. Abstracts of the selected papers for presentation in FFP12 Symposium are published FFP12 Europhysics Conference Abstract Booklet (ISBN 2-914771-61-4).

Part I of this volume contains four papers that represent the general aspects of frontier treated in the Symposium on crystal lattices and theory of gravity and on mass generation, an history of physics contribution on Enrico Fermi and Ettore Majorana and a research-based contribution on teacher education. Parts II–XI contain the best papers received, respectively, on Astronomy and Astrophysics, Particle physics and high energy physics, gravitation and cosmology, condensed matter physics, statistical physics, theoretical physics, mathematical physics, computational physics, Physics Teaching/Learning and teacher formation, and popularization of physics.

This book is very rich in important results of the current research in physics, highlighting the great fertility and liveliness of research in the various fields of this discipline. The last three parts of this volume exemplify how the physics education research and the teacher education research contribute to the social responsibility of scientists to form a scientific culture in future generations, because science is a part of the cultural heritage of all citizens. The last part of the book provides examples of the important contributions of scientists to society through scientific popularization. We hope that this book is useful for the mutual understanding between scientists and research forming the basis for the cross-fertilization between them.

Burra G. Sidharth
Marisa Michelini
Lorenzo Santi

Part I
General Talks

Chapter 1
Two-Dimensional Anharmonic Crystal Lattices: Solitons, Solectrons, and Electric Conduction

Manuel G. Velarde, Werner Ebeling and Alexander P. Chetverikov

Abstract Reported here are salient features of soliton-mediated electron transport in anharmonic crystal lattices. After recalling how an electron-soliton bound state (solectron) can be formed we comment on consequences like electron surfing on a sound wave and *ballistic* transport, possible percolation in 2d lattices, and a novel form of electron pairing with strongly correlated electrons both in real space and momentum space.

1.1 Introduction

Electrons, holes, or their dressed forms as "quasiparticles", in the approach introduced by Landau [28, 29], play a key role in transferring charge, energy, information or signals in technological and biological systems [38]. Engineers have invented ingenious methods for, e.g., long range electron transfer (ET) such that an electron and its "carrier", forming a quasiparticle, go together all along the path hence with space and time synchrony. Figure 1.1 illustrates the simplest geometry between a donor (D) and an acceptor (A). Velocities reported are in the range of sound velocity which in bio-systems or in *GaAs* layers are about Angstrom/picosecond (Km/s). Such values are indeed much lower than the velocity of light propagation in the medium. Thus at first sight, leaving aside a deeper discussion concerning specific purposes [12, 35], controlling electrons seems to be more feasible with sound (or even supersonic) waves than with photons. Electron *surfing* on an appropriate highly monochromatic, quite strong albeit linear/harmonic wave has recently being observed [27, 32]. Earlier the present authors have proposed the *solectron* concept as a new "quasiparticle" [3, 4, 6–9, 17, 25, 40–43, 45, 46] encompassing lattice anharmonicity (hence invoking nonlinear elasticity beyond Hooke's law) and (Holstein-Fröhlich) electron-lattice interactions thus generalizing the polaron concept and quasiparticle

M. G. Velarde (✉) · W. Ebeling · A. P. Chetverikov
Instituto Pluridisciplinar, Universidad Complutense de Madrid, Madrid, Spain

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_1,

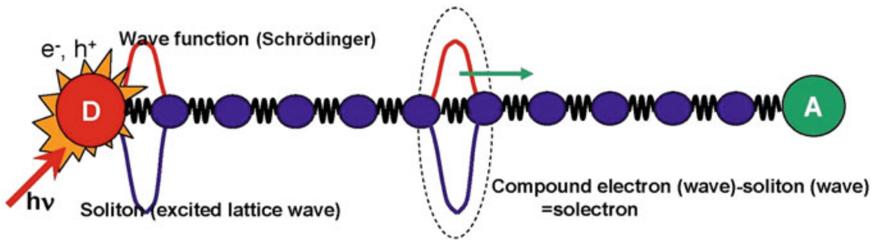


Fig. 1.1 Electron transfer from a donor (D) to an acceptor (A) along a 1d crystal lattice. The springs mimic either harmonic interactions or otherwise. In this text they are assumed to correspond to (anharmonic) Morse potentials. The figure also illustrates the electron-soliton bound state (solectron) formation. Depending on the material ways other than photoexcitation at the donor site could lead to the same consequences

introduced by Landau and Pekar [16, 29, 34]. Anharmonic, generally supersonic waves are naturally robust due to, e.g., a balance between nonlinearity and dispersion (or dissipation). In the following Sections we succinctly describe some of our findings and predictions for one-dimensional (1d) crystal lattices for which exact analytical and numerical results exist (Sect. 1.2) and, subsequently, for two-dimensional (2d) lattices for which only numerical results are available (Sects. 1.3, 1.4 and 1.5). Comments about theory and experiments are provided in Sect. 1.6 of this text.

1.2 Soliton Assisted Electron Transfer in 1d Lattices

Although the basic phenomenological theory exists [31] yet long range ET (beyond 20 Å) in biomolecules is an outstanding problem [23, 24]. Recent experiments by Barton and collaborators [38] with synthetic DNA show an apparent *ballistic* transport over 34 nm for which no theory exists. Let us see how we can address this question building upon our *solectron* concept (Fig. 1.1).

We consider the 1d-crystal lattice with *anharmonic* forces described by the following Hamiltonian

$$H_{lattice} = \sum_n \left\{ \frac{M v_n^2}{2} + D\left(1 - \exp\left[-B\left(x_n - x_{n-1} - \sigma\right)\right]\right)^2 \right\}, \qquad (1.1)$$

where x_n, v_n, M, D, B and σ denote, respectively, space lattice coordinates/sites, lattice particle/unit velocities, unit masses (all taken equal), the potential depth or dissociation energy of the Morse potential (akin to the 12–6 Lennard-Jones potential), lattice stiffness constant and interparticle equilibrium distance or initial lattice spacing. For our purpose here we introduce suitably rescaled relative lattice displacements, $q_n = B(x_n - n\sigma)$. Around the minimum of the potential well we can define $\omega_0 = \left(2DB^2/M\right)^{1/2}$ as the linear (harmonic) vibration frequency. For biomolecules like azurin $\omega_0 = 10^{13}\, s^{-1}$, $M \approx 100$ amu. Then for $D = 0.1\ eV$ $(D = 1.0\ eV)$ we can set $B = 2.3\,\text{Å}^{-1}$ $(B = 0.72\,\text{Å}^{-1})$ [46].

If an *excess* electron is added to the lattice we can take it in the tight binding approximation (TBA) and hence

$$H_{electron} = \sum_n E_n C_n^* C_n - \sum_n V_{n,n-1}\,(q_k)\left(C_n^* C_{n-1} + C_n C_{n-1}^*\right), \tag{1.2}$$

with n denoting the lattice site where the electron is placed (in probability density, $w_n = |C_n|^2$, $\sum_n w_n = 1$). We want to emphasize the significance of hopping in the transport process relative to effects due to onsite energy shifts and hence we assume $E_n = E_0$ for all n save those referring to D and A. The quantities $V_{n,n-1}$ belong to the transfer matrix or overlapping integrals. They depend on actual relative lattice displacements, and we can set [37]

$$V_{n,n-1} = V_0 \exp\left[-\alpha\,(q_n - q_{n-1})\right], \tag{1.3}$$

where V_0 and α account for the electron-lattice coupling strength. Accordingly, $\tau = V_0/\hbar\omega_0$ provides the ratio of the two dynamical time scales (electronic over mechanical/sound).

From (1.1)–(1.3) follow the equations of motion in suitable dimensionless form:

$$\begin{gathered}\frac{d^2 q_n}{dt^2} = \left[1 - e^{(q_n - q_{n+1})}\right] e^{(q_n - q_{n+1})} - \left[1 - e^{(q_{n-1} - q_n)}\right] e^{(q_{n-1} - q_n)} + \\ +2\alpha V\left[Re\left(C_{n+1} C_n^*\right) e^{\alpha(q_n - q_{n+1})} - Re\left(C_n C_{n-1}^*\right) e^{\alpha(q_{n-1} - q_n)}\right]\end{gathered} \tag{1.4}$$

$$\frac{dC_n}{dt} = i\tau\left[C_{n+1} e^{\alpha(q_n - q_{n+1})} + C_{n-1} e^{\alpha(q_{n-1} - q_n)}\right]. \tag{1.5}$$

It is worth recalling that if rather than the Morse potential (1) we use a similar potential introduced by Toda the lattice dynamic problem defined by Eq. (1.4) in the absence of the added electron ($\alpha = 0$) is exactly solvable [5, 39]. Thus we know analytical expressions for lattice motions and, moreover, for the thermodynamics/statistical mechanics (including specific heats, dynamic structure factor, etc.) of such 1d many-body problem. For the Morse potential (1) it has been numerically shown that no significant differences exist for lattice motions and other physical quantities [15, 36]. Temperature can be incorporated in the dynamics by adding to Eq. (1.4) Langevin sources by using an appropriate heat bath (delta-correlated Gaussian white noise) and using Einstein's relation between noise strength and temperature. To avoid redundancy we illustrate this point in Sect. 1.3.

The implementation of the scheme shown in Fig. 1.1 is one prediction with velocities in the sonic and supersonic range. Figure 1.2 illustrates the possibility using Eqs. (1.4)–(1.5) of extracting an electron placed in a potential well in the 1d Morse lattice by a generally supersonic soliton. For the geometry of Fig. 1.1 we can use it to

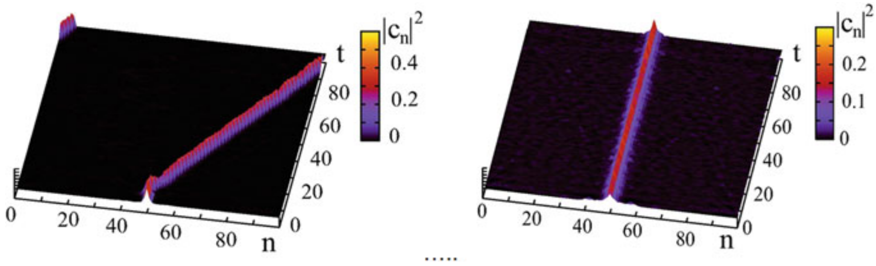


Fig. 1.2 Extraction of an electron from a potential well (a donor) and ballistic transport to an acceptor observed using the electron probability density$|C_n|^2$. *Left panel*: shallow well$|E| = 10$, extraction 100 %. *Right panel*: deep well$|E| = 18$, no extraction. Parameter values: $\alpha = 1.75$, $V_0 = 0.35$ and $\tau = 10$

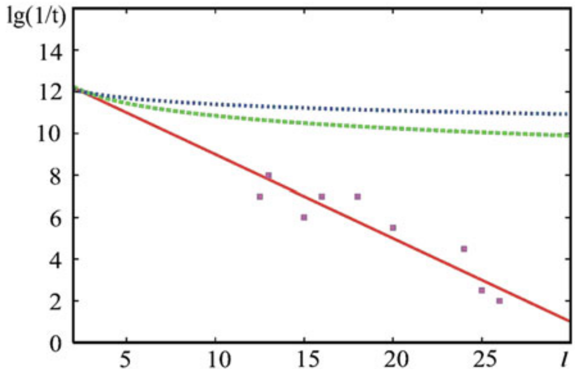


Fig. 1.3 Logarithm of reciprocal time lapse (in seconds) which an electron bound to a soliton needs to travel a distance$l$ (in Angstrom) for the geometry of Fig. 1.1. The upper dotted (*blue*) curve corresponds to a sound velocity of 17 Angstrom/ps, illustrating a ballistic transport. The second dotted (*green*) curve from above shows the reciprocal time needed if the electron hops stochastically between thermally excited solitons. The bottom solid line embraces data illustrating a tunneling process. The dots are reciprocal times measured for natural bio-molecules [23], [24]. The transfer times found for synthetic DNA are much shorter [38] bearing similarity to our model findings—upper dotted (blue) line—for solectron transfer

estimate the *ballistic* process time lapse to go from the donor to the acceptor. For the computation with a lattice of $N = 100$ units the well is assumed Gaussian of depth $|E|$ (in units of $\hbar\omega_0$) with $E < O$ localized at site 50. The soliton initially spans a few lattice sites (two or three) excited at site 40. If the well depth is shallow enough the extraction is ensured up to 100 % whereas if the well is too deep no extraction occurs. Needless to say extraction is possible with probability varying from zero to unity as the well depth is decreased. Time lapse from D to A is obtained by simply dividing length over soliton speed. Illustration is provided in Fig. 1.3 where "ℓ" (see Fig. 1.1) accounts for the distance travelled (in principle from D to an appropriately placed acceptor A). Comparison is provided between the *ballistic* case and other possibilities like *diffusion*-like transport with thermally (hence randomly) excited solitons [7] and *tunneling* transport [13].

1.3 Two-Dimensional Crystal Lattices

Recently, two groups of experimentalists have observed how an electron can "surf" on a suitably strong albeit linear, highly monochromatic sound wave (in *GaAs* layers at 300 mK). Sound demands lattice compressions and hence is accompanied by electric/polarization fields which for piezoelectric crystals integrate to macroscopic level. Our theoretical *solectron* approach targets sound-wave electron surfing due to nonlinear soliton excitations in 2d-anharmonic crystal lattice layers, with velocities ranging from supersonic to sub-sonic [5, 25, 30]. We do not pretend here to explain the *GaAs* experimental results. We simply wish to point out that appropriate sound waves in suitable nonlinear crystalline materials, could provide long range ET in 2d, with sonic or supersonic velocities for temperatures much higher than that so far achieved in experiments.

In 2d the Morse potential needs to be truncated to avoid overcounting lattice sites. Then using complex coordinates $Z = x + iy$, where x and y are Cartesian coordinates, the equations of motion replacing Eq. (1.4) are

$$\frac{d^2 Z_n}{dt^2} = \sum_k F_{nk}\left(|Z_{nk}|\right) z_{nk} + \left[-\gamma \frac{dZ_n}{dt} + \sqrt{2D_v}\left(\xi_{nx} + i\xi_{ny}\right)\right], \tag{1.6}$$

with $F_{nk}\left(|Z_{nk}|\right) = -\left(dV/dr\right)_{r=|Z_{nk}|}$, $z_{nk} = (Z_n - Z_k)/|Z_n - Z_k|$. In Eq. (1.6) we have incorporated thermal effects. The quantities γ (friction coefficient), D_v (diffusion coefficient) and the ξ_s (noise generators) characterize the Gaussian noise. $D_v = k_B T\gamma/M$ is Einstein's relation with k_B, Boltzmann constant.

To illustrate lattice motions we consider each lattice unit as a sphere representing the core electron Gaussian distribution at the corresponding site: $\rho(Z,t) = \sum_{|Z-Z_i(t)|<1.5} \exp\left(-\left|Z - Z_j(t)\right|^2/2\lambda^2\right)$ with λ a parameter. Thus overlapping of two such Gaussians permit to "detect" the expected "mechanical" compression of two lattice units as Fig. 1.4 illustrates [10, 11]. The evolution of the electron follows Eq. (1.5) for the 2d lattice geometry.

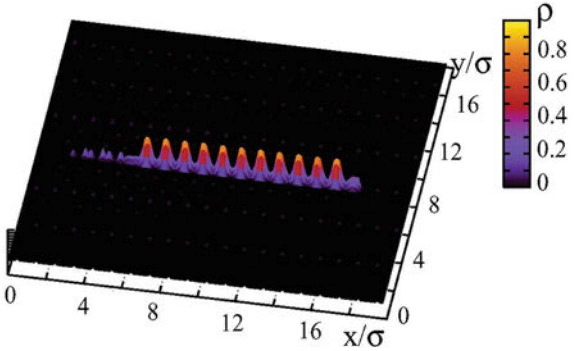


Fig. 1.4 Cumulative sequence of snapshots using $\rho(Z,t)$ to track a soliton running along the x-axis of a triangular lattice using Eqs. (1.4), (1.6). Parameter values: $B\sigma = 4$, $T = 0.001$

1.4 Two-Dimensional Crystal Lattices. Pauli's Master Equation Approach

Continuing with the 2d case, we now consider an alternative approach to using the Schrödinger Eq. (1.5). We shall consider how transport is achieved following Pauli's master equation approach [17]. Equation (1.2) is now considered with $V_{nn'}\left(Z_{n,n'}\right)$. The energy levels are taken in the *polarization* approximation $E_n = E_0 - \sum_n \left\{U_e h^4/\left[\left|Z_{n,n'}\right|^2 + h^2\right]^2\right\}$, where U_e is the electric potential strength and h defines the range of the electric field polarization interaction. Rather than relying on the Schrödinger description of the TBA we follow Pauli's master equation approach with transition probabilities

$$W_{n,n'} = \left(t_0^2/\hbar\right) \exp\left[-2\alpha\left|Z_{n,n'}\right|\right] E\left(n, n'; \beta\right), \qquad (1.7)$$

$$\frac{dw_n}{dt} = \sum \left[W_{nn'}w_{n'} - W_{n'n}w_n\right], \qquad (1.8)$$

where $E\left(n, n'; \beta\right) = 1$ if $E_n < E_{n'}$ and $E\left(n, n'; \beta\right) = \exp\left[-\beta\left(E_n - E_{n'}\right)\right]$ if $E_n > E_{n'}$, $\beta = 1/k_B T$. Equations. (1.7)–(1.8) are solved with Eq. (1.6) to obtain the electron probability density $w_n(t)$ neglecting the feedback of the electron on the lattice dynamics.

Figures 1.5 and 1.6 illustrate electron and solectron evolution along a 2d lattice. Figure 1.5 refers to electron taken alone while Fig. 1.6 illustrates how, after switching-on the electron-lattice interaction, the soliton from Fig. 1.4 is able to trap the electron from Fig. 1.5 and after forming the solectron transports charge along the lattice (see also [42]).

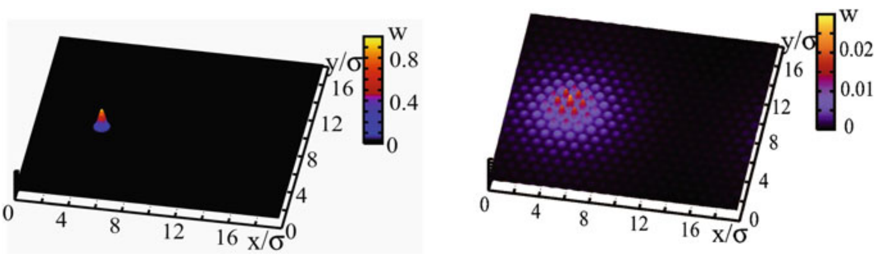


Fig. 1.5 An electron alone placed at a given lattice site (*left panel*). The quantity w here accounts for the probability density (otherwise $|C_n|^2$ in Fig. (1.3). As time progresses the electron spreads over the slightly heated lattice ($T = 0.002D$) following Pauli's equation from the initial condition (*left panel*) to a subsequent time instant (*right panel*)

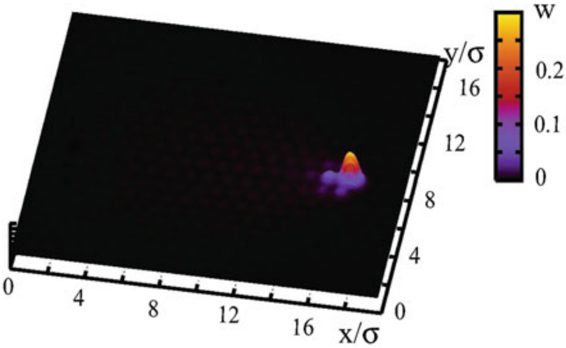


Fig. 1.6 Solectron formation and eventual evolution when the electron-phonon (here electron-soliton) interaction is switched-on ($\alpha \neq 0$ in the corresponding to Eq. (1.4) for the 2d case). We see that the electron is trapped by a soliton (like that of Fig. 1.4) thus forming the solectron which transfers the electron probability density without spreading at variance with the result illustrated in Fig. 1.5 (*right panel*)

1.5 Percolation and Other Features in 2d-Lattices

Solitons can be excited in a crystal lattice by several actions. One is to add finite momentum to a group of nearby lattice units, another is by heating the crystal alltogether. Then one expects quite many excitations including phonons and solitons randomly appearing along the 2d lattice and having finite life times thus leaving finite-length traces. Figure 1.7 illustrates thermal excitations leading to spots of instant electron density $n_e\,(x,\,y;\,t)$ due to higher than equilibrium electric/polarization field maxima. Here in the simplest Boltzmann approximation

$$n_{el}\,(Z/x,\,y;\,t) = \left\{\exp - \left[U\,(Z,t)/k_B T\right]/n^0_{el}\right\}, \tag{1.9}$$

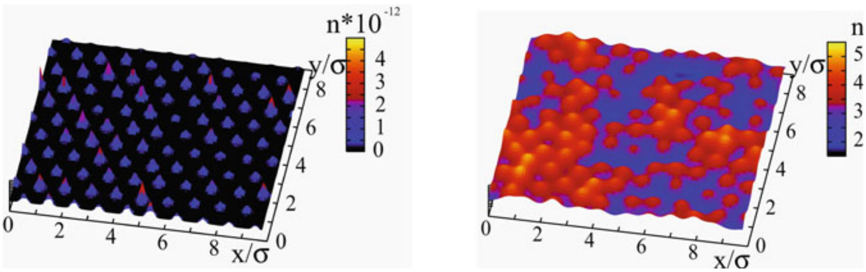


Fig. 1.7 Towards percolation. Instantaneous space distribution of electron probability density $n\,(x,\,y)$ associated to lattice solitons (sound) in a triangular Morse lattice ($N = 100$) at, respectively, low ($T = 0.02$ D) (*left panel*) and high ($T = 0.4$ D) (*right panel*) temperatures. The latter exhibits an almost percolating path. Parameter values: $B\,\sigma = 3$, $U_e = 0.4$D, $h = 0.7\,\sigma$

With n_{el}^0 the normalizing factor, $n_{el}^0 = \int \exp\left[-U(Z,t)/k_B T\right] dZ$.

Only at temperatures high enough one expects a distribution of "local" spots permitting in kind of zig-zag the occurrence of an "infinite" path thus percolating from side to side of the 2d lattice [6, 8]. Indeed by increasing temperature one increases the significance as well as the "density" of soliton excitations/traces. If percolation does occur by adding an excess electron and playing with an external field we have a novel way of one-sided electric conduction mediated by the solitons. We have just explored this possibility but have not yet been able to draw conclusions about the scaling laws of the process. On the other hand since percolation is expected as a second-order phase transition it seems worth investigating the possible connection with the pseudo-gap transition observed in such superconducting materials as cuprates.

1.6 Concluding Remarks

We have illustrated how lattice solitons arising from *finite* amplitude compression-expansion longitudinal motions bring sound and also create electric polarization fields [29]. The latter are able to trap charges and provide long-range ET in a wide range of temperatures (up to e.g., 300K for bio-molecules). Such "sound" waves could exhibit subsonic, sonic or supersonic velocity, whose actual value depends on the strength of the electron-phonon/soliton interaction. Noteworthy is that such interaction and subsequent electric transport, in the most general case, embraces a genuine *polaron* effect [16, 34] and also a genuine soliton/solectron effect [3]. For piezoelectric materials like *GaAs* that sound waves can transport electrons there is now experimental evidence [27, 32]. This was achieved by means of strong albeit linear/infinitesimal, highly monochromatic waves appropriately creating the electric/polarization field that due to the specificity of the crystal symmetry and other features integrate to macroscopic level. These experiments done at 300 mK due to quantum limitations imposed to the set-up provide hope for similar long-range ET at "high" temperatures. Indeed the limitations are only due to the electron entry and exit/detector gates. The *solectron* theory predicts such a possibility in appropriate non-linearly elastic crystal materials capable of sustaining lattice solitons. Recent experiments using synthetic DNA [38] show a kind of *ballistic* ET over 34 nm which as Fig. 1.3 illustrates bears similarity with a prediction of our *solectron* theory [42]. In 2d crystal lattices the *solectron* theory predicts the possibility of percolation as a way of long range charge transport when the material is heated up to the range of robustness/stability of lattice solitons, as Fig. 1.7 illustrates. Work remains to be carried out to assess the corresponding percolation scaling laws.

Finally, we have recently shown that the *solectron* theory offers a new way of electron *pairing* by having two electrons strongly correlated (both in real space and in momentum space with due account of Pauli's exclusion principle and Coulomb repulsion using Hubbard's local approximation) due to their trapping by lattice solitons [2, 26, 41, 44, 47–49]. This feature shows the quite significant role played by the lattice dynamics well beyond the role played in the formation of Cooper pairs (in momentum space) underlying the BCS theory [14] or in the bipolaron theory [1] and much in the spirit of Fröhlich approach to the problem unfortunately using a harmonic lattice Hamiltonian at a time before (lattice) solitons were known [19–22, 33, 39, 50]; see also[51]. Incidentally, Einstein [18] was the first who used the concept of molecular conduction chains trying to understand superconduction. Thus it is reasonable to expect that a soliton-mediated Bose-Einstein condensation could take place in appropriate 2d *anharmonic* crystal lattices well above absolute zero. This is yet to be shown.

Acknowledgments The authors are grateful to A. S. Alexandrov, E. Brändäs, L. Brizhik, L. Cruzeiro, F. de Moura, J. Feder, D. Hennig, R. Lima, R. Miranda, R. McNeil, D. Newns, G. Röpke and G. Vinogradov for enlightening discussions. This research has been sponsored by the Spanish Ministerio de Ciencia e Innovación, under Grants EXPLORA FIS2009-06585 and MAT2011-26221.

References

1. Alexandrov AS, Alexandrov AS (eds) (2007) Polarons in advanced materials. Springer, Dordrecht
2. Brizhik L, Chetverikov AP, Ebeling W, Röpke G, Velarde MG (2012) Electron pairing and Coulomb repulsion in one-dimensional anharmonic lattices (in preparation)
3. Cantu Ros OG, Cruzeiro L, Velarde MG, Ebeling W (2011) On the possibility of electric transport mediated by long living intrinsic localized solectron modes. Eur Phys J B 80:545–554
4. Chetverikov AP, Ebeling W, Velarde MG (2006a) Nonlinear excitations and electric transport in dissipative Morse-Toda lattices. Eur Phys J B 51:87–99
5. Chetverikov AP, Ebeling W, Velarde MG (2006b) Disipative solitons and complex currents in active lattices. Int J Bifurcation Chaos 16:1613–1632
6. Chetverikov AP, Ebeling W, Velarde MG (2009) Local electron distributions and diffusion in anharmonic lattices mediated by thermally excited solitons. Eur Phys J B 70:217–227
7. Chetverikov AP, Ebeling W, Velarde MG (2010) Thermal solitons and solectrons in nonlinear conducting chains. Int J Quantum Chem 110:46–61
8. Chetverikov AP, Ebeling W, Velarde MG (2011a) Soliton-like excitations and solectrons in two dimensional nonlinear lattices. Eur Phys J B 80:137–145
9. Chetverikov AP, Ebeling W, Röpke G, Velarde MG (2011b) Hopping transport and stochastic dynamics of electrons in plasma layers. Contrib Plasma Phys 51:814–829
10. Chetverikov AP, Ebeling W, Velarde MG (2011c) Localized nonlinear, soliton-like waves in two-dimensional anharmonic lattices. Wave Motion 48:753–760
11. Chetverikov AP, Ebeling W, Velarde MG (2011d) Propeties of nano-scale soliton-like excitations in two-dimensional lattice layers. Physica D 240:1954–1959
12. Chetverikov AP, Ebeling W, Velarde, MG (2012a) Controlling electron transfer at the nanoscale by soliton-like excitations in two-dimensional, nonlinear lattices (in preparation)
13. Chetverikov AP, Cruzeiro L, Ebeling W, Velarde, MG (2012b) Studies on electron transfer and tunneling in inhomogeneous, nonlinear lattices (in preparation)
14. Cooper LN, Feldman D (2011) BCS: 50 years. World Scientific, London
15. Dancz J, Rice SA (1977) Large amplitude vibrational motion in a one dimensional chain: coherent state representation. J Chem Phys 67:1418–1426
16. Devreese JTL, Devreese JTL (eds) (1972) Polarons in ionic crystals and polar semiconductors. North Holland, Amsterdam

17. Ebeling W, Velarde MG, Chetverikov AP (2009) Bound states of electrons with soliton-like excitations in thermal systems—adiabatic approximations. Condensed Mat Phys 12:633–645
18. Einstein A (1922) Theoretische Bemerkungen zur Supraleitung der Metalle, in Het Natuurkundig Laboratorium der Rijksuniversiteit te Leiden, in de Jaren, 429, 1904–1922. Eduard Ijdo, Leiden
19. Fröhlich H (1950) Theory of the superconducting state. I. The ground state at the absolute zero of temperature. Phys Rev 79:845–856
20. Fröhlich H (1952) Interaction of electrons with lattice vibrations. Proc Roy Soc London 215:291–298
21. Fröhlich H (1954a) On the theory of superconductivity: the one-dimensional case. Proc Roy Soc London A223:296–305
22. Fröhlich H (1954b) Electrons in lattice fields. Adv Phys 3:325–361
23. Gray HB, Winkler JR (2003) Electron tunneling through proteins. Quart Rev Biophys 36:341–372
24. Gray HB, Winkler JR (2005) Long-range electron transfer. PNAS 102:3534–3539
25. Hennig D, Neissner C, Velarde MG and Ebeling W (2006) Effect of anharmonicity on charge transport in hydrogen-bonded systems. Phys Rev B 73, 024306–1-10
26. Hennig D, Velarde MG, Ebeling W, Chetverikov AP (2008) Compounds of paired electrons and lattice solitons moving with supersonic velocity. Phys Rev E 78, 066606–1-9
27. Hermelin S, Takada S, Yamamoto M, Tarucha S, Wieck AD, Saminadayar L, Bäuerle C, Meunier T (2011) Electrons surfing on a sound wave as a platform for quantum optics with flying electrons. Nature 477:435–438
28. Kaganov MI, Lifshits IM (1979) Quasiparticles. Mir, Moscow
29. Landau LD (1933) Electron motion in crystal lattices. Phys Z Sowjetunion 3:664
30. Makarov VA, Velarde MG, Chetverikov AP, Ebeling W (2006) Anharmonicity and its significance to non-Ohmic electric conduction. Phys Rev E 73, 066626–1-12
31. Marcus RA, Sutin N (1985) Electron transfer in chemistry and biology, Biochim. Biophys. Acta 811, 265–322. See also (1999). Adv Chem Phys 106:1–6
32. McNeil RPG, Kataoka M, Ford CJB, Barnes CHW, Anderson D, Jones GAC, Farrer I, Ritchie DA (2011) On-demand single-electron transfer between distant quantum dots. Nature 477:439–442
33. Nayanov VI (1986) Surface acoustic cnoidal waves and solitons in a $LiNbO_3$ -(SiO film) structure. JETP Lett 44:314–317
34. Pekar SI (1954) Untersuchungen über die Elektronentheorie. Akademie Verlag, Berlin
35. Pomeau Y, Le Berre M (2007) Optical soliton as quantum objects. Chaos 17:037118
36. Rolfe TJ, Rice SA, Dancz J (1979) A numerical study of large amplitude motion on a chain of coupled nonlinear oscillators. J Chem Phys 70:26–33
37. Slater JC (1974) Quantum theory of molecules and solids, vol 4. McGraw-Hill, New York
38. Slinker JD, Muren NB, Renfrew SE, Barton JK (2011) DNA charge transport over 34 nm. Nat Chem 3:228–233
39. Toda M (1989) Theory of nonlinear lattices, 2nd ed. Springer-Verlag, Berlin
40. Velarde MG, Ebeling W, Chetverikov AP (2005) On the possibility of electric conduction mediated by dissipative solitons. Int J Bifurcation Chaos 15:245–251
41. Velarde MG, Ebeling W, Hennig D, Neissner C (2006) On soliton-mediated fast electric conduction in a nonlinear lattice with Morse interactions. Int J Bifurcation Chaos 16:1035–1039
42. Velarde MG, Ebeling W, Chetverikov AP, Hennig D (2008a) Electron trapping by solitons. Classical versus quantum mechanical approach. Int J Bifurcation Chaos 18:521–526
43. Velarde MG, Ebeling W, Chetverikov AP (2008b) Thermal solitons and solectrons in 1d anharmonic lattices up to physiological temperatures. Int J Bifurcation Chaos 18:3815–3823
44. Velarde MG, Neissner C (2008c) Soliton-mediated electron pairing. Int J Bifurcation Chaos 18:885–890
45. Velarde MG (2010a) From polaron to solectron: the addition of nonlinear elasticity to quantum mechanics and its possible effect upon electric transport. J Computat Applied Maths 233:1432–1445

46. Velarde MG, Chetverikov AP, Ebeling W, Hennig D, Kozak JJ (2010b) On the mathematical modeling of soliton-mediated long-range electron transfer. Int J Bifurcation Chaos 20:185–194
47. Velarde MG, Ebeling W, Chetverikov AP (2011a) Numerical evidence of soliton-mediated electron pairing in heated anharmonic crystal lattices. Int J Bifurcation Chaos 21:1595–1600
48. Velarde MG, Brizhik L, Chetverikov AP, Cruzeiro L, Ebeling W, Röpke G (2011b) Electron pairing in one-dimensional anharmonic crystal lattices. Int. J. Quantum Chem, doi:10.1002/qua.23008
49. Velarde MG, Brizhik L, Chetverikov AP, Cruzeiro L, Ebeling W, Röpke G (2011c) Quartic lattice interactions, soliton-like excitations, and electron pairing in one-dimensional anharmonic crystals. Int J Quantum Chem, doi:10.1002/qua.23282
50. Zabusky NJ, Kruskal MD (1965) Interaction of solitons in a collisionless plasma and the recurrence of initial states. Phys Rev Lett 15:57–62
51. Zabusky NJ (2005) Fermi-Pasta-Ulam, solitons and the fabric of nonlinear and computational science: history, synergetics, and visiometrics. Chaos 15:015102

Chapter 2
Generating the Mass of Particles from Extended Theories of Gravity

Salvatore Capozziello and Mariafelicia De Laurentis

Abstract A geometrical approach to produce the mass of particles is derived. The results could be suitably tested at LHC. Starting from a 5D unification scheme, we show that all the known interactions could be induced by a symmetry breaking of the non-unitary $GL(4)$-group of diffeomorphisms. The further gravitational degrees of freedom, emerging from the reduction mechanism in 4D, eliminate the hierarchy problem generating a cut-off comparable with electroweak scales.

2.1 Introduction

The *Standard Model of Particles* can be considered a successful relativistic quantum field theory both from particle physics and group theory points of view. Technically, it is a non-Abelian gauge theory (a Yang-Mills theory) associated to the tensor product of the internal symmetry groups $SU(3) \times SU(2) \times U(1)$, where the $SU(3)$ color symmetry for quantum chromodynamics is treated as exact, whereas the $SU(2) \times U(1)$ symmetry, responsible for the electro-weak gauge fields, is considered spontaneously broken. So far, as we know, there are four fundamental forces in Nature; namely, electromagnetic, weak, strong and gravitational forces. The Standard Model well represents the first three, but not the gravitational interaction. On the other hand, General Relativity (GR) is a geometric theory of the gravitational field which is described by the metric tensor $g_{\mu\nu}$ defined on pseudo-Riemannian space-times. The Einstein field equations are nonlinear and have to be satisfied by the metric tensor. This nonlinearity is indeed a source of difficulty in quantization of GR. Since the Standard Model is a gauge theory where all the fields mediating the interactions are represented by gauge potentials, the question is why the fields mediating the gravitational interaction are different from those of the other fundamental forces. It

S. Capozziello (✉) · M. De Laurentis
Dipartimento di Scienze Fisiche, Università di Napoli "Federico II" and INFN Sezione di Napoli, Complesso Universitario di Monte Sant'Angelo, Edificio G, Via Cinthia, I-80126 Napoli, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_2,

is reasonable to expect that there may be a gauge theory in which the gravitational fields stand on the same footing as those of the other fields. As it is well-known, this expectation has prompted a re-examination of GR from the point of view of gauge theories. While the gauge groups involved in the Standard Model are all internal symmetry groups, the gauge groups in GR is associated to external space-time symmetries. Therefore, the gauge theory of gravity cannot be dealt under the standard of the usual Yang-Mills theories.

Nevertheless, the idea of an unification theory, capable of describing all the fundamental interactions of physics under the same standard, has been one of the main issues of modern physics, starting from the early efforts by Einstein, Weyl, Kaluza and Klein until the most recent approaches. In any case, the large number of ideas, up to now proposed, results unsuccessful due to several reasons: the technical difficulties connected with the lack of a unitary mathematical description of all the interactions; the huge number of parameters introduced to "build up" the unified theory and the fact that most of them cannot be observed neither at laboratory nor at astrophysical (or cosmological) scales; the very wide (and several times questionable since not-testable) number of extra-dimensions requested by several approaches. Due to this situation, it seems that unification is a useful (and aesthetic) paradigm, but far to be achieved, if the trend is continuing to unify interactions by adding and adding new particles and new parameters (e.g. dark matter forest).

A different approach could be to consider the very essential physical quantities and try to achieve unification without any *ad hoc* new ingredients. This approach can be pursued starting from straightforward considerations which lead to reconsider modern physics under a sort of economic issue aimed to unify forces without adding new parameters. A prominent role in this view deserves conservation laws and symmetries.

As a general remark, the Noether Theorem states that, for every conservation law of Nature, a symmetry *must* exist. This leads to a fundamental result also from a mathematical point of view since the presence of symmetries technically reduces dynamics (i.e. gives rise to first integrals of motion) and, in several cases, allows to get the general solution. With these considerations in mind, we can try to change our point of view and investigate what will be the consequences of the absolute validity of conservation laws without introducing any arbitrary symmetry breaking [1].

In order to see what happens as soon as we ask for the absolute validity of conservation laws, we could take into account the Bianchi identities. Such geometrical identities work in every covariant field theory (e.g. Electromagnetism or GR) and can be read as equations of motion also in a fiber bundle approach [2]. It is possible to show that, the absolute validity of conservation laws, intrinsically contains symmetric dynamics; moreover, reducing dynamics from 5D to 4D, it gives rise to the physical quantities characterizing particles as the mass [3].

The *minimal* ingredient which we require is the fact that a 5-dimensional, singularity free space, where conservation laws are always and absolutely *conserved*, has to be defined. Specifically, in such a space, Bianchi identities are asked to be always valid and, moreover, the process of reduction to 4D-space *generates* the mass spectra of particles. In this sense, a dynamical unification scheme will be achieved where a fifth dimension has the physical meaning of *inducing the mass of particles*. In other words, "effective" scalar fields coming from dimensional reduction mechanisms are related to the $GL(4)$-group of diffeomorphisms. In this sense, we do not need any spontaneous symmetry breaking but just a self-consistent way to classify space-time deformations and reductions as "gauge bosons" [3].

2.2 The 5D-space and the Reduction to 4D-dynamics

Let us start with a 5D-variational principle with

$$\delta \int d^{(5)}x\sqrt{-g^{(5)}}\left[{}^{(5)}R + \lambda(g_{44} - \varepsilon\Phi^2)\right] = 0, \tag{2.1}$$

where λ is a Lagrange multiplier, Φ a scalar field and $\varepsilon = \pm 1$. This approach is completely general and used in theoretical physics when we want to put in evidence some specific feature [4]. In this case, we need it in order to derive the physical gauge for the 5D-metric. We can write the metric as

$$dS^2 = g_{AB}dx^A dx^B = g_{\alpha\beta}dx^\alpha dx^\beta + g_{44}\left(dx^4\right)^2 = g_{\alpha\beta}dx^\alpha dx^\beta + \varepsilon\Phi^2\left(dx^4\right)^2, \tag{2.2}$$

from which we obtain directly particle-like solutions ($\varepsilon = -1$) or wave-like solutions ($\varepsilon = +1$) in the 4D-reduction procedure. The standard signature of 4D-component of the metric is $(+ - - -)$ and $\alpha, \beta = 0, 1, 2, 3$. Furthermore, the 5D-metric can be written in a Kaluza-Klein fashion as the matrix

$$g_{AB} = \begin{pmatrix} g_{\alpha\beta} & 0 \\ 0 & \varepsilon\Phi^2 \end{pmatrix}, \tag{2.3}$$

and the 5D-curvature Ricci tensor is

$${}^{(5)}R_{\alpha\beta} = R_{\alpha\beta} - \frac{\Phi_{,\alpha;\beta}}{\Phi} + \frac{\varepsilon}{2\Phi^2}\left(\frac{\Phi_{,4}g_{\alpha\beta,4}}{\Phi} - g_{\alpha\beta,44} + g^{\lambda\mu}g_{\alpha\lambda,4}g_{\beta\mu,4} - \frac{g^{\mu\nu}g_{\mu\nu,4}g_{\alpha\beta,4}}{2}\right), \tag{2.4}$$

where $R_{\alpha\beta}$ is the 4D-Ricci tensor. The expressions for ${}^{(5)}R_{44}$ and ${}^{(5)}R_{4\alpha}$ can be analogously derived. After the projection from 5D to 4D, $g_{\alpha\beta}$, derived from g_{AB}, no longer explicitly depends on x^4. From Eq. 2.4, a useful expression for the Ricci scalar can be derived:

$${}^{(5)}R = R - \frac{1}{\Phi}\Box\Phi \tag{2.5}$$

where the dependence on ε is explicitly disappeared and $\Box$ is the 4D-d'Alembert operator. The action in Eq. 2.1 can be recast in a 4D-reduced Brans-Dicke form

$$\mathcal{A} = -\frac{1}{16\pi G_N}\int d^4x\sqrt{-g}\,[\Phi R + \mathcal{L}_\Phi], \tag{2.6}$$

where the Newton constant is given by

$$G_N = \frac{^{(5)}G}{2\pi l} \tag{2.7}$$

with l a characteristic length in 5D which can be related to a suitable Compton length. Defining a generic function of a 4D-scalar field ϕ as

$$-\frac{\Phi}{16\pi G_N} = F(\phi)\,, \tag{2.8}$$

we get, in 4D, a general action in which gravity is non-minimally coupled to a scalar field, that is

$$\mathcal{A} = \int_{\mathcal{M}} d^4x\sqrt{-g}\left[F(\phi)\,R + \frac{1}{2}g^{\mu\nu}\phi_{;\mu}\phi_{;\nu} - V(\phi)\right] + \int_{\partial\mathcal{M}} d^3x\sqrt{-b}K, \tag{2.9}$$

where the form and the role of $V(\phi)$ are still general. The second integral is a boundary term where $K \equiv h^{ij}\,K_{ij}$ is the trace of the extrinsic curvature tensor K_{ij} of the hypersurface $\partial\mathcal{M}$ which is embedded in the 4D-manifold $\mathcal{M}$; b is the metric determinant of the 3D-manifold. The Einstein field equations can be derived by varying with respect to the 4D-metric $g_{\mu\nu}$

$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = \tilde{T}_{\mu\nu}\,, \tag{2.10}$$

where

$$\tilde{T}_{\mu\nu} = \frac{1}{F(\phi)}\left\{-\frac{1}{2}\phi_{;\mu}\phi_{;\nu} + \frac{1}{4}g_{\mu\nu}\phi_{;\alpha}\phi^{;\alpha} - \frac{1}{2}g_{\mu\nu}V(\phi) - g_{\mu\nu}\Box F(\phi) + F(\phi)_{;\mu\nu}\right\} \tag{2.11}$$

is the effective stress–energy tensor containing the non-minimal coupling contributions, the kinetic terms and the potential of the scalar field ϕ. In the case in which $F(\phi)$ is a constant F_0 (in our units, $F_0 = -1/(16\pi G_N)$), we get the stress–energy tensor of a scalar field minimally coupled to gravity, that is

$$T_{\mu\nu} = \phi_{;\mu}\phi_{;\nu} - \frac{1}{2}g_{\mu\nu}\phi_{;\alpha}\phi^{;\alpha} + g_{\mu\nu}V(\phi)\,. \tag{2.12}$$

By varying with respect to ϕ, we get the 4D-Klein–Gordon equation

$$\Box\phi - RF'(\phi) + V'(\phi) = 0 \tag{2.13}$$

where $F'(\phi) = dF(\phi)/d\phi$ and $V'(\phi) = dV(\phi)/d\phi$. It is possible to show that Eq. (2.13) is nothing else but the contracted Bianchi identity. This feature shows that the effective stress–energy tensor at right hand side of (2.10) is a zero–divergence tensor and this fact is fully compatible with Einstein theory of gravity also if we started from a 5D-space. Specifically, the reduction procedure, which we have used, preserves the standard features of GR since we are in the realm of the conformal-affine structure [3].

In order to give a physical meaning to the fifth dimension, let us recast the above Klein-Gordon Eq. (2.13) as

$$\left(\Box + m_{eff}^2\right)\phi = 0, \tag{2.14}$$

where

$$m_{eff}^2 = \left[V'(\phi) - RF'(\phi)\right]\phi^{-1}, \tag{2.15}$$

is the effective mass, i.e. a function of ϕ, where self-gravity contributions, $RF'(\phi)$, and scalar field self-interactions, $V'(\phi)$, are taken into account. In any quantum field theory formulated on curved space-times, these contributions, at one-loop level, have the same "weight" [5]. This toy model shows that a "natural" way to generate particle masses can be achieved starting from a 5D picture. In other words, the concept of *mass* can be derived from a very geometrical viewpoint.

2.3 Massive Gravitational States and the Induced Symmetry Breaking

The above results could be interesting to investigate quantum gravity effects and symmetry breaking in the range between GeV and TeV scales. Such scales are actually investigated by the today running experiments at LHC. It is important to stress that any ultra-violet model of gravity (e.g. at TeV scales) have to explain also the observed weakness of gravitational effects at largest (infra-red) scales. This means that massless (or quasi-massless) modes have to be considered in any case.

The above 5D-action is an example of higher dimensional action where the effective gravitational energy scale (Planck scale) can be "rescaled" according to Eqs. (2.7) and (2.8). In terms of mass, being $M_p^2 = \frac{c\hbar}{G_N}$ the constraint coming from the ultra-violet limit of the theory (10^{19} GeV), we can set $M_p^2 = M_\sharp^{D-2}V_{D-4}$, where V_{D-4} is the "volume" coming from the extra dimension. It is easy to see that V_{D-4}, in the 5D case, is related to the fifth component of Φ. $M_\sharp$ is a cut-off mass that becomes relevant as soon as the Lorentz invariance is violated. Such a scale could be of TeV order. As we have shown, it is quite natural to obtain effective theories containing scalar fields of gravitational origin. In this sense, $M_\sharp$ is the result of a dimensional reduction. To be more explicit, the 4D dynamics is led by the effective potential $V(\phi)$ and the non-minimal coupling $F(\phi)$. Such functions could be experimentally tested since related to massive states. In particular, the effective model, produced by the reduction mechanism from 5D to 4D, can be chosen as

$$\mathcal{A} = \int d^4x\sqrt{-g}\left[-\frac{\phi^2}{2}R + \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V\right] \tag{2.16}$$

plus contributions of ordinary matter terms. The potential for ϕ can be assumed as

$$V(\phi) = \frac{M_\sharp^2}{2}\phi^2 + \frac{\lambda}{4}\phi^4\,, \tag{2.17}$$

where massive and self-interaction terms are present. This is the standard choice of quantum field theory which perfectly fits with the arguments of dimensional reduction. Let us recall again that the scalar field ϕ is not put *by hand* into dynamics but it is given by the extra degrees of freedom of gravitational field generated by the reduction process in 4D. It is easy to derive the vacuum expectation value of ϕ, being

$$M_\sharp^2 = 2\lambda M_p^2, \tag{2.18}$$

which is a fundamental scale of the theory. Such a scale can be confronted with the Higgs vacuum expectation value which is 246 GeV and then with the *hierarchy problem*. If $M_\sharp$ is larger than the Higgs mass, the problem is obviously circumvented. It is important to recall that hierarchy problem occurs when couplings and masses of effective theories are very different than the parameters measured by experiments. This happen since measured parameters are related to the fundamental parameters by renormalization and fine cancellations between fundamental quantities and quantum corrections are necessary. The hierarchy problem is essentially a fine-tuning problem. In particle physics, the question is why the weak force is stronger and stronger than gravity. Both of these forces involve constants of Nature: Fermi's constant for the weak force and Newton's constant for gravity. From the Standard Model, it appears that Fermi's constant is unnaturally large and should be closer to Newton's constant.

Technically, the question is why the Higgs boson is so much lighter than the Planck mass (or the grand unification energy). In fact, researchers are searching for Higgs masses ranging from 115 up to 350 GeV with different selected decay channels from $b\bar{b}$ to $t\bar{t}$ (see for example [6] and references therein). One would expect that the large quantum contributions to the square of the Higgs boson mass would inevitably make the mass huge, comparable to the scale at which new physics appears, unless there is an incredible fine-tuning cancellation between the quadratic radiative corrections and the bare mass. With this state of art, the problem cannot be formulated in the context of the Standard Model where the Higgs mass cannot be calculated. In a sense, the problem is solvable if, in a given effective theory of particles, where the Higgs boson mass is calculable, there are no fine-tunings. If one accepts the *big-desert* assumption and the existence of a hierarchy problem, some new mechanism (at Higgs scale) becomes necessary to avoid fine-tunings.

The model which we are discussing contains a "running" scale that could avoid to set precisely the Higgs scale. If the mass of the field ϕ is in TeV region, there is no hierarchy problem being ϕ a gravitational scale. In this case, the Standard Model holds up plus an extended gravitational sector derived from the fifth dimension.

In other words, the Planck scale can be dynamically derived from the vacuum expectation value of ϕ. In some sense, our model, in its low energy realization, works like the model proposed by Antoniadis et al. [7]. The Planck scale can be recovered, as soon as the coupling λ is of the order 10^{-31}. Action (2.16) is an effective model valid up to a cutoff scale of a few $M_\sharp \sim$ TeV (see also [8]). The tiny value of λ, coming from the extra dimension, allows the presence of physical (quasi-) massless gravitons with very large interaction lengths [9].

Also the string theory limit corresponds to a large scalar field vacuum expectation value at TeV [7]. It is important to stress that, by a conformal transformation from the Jordan frame to the Einstein frame, the Planck scale is decoupled from the vacuum expectation of the scalar field ϕ. However the scalar field redefinition has to preserve the vacuum of the theory. Besides, the gauge couplings and masses depend on the vacuum expectation value of ϕ and are dynamically determined. This means that Standard Model and Einstein Gravity (in the conformal-affine sense [3]) could be recovered *without the hierarchy problem*. As discussed in [10], it is possible to show that the operators generated by the self-interaction of the scalar field are of the form

$$\frac{1}{M_\sharp^{N-4}}\lambda^{\frac{N}{2}}\phi^N \tag{2.19}$$

and they are always suppressed by the small parameter λ and do not destabilize the potential of the theory. This result holds also for perturbative corrections coming from quantum gravity.

Considering again the problem of mass generation, one can assume that particles of Standard Model have sizes related to the cut off, that is $M_\sharp^{-1}$, and their collisions could lead to the formation of bound states as in [7]. Potentially, such a phenomenon could mimic the decay of semi-classical quantum black holes and, at lower energies, it could be useful to investigate substructures of StandardModel. This means that we should expect some strong scattering effects in the TeV region involving the coupling of ϕ to the Standard Model fields. The "signature" of this phenomenon could lead to polarization effects of the particle beam. Furthermore the strong dynamics derived from the phenomenon could resemble compositeness as discussed in [11]. Furthermore, bounds on the production of mini-black holes can be derived from astroparticle physics. In [12] a bound on the cross-section is

$$\sigma_{\nu N\to BH+X} < \frac{0.5}{\mathrm{TeV}^2}. \tag{2.20}$$

Assuming, in our case, the cross-section $\sigma = M_\sharp^{-2}$, we get a bound of TeV order. If the fundamental scale of our theory is of this order, strong scattering processes at

LHC would have the cross-section

$$\sigma_{(pp\rightarrow grav.ghosts+X)} \sim 1 \times 10^{7}\text{fb} \tag{2.21}$$

and would dominate the cross-sections expected from the Standard Model. In this case, the Higgs boson could not be detected and no hierarchy problem would be present.

In summary, Higgs mechanism is an approach that allows: (i) to generate the masses of electroweak gauge bosons; (ii) to preserve the perturbative unitarity of the S-matrix; (iii) to preserve the renormalizability of the theory. The masses of the electroweak bosons can be written in a gauge invariant form using either the non-linear sigma model [14] or a gauge invariant formulation of the electroweak bosons. However if there is no propagating Higgs boson, quantum field amplitudes describing modes of the electroweak bosons grow too fast violating the unitarity around TeV scales [16]. There are several ways in which unitarity could be restored but the Standard Model without a Higgs boson is non-renormalizable at perturbative level.

A possibility is that the weak interactions become strongly coupled at TeV scales and then the related gauge theory becomes unitary at non-perturbative level. Another possibility for models without a Higgs boson consists in introducing weakly coupled new particles to delay the unitarity problem into the multi TeV regime where the UV limit of the Standard Model is expected to become relevant. In [17], it is proposed that, as black holes in gravitational scattering, classical objects could form in the scattering of longitudinal W-bosons leading to unitary scattering amplitude.

These ideas are very intriguing and show several features of electroweak interactions. First of all, the Higgs mechanism is strictly necessary to generatemasses for the electroweak bosons. Beside, some mechanisms can be unitary but not renormalizable or vice-versa. In summary, the paradigm is that three different criteria should be fulfilled: (i) gauge invariant generation of masses of electroweak bosons, (ii) perturbative unitarity; (iii) renormalizability of the theory.

Here we have proposed an alternative approach, based on Extended Theories of Gravity deduced from a 5D- manifold, where the Standard Model is fully recovered enlarging the gravitational sector but avoiding the Higgs boson and the hierarchy problem.

It is important to point out that, in both the non-linear sigma model and in gauge invariant formulation of Standard Model, it is possible to define an action in terms of an expansion in the scale of the electroweak interactions v. The action can be written as

$$\mathcal{A} = \mathcal{A}_{SMw/oHiggs} + \int d^4x \sum_i \frac{C_i}{v^N} O_i^{4+N}, \tag{2.22}$$

where O_i^{4+N} are operators compatible with the symmetries of the model. The electroweak bosons are gauge invariant fields defined by

$$\underline{W}^i_\mu = \frac{i}{2g} \mathrm{Tr}\, \Omega^\dagger \overleftrightarrow{D}_\mu \Omega \tau^i \,, \tag{2.23}$$

with and $D_\mu = \partial_\mu - igB_\mu(x)$

$$\Omega = \frac{1}{\sqrt{\phi^\dagger \phi}} \begin{pmatrix} \phi_2^* & \phi_1 \\ -\phi_1^* & \phi_2 \end{pmatrix}, \tag{2.24}$$

where

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}. \tag{2.25}$$

is a $SU(2)_L$ doublet scalar field which is considered to be a dressing field and does not need to propagate. The same approach can be applied to fermions [19].

The analogy between the effective action for the electroweak interactions (2.22) and that of Extended Gravity is striking. Considering only the leading terms, the above theory can be written as a Taylor series of the form

$$f(R) \simeq \Lambda + f_0^{'} R + \frac{1}{2!} f_0^{''} R^2 + \frac{1}{3!} f_0^{'''} R^3 + ... \tag{2.26}$$

where the coefficients are the derivatives of $f(R)$ calculated at a certain value of R. Clearly, the extra gravitational degrees of freedom can be suitably transformed into scalar fields ϕ which allows to avoid the hierarchy problem [3]. Both electroweak theory and Extended Gravity have a dimensional energy scale which defines the strength of the interactions. The Planck mass sets the strength of gravitational interactions while the weak scale λ determines the range and the strength of the electroweak interactions. As shown above, these scales can be compared at TeV energies.

In other words, the electroweak bosons are not gauge bosons in standard sense but they can be "derived" from the further gravitational degrees of freedom emerging in Extended Gravity. The local $SU(2)_L$ gauge symmetry is imposed at the level of quantum fields. However there is a residual global $SU(2)$ symmetry, i.e. the "custodial symmetry". In the case of gravitational theories formulated as the $GL(4)$-group of diffeomorphisms, tetrads are an unavoidable feature necessary to construct the theory. They are gauge fields which transform under the local Lorentz group $SO(3, 1)$ and under general coordinate transformations, the metric $g_{\mu\nu} = e^a_\mu e^b_\nu \eta_{ab}$ which is the field that is being quantized, transforms under general coordinate transformations which is the equivalent of the global $SU(2)$ symmetry for the weak interactions (in our case the residual $GL(2) \supset SU(2)$) . Such an analogy between the tetrad fields and the Higgs field is extremely relevant. We can say that the Higgs field has the same role of the tetrads for the electroweak interactions while the electroweak bosons have the same role of the metric.

A gravitational action like (2.26) is, in principle, non-perturbatively renormalizable if, as shown by Weinberg, there is a non-trivial fixed point which makes the gravity asymptotically free [24]. This scenario implies that only a finite number of the Wilson coefficients in the effective action would need to be measured and the theory would thus be predictive and probed at LHC.

Measuring the strength of the electroweak interactions in the electroweak W-boson scattering could easily reveal a non-trivial running of the electroweak scale v. If an electroweak fixed point exists, an increase in the strength of the electroweak interactions could be found, as in the strongly interacting W-bosons scenario, before the electroweak interactions become very weak and eventually irrelevant in the fixed point regime. In analogy to the non-perturbative running of the non-perturbative Planck mass, it is possible to introduce an effective weak scale

$$v_{eff}^2 = v^2 \left(1 + \frac{\omega}{8\pi} \frac{\mu^2}{v^2}\right), \tag{2.27}$$

where μ is an arbitrary mass scale, ω a non-perturbative parameter which determines the running of the effective weak scale and v is the weak scale measured at low energies. If ω is positive, the electroweak interactions would become weaker with increasing center of mass energy. This asymptotically free weak interaction would be renormalizable at the non-perturbative level without having a propagating Higgs boson again in analogy to Extended Gravity [9].

The asymptotically free weak interaction scenario could also solve the unitarity problem of the Standard Model without a Higgs boson. In this case, there are five amplitudes contributing at tree-level to the scattering of two longitudinally polarized electroweak W-bosons. Summing these five amplitudes, one finds at order s/M_W^2.

$$\mathcal{A}\left(W_L^+ + W_L^- \to W_L^+ + W_L^-\right) = \frac{s}{v_{eff}^2}\left(\frac{1}{2} + \frac{1}{2}\cos\theta\right), \tag{2.28}$$

where s is the center of mass energy squared and θ is the scattering angle. Clearly if v_{eff} grows fast enough with energy, the ultra-violet behaviour of these amplitudes can be compensated and the summed amplitude can remain below the unitary bound. A similar proposal has been made to solve problems with unitarity in extra-dimensional models [20].

It is important to stress that our approach does not require new physics but takes only into account the whole budget of gravitational degrees of freedom. The monitoring of the strength of the electroweak interactions in the W-bosons scattering at LHC could establish the existence of a fixed-point in the weak interactions. Using the one-loop renormalization group of the weak scale could help in formalizing this picture [21]. To be more precise, let us consider the scale of electroweak interactions

$$v(\mu) = v_0 \left(\frac{\mu}{\mu_0}\right)^{\frac{\gamma}{16\pi^2}}, \tag{2.29}$$

where

$$\gamma = \frac{9}{4}\left(\frac{1}{5}g_1^2 + g_2^2\right) - Y_2\left(S\right) \tag{2.30}$$

and

$$Y_2\left(S\right) = \mathrm{Tr}\left(3\mathrm{Y}_u^{\dagger}\mathrm{Y}_u + 3\mathrm{Y}_d^{\dagger}\mathrm{Y}_d + \mathrm{Y}_e^{\dagger}\mathrm{Y}_e\right), \tag{2.31}$$

where Y_i are the respective Yukawa matrices. If the theory is in the perturbative regime e.g. at m_W, the Yukawa coupling of the top dominates since at this scale $g_1 = 0.31$ and $g_2 = 0.65$ and γ is negative. In this case, the scale of the weak interactions become smaller. If the weak interactions become strongly coupled at TeV region, g_2 becomes large and γ is expected to become positive. We obtain the expected running and the weak scale becomes larger. This is not possible in the framework of a perturbative approach. This result could represent a "signature" for the approach presented here. However, we stress once again that there are indications of a non-trivial fixed point for thenon-linear sigma model using exact renormalization group techniques [22]. In conclusion, the unitarity problem of the weak interactions could be fixed by a non-trivial fixed point in the renormalization group of the weak scale. A similar mechanism could also fix the unitarity problem for fermions masses [23] if their masses are not generated by the standard Higgs mechanism but in the same way considered here (let us remind that also $SU(3)$ could be generated by the splitting of $GL(4)$-group). In the case of electroweak interactions this approach could be soon checked at LHC but good indications are also available for QCD [25].

2.4 Discussion and Conclusions

The goal of the present approach is to pursue a unification scheme of fundamental interactions based on (i) a non-perturbative dynamics, (ii) the non-introduction of ad hoc hypotheses and (iii) the consideration of the minimal necessary number of free parameters and dimensions. In principle, different Extended Theories of Gravity can be conformally related each other and derived from a 5D manifold where the fifth dimension can assume the meaning of "mass generator". In other words, it is possible to derive a unification scheme based on the assumption that a 5D-space can be defined where conservation laws are always and absolutely conserved. Such a General Conservation Principle [15] holds since we ask for the validity of the 5D-Bianchi identities which must be always non-singular and invariant for every diffeomorphism. The 5D-space is a smooth, connected and compact manifold where we can derive field equations, geodesic equations and a globally defined Lorentz structure. The standard physics emerges as soon as we reduce from 5D to 4D-space recovering the $GL(4)$-group of diffeomorphisms. By the reduction procedure one is capable of generating the masses of particles and their organization in families. The byproduct is a 4D effective theory of gravity where further gravitational degrees of freedom naturally emerge, induced by the fifth dimension. In other words, we do not recover the Standard GR but Extended Theories of Gravity where non-minimal coupling, scalar field self-interaction potentials and higher-order curvature terms have to be considered. These theories can be confronted and related by conformal transformations.

Furthermore, the $GL(4)$ group of diffeomorphisms can be suitably split generating the fundamental groups of physical interactions. In this respect, a possible the group splitting can be

$$\underbrace{\underbrace{GL\,(4)}_{4\times 4}}_{\text{diffeom.}} \supset \underbrace{\underbrace{SU\,(3)}_{3^2-1}}_{\text{gluons}} \otimes \underbrace{\underbrace{SU\,(2)}_{2^2-1}}_{\text{vec. bosons}} \otimes \underbrace{\underbrace{U\,(1)}_{1}}_{\text{photon}} \otimes \underbrace{\underbrace{GL\,(2)}_{2\times 2}}_{\text{gravitons}} \tag{2.32}$$

with further gravitational degrees of freedom [9].

The main feature of this approach is that higher-order terms or induced scalar fields enlarge the gravitational sector giving rise to massless, massive spin-2 gravitons and massive spin-0 gravitons [3, 29]. Such gravitational modes results in 6 polarizations, according to the prescription of the Riemann theorem stating that in a given N-dimensional space, $N(N-1)/2$ degrees of freedom are possible. The massive spin-2 gravitational states are ghost particles. Their role result relevant as soon as we can define a cut-off mass at TeV scale (the vacuum state of the scalar field) that allows both to circumvent the hierarchy problem and the detection of the Higgs boson. In such a case, the Standard Model of particles should be confirmed without recurring to perturbative, renormalizable schemes involving new particles. The weakness of self-interaction coupling would guarantee the fact that gravity could be compared, at TeV scale, with electroweak interaction.

However, some crucial points have to be considered in order to improve of the proposed approach. The main goal of our scenario is that the Standard Model of particles could be generated by the effective gravitational interactions coming from higher dimensions. In particular the gauge symmetries and mass generations could be achieved starting from conservation laws in 5D. It is important to stress that the Standard Model does not mean only the gauge interaction but also quarks and leptons with their mass matrices that have to be exactly addressed. In particular, the fermion sector has to be recovered.

It is well-known that the standard gauge interactions contains the chiral gauge interactions, which, in our picture, have to be generated from the gravitational interactions otherwise there is no possibility to distinguish between the left-handed and the right-handed particles. In particular, the $SU(2)$ part of the standard gauge interactions, generated from $GL(4)$, has to be chiral and, consequently, fermions acquire a chiral representation. To this end, torsion fields have to be incorporated for the following reasons. As discussed in [13], the Cartan torsion tensor plays the role of spin source in the gravitational field equations where the affine connection is not simply Levi-Civita. Furthermore, as demonstrated in [2], torsion plays an important role in Extended Theories of Gravity since brings further gravitational degrees of freedom responsible of chiral interactions. In other words, torsion is not only the source of spin but, thanks to the non-trivial structure of connections $\Gamma^{\alpha}_{\beta\gamma}$, can give rise to chiral interactions of geometric origin [26]. This means that the $SU(2)_L$ and $SU(3)$ could be related to a gravitationally induced symmetry breaking process where torsion plays a fundamental role. Due to these facts, the present approach has to be generalized including torsion fields. Phenomenological studies considering torsion and fermion interactions are already reported in [27].

Furthermore, as shown in [26], space-like, time-like and null torsion tensors, generated by non-trivial combinations of vector and bi-vector fields, can be classified and represented by matrices which could explain mass matrices of fermions and the hierarchy in the generation of quarks. This approach agrees with other approaches where the effects of gluonic condensates in holographic QCD can be encoded in suitable deformations of 5D metrics (see e.g. [18]).

A detailed study in this sense will be the argument of forthcoming studies.

The validity of the presented scheme could be reasonably checked at LHC in short time, due to the increasing luminosities of the set up. In fact, the LHC experiments (in particular ATLAS and CMS) are indicating, very preliminary, the presence of resonances and condensate states that confirm the Standard Model but, up to now, cannot be considered as evidences for the Higgs boson [25]. Similar results, but with larger integrated luminosity, are reported also by the CDF collaboration at Fermi Lab [28]. The interpretation of such data could be that the further gravitational modes discussed here would induce the formation of resonances and condensates giving rise to a sort of gravitational Higgs mechanism.

References

1. Basini G, Capozziello S (2003) Gen Rel Grav 35:2217
2. Capozziello S, Cianci R, Stornaiolo C, Vignolo S (2008) Int J Geom Meth Mod Phys 5:765
3. Capozziello S, Basini G, De Laurentis Eur M (2011) Phys J C 71:1679
4. Capozziello S, Matsumoto J, Nojiri S, Odintsov SD (2010) Phys Lett B 693:198
5. Birrell N, Davies PC (1984) Quantum fields in curved space. Cambridge University, Cambridge
6. Pich A (2007) arXiv: 0705.4264[hep-ph] (2007)
7. Antoniadis I, Dimopoulos S, Giveon A (2001) JHEP 0105:055
8. Calmet X, Hsu SDH, Reeb Phys D (2008) Rev D 77:125015
9. Capozziello S, De Laurentis M (2011) Phys Rept 509:167
10. Calmet X, Hsu SDH (2008) Phys Lett B 663:95
11. Meade P, Randall L (2008) J High Energy Phys 0805:003
12. Anchordoqui LA, Feng JL, Goldberg H, Shapere AD (2002) Phys Rev D 65:124027
13. Hehl FW et al (1970) J Math Phys 12:1334
14. Coleman S, Wess J, Zumino B (1969) Phys Rev 117:2239
15. Basini G, Capozziello S, Longo G (2003) Phys Lett A 311:465
16. Llewellyn Smith CH (1973) Phys Lett B 46:233
17. Dvali G, Giudice GF, Gomez C, Kehagias A (2010) CERN-PH-TH-226
18. Cappiello L, D'Ambrosio G (2010) Eur Phys J C 69:315
19. 't Hooft G Topological aspects of quantum chromodynamics, Lectures given at International School of Nuclear Physics: 20th Course: Heavy Ion Collisions from Nuclear to Quark Matter (Erice 98), Erice, Italy, 17–25 Sep 1998, Topological aspects of quantum chromodynamics, hep-th/9812204, published in Erice 1998, From the Planck length to the Hubble radius, pp.216–236

20. Hewett J, Rizzo T (2007) JHEP 0712:009
21. Arason H, Castano DJ, Keszthelyi B, Mikaelian S, Piard EJ, Ramond P, Wright BD (1992) Phys Rev D 46:3945
22. Fabbrichesi M, Percacci R, Tonero A, Zanusso O (2011) Phys Rev D83:025016
23. Appelquist T, Chanowitz MS (1987) Phys Rev Lett 59, 2405 [Erratum-ibid. 60, 1589 (1988)]
24. Weinberg S, in Understanding the Fundamental Constituents of Matter, ed. A. Zichichi (Plenum Press, New York, 1977); in General Relativity, ed. S. W. Hawking and W. Israel (Cambridge University Press, 1979): 700
25. Aad G et al (2011) The ATLAS collaboration. Eur Phys J C 71:1512
26. Capozziello S, Lambiase G, Stornaiolo C (2001) Ann Phys 10:713
27. Belyaev AS, Shapiro IL, do Vale MBA(2007) Phys Rev D 75: 034014
28. Aaltonen T et al. (2011) arXiv: 1104.0699 [hep- ex]
29. Bogdanos C, Capozziello S, De Laurentis M, Nesseris S (2010) Astropart Phys 34:236–244

Chapter 3
Enrico Fermi and Ettore Majorana: So Strong, So Different

Francesco Guerra and Nadia Robotti

Abstract By exploiting primary sources we will analyze some of the aspects of the very complex relationship between Enrico Fermi and Ettore Majorana, from 1927 (first contacts of Majorana with the Institute of Physics of Rome, and with Fermi) until 1938 (disappearance of Majorana). The relationship between Fermi and Majorana can not be interpreted in the simple scheme Teacher-Student. Majorana, indeed, played an important role in the development of research in Rome in the field of the statistical model for the atom and in nuclear physics.

Our current research concerns the development of Nuclear Physics in Italy in the Thirties of Twentyth Century, and is based exclusively on primary sources (archive documents, scientific literature printed on the journals of the time, and so on). In this framework, we will try to outline some aspects of the complex topic concerning the relationship between Enrico Fermi and Ettore Majorana. Of course, this chapter will touch only some of the most important issues. For convenience, our exposure will be connected to a periodization of Majorana scientific activity, which we used in previous works.

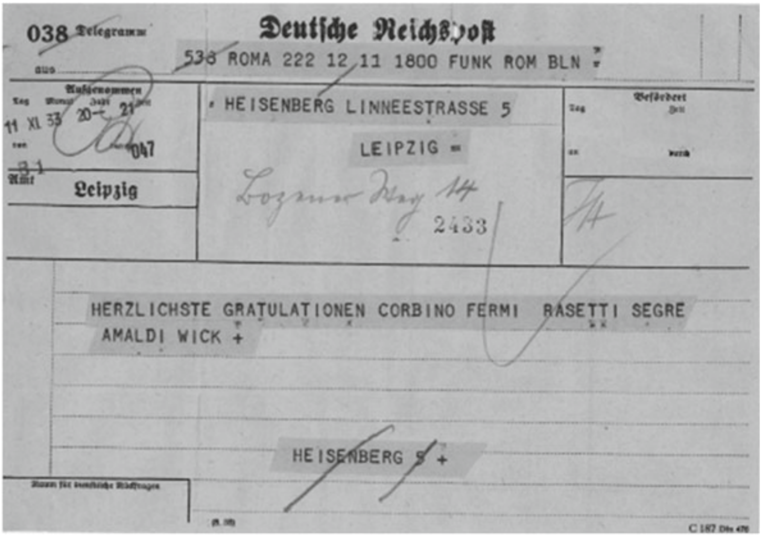
One of our important results is the reassessment of the role played by Majorana for the decisive orientation of the research in Rome (statistical model of the atom, nuclear physics). This role is obscured in discussions which place emphasis on the (alleged) "genius" of Majorana, usually associated with his (alleged) "lack of common sense". Perhaps we could summarize, in a very concise and expressive manner, the nature of the relationship between Majorana and Fermi, and in general the Physical Institute of Rome, presenting these two important documents stored in the "Archive Heisenberg", at the Max Planck Institute of Munich.

On November 9, 1933, the Sweden Academy of Sciences announced that the Nobel Prize for Physics for the year 1932 was awarded to Werner Heisenberg.

F. Guerra (✉)
Department of Physics, Sapienza University of Rome and INFN, Rome, Italy

N. Robotti
Department of Physics, University of Genova, Rome, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_3,


038 Telegramm

Deutsche Reichspost

538 ROMA 222 12,11 1800 FUNK ROM BLN

Aufgenommen

11 XI 33 20— 21

047

Leipzig

HEISENBERG LINNEESTRASSE 5

LEIPZIG =

2433

Befördert

HERZLICHSTE GRATULATIONEN CORBINO FERMI RASETTI SEGRE
AMALDI WICK +

HEISENBERG 5 +

Fig. 3.1 "Roma" congratulations (Heisenberg archive, Max Planck institute, Munchen)

Immediately, all the exponents of world culture send their congratulations, in various forms. Heisenberg stored in a folder of his personal archive all received messages. Among them, a telegram by the Deutsche Reichspost, from Rome and dated 11.11.33 (Fig. 3.1) written in a very cold and formal style, in German: "Most cordial congratulations corbino fermi rasetti segre amaldi wick."

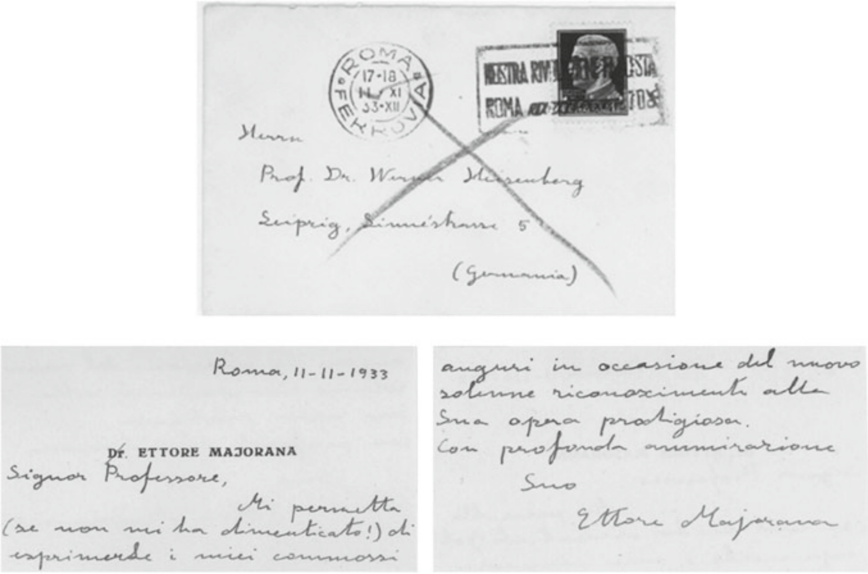
We note the order of signing according to the close rank of academic seniority at the Institute of Physics in Rome (Orso Mario Corbino, Enrico Fermi, Franco Rasetti, Emilio Segre', Edoardo Amaldi, Gian Carlo Wick).

Ettore Majorana (who is in Rome) is not included in the list, not even at the very last place. But Majorana, who had left Leipzig in early August, sends his "Gratulationen", according to his style. He sends a small personal business card, dated Rome, 11.11.1933 (Fig. 3.2) with the title "Dr." canceled by hand, written in a poignant Italian. It is a very intense letter.

"Dear Professor, Let me (if you have not forgotten me !) allow to express my greetings on the occasion of the new formal recognition of your prodigious work. With deep admiration Yours Ettore Majorana."

3.1 The Formation Years Until the Doctoral Degree (1929)

The scientific research activities of Ettore Majorana develop immediately at the highest international standard, while it is still a university student at the School of Engineering, in 1928.


Herrn
Prof. Dr. Werner Heisenberg
Leipzig,
(Germania)

Roma, 11-11-1933

Dr. ETTORE MAJORANA

Signor Professore,
Mi permetta (se non mi ha dimenticato!) di esprimerle i miei commossi auguri in occasione del nuovo solenne riconoscimento alla Sua opera prodigiosa.
Con profonda ammirazione
Suo
Ettore Majorana

Fig. 3.2 Majorana congratulations (Heisenberg archive, Max Planck institute, Munchen)

His first research concern the applications, the improvement, and the extension of the statistical model for atoms, introduced by Enrico Fermi at the end of 1927, only few months before! [1]. This model is now known as the "Thomas-Fermi model".

As it is well known, Enrico Fermi was called as full Professor in Theoretical Physics in Rome in 1927, through the effort of the Director of the Institute Orso Mario Corbino to develop advanced modern physics in Rome. The main Fermi achievements concern: the so called "Fermi-Dirac statistics", the statistical model for the atom, the theory of weak interactions, the discovery of the neutron-induced radioactivity, the effect of the slowing down of neutrons, (after the 1938 Nobel Prize and the emigration to the USA) the atomic pile, the Manhattan project, the elementary particle physics, the computers, and so on).

The first involvement of Majorana on the statistical model subject is a paper in collaboration with his friend Giovanni Gentile jr., published on the Proceedings of the "Accademia dei Lincei", presented on July 24th, 1928 by Orso Mario Corbino [3].

They calculate the splitting of the spectroscopic energy levels due to the hypothesis of the spinning electrons as recently developed by Dirac. It is a well received paper, developed completely in the frame of Fermi aproach.

Then Majorana continues his research alone, with full autonomy and effectiveness.

He proposes an improvement of the model (he changes the expression of the effective potential acting on the optical electron) and includes also positive ions (it is the first treatment made). Some of the results are communicated to the 22nd General Meeting of the Italian Society of Physics (Rome, 28–30 December 1928).

At that time Majorana was still a student, and had recently officially moved from Engineering to Physics. The communication was presented by Majorana in front of an audience of famous Physicists and Mathematicians, as O. M. Corbino, T. Levi Civita, V. Volterra, G. Polvani, Q. Majorana, A. Carrelli, E. Fermi, (he was not a timid person!)

This communication is regularly published on "IL Nuovo Cimento", the Journal of the Society [8], but it has been never mentioned in any scientific paper, nor by any of the many historians who wrote on the life and activity of Majorana. Historical analysis is in a paper by F. G. and N. R. [5]. The acknowledgements are very interesting: "The A. thanks Professor Fermi, for the advice and suggestions around new applications of this statistical method that has thrown much light on the atomic physics, and whose fertility, appears far from exhausted, still waiting to be ventured into investigation of fields of larger scope and more full of promise."

The results contained in Majorana notebooks (at Domus Galiaeana in Pisa) and in the communication show that Majorana reached a fully developed scientific personality, completely independent from Fermi. However, Majorana does not publish the results announced in the communication, nor the other results contained in his notebooks. An enlarged paper on the subject would have given a complete representation of the statistical model, his applications and extensions. As a matter of fact, Majorana does not work anymore on these subjects.

Fermi convinces himself of Majorana improvement only in late 1933 and puts it at the basis of the monumental conclusive paper of the Rome School, co-authored by Fermi and Amaldi in 1934, without any reference to Majorana [2]. It is amazing to note that all formulas of the general Fermi-Amaldi paper (1934) coincide with the corresponding formulas of Majorana communication (1929).

Majorana earns his doctoral degree in Physics (July 6th, 1929) with a Thesis on Nuclear Physics (with the title "On the mechanics of radioactive nuclei"). Fermi is the supervisor. Majorana Thesis is the first work on Nuclear Physics in Rome, and also in Italy. Majorana gives a rigorous justification to Gamow model of alpha decay based on the quantum tunnel effect. Anyway, even if original and internationally competitive results are achieved, they are not published.

3.2 From the Doctoral Degree to the Private Professorship (1929–1932)

After graduating a short period of silence: he does not deal at the moment with Nuclear Physics, but maturates new lines of research (Atomic Physics, Molecular Physics and Elementary Particle Physics) in complete autonomy from Fermi (which deals with Quantum Electrodynamics and Hyperfine Structure of Atomic Spectra).

Then follows a very intense activity, oriented toward the "IL Nuovo Cimento".

Between the end of 1930 and January 1931 there are ready two papers [9, 10] on the quantum explanation of the chemical bond (formation of Helium molecular ion and Hydrogen molecule). The first is presented by Corbino at Accademia dei Lincei, December 7, 1930: "I warmly thank Professor Enrico Fermi, who gave me precious advice and aid".

Majorana becomes a pioneer in Theoretical Chemistry (this discipline will develop in Italy only in the 50s of twentyth Century).

In 1931, other two papers will be published, this time in Spectroscopy. Majorana provides a theoretical interpretation of two new lines of Helium, recently discovered [11] and some triplets of calcium [12]. These two works are appreciated as elegant examples of applications of group theory. In reality they have a direct physical interest: for the first time the role of the phenomenon of self-ionization of atoms is recognized in Spectroscopy.

Follows a very important work (published in 1932): "Atoms oriented in magnetic field" [13]. In this chapter, he proposes an optimal arrangement of the magnetic field to show the sudden flip in the spatial quantization of the spin of atoms, and other related effects. This arrangement is immediately adopted with great success in the laboratory of Otto Stern in Hamburg (where there is temporarily Emilio Segrè). Majorana gives the "Announcement" (never quoted in Literature), in the Journal "Ricerca Scientifica" [14], so as to publicize immediately his results and have priority of discovery!

Then the fundamental work: "Theory of relativistic particles with arbitrary intrinsic momentum" [15]. He formulates a relativistic generalization of the Schrödinger equation that completely eliminates the existence of negative-energy solutions (provided instead by the Dirac equation) and valid for particles with arbitrary spin (Dirac equation instead is valid only for $s = 1/2$). It is expanded and systematized by Wigner in 1939, which recognizes the pioneering role of Majorana. He thanks Fermi: "I especially thank Professor E. Fermi for the discussion of this theory."

The period 1930–1932 is therefore scientifically very intense. The activity is carried out in complete independence. In this period we mark an additional peculiar aspect of Majorana scientific career. After earning the doctoral degree he does not receive any position (all other brilliant young "Panisperna boys" are immediately hired in the University, at the beginning with temporary "assistant" positions). Majorana "frequents freely the Institute, by following the scientific movement". In November 1932 he earns the abilitation to the private professorship ("libera docenza") in Theoretical Physics. Fermi is the Chairman of the Minister examination Committee.

At the end of 1932, the "strict relationship" between Ettore Majorana and the Rome Institute of Physics (in particular Enrico Fermi) come to the end, as he explicitly remarks in his 1937 curriculum presented in the application for the Palermo professorship.

3.3 The Visit to Lipsia (1933)

It is a strategic decision! With the support of Fermi, he applies for a fellowship to be exploited abroad to the National Council for Research (C.N.d.R.). He gets the fellowship: 12,000 Lire for six months (the salary of an industry worker was less than

1,000 Lire/month, for a full professor was around 25,000 Lire/year). His program is very advanced: Nuclear Physics and Elementary Particle Theory. Fermi support letter to Majorana project is peculiar (more conservative): he says that Majorana will continue with profit his research on atomic physics, and applications of group theory. He does not seem aware of the advanced Majorana programs: Nuclear Physics and Elementary Particles

To understand the apparent discrepancy between two descriptions of the planned program, it is necessary to recall some crucial aspects of the situation in Rome, in the strategic year 1932. This is addressed in depth in our articles [4, 5, 7] and in the monograph [6]. Here we are going to some brief schematic remarks. Since the beginning of the Thirties, it was clear to Orso Mario Corbino and Enrico Fermi that the thrust of Atomic Physics, even in its newest quantum aspects, was running low, and that the new open frontier was that of the physical study of the atomic nucleus. Hence the decision to organize in Rome, in October 1931, an International Conference of Nuclear Physics, the first of its kind in the world, which took place with great success with the support of the Academy of Italy and Volta Foundation, with a budget of about 200,000 lire (full original documentation is at the Accademia dei Lincei). But after the Conference the difficulties in starting the actual research in Nuclear Physics were evident. To witness the deep atmosphere of indecision we report about significant passages of the letter, dated September 30, 1932, sent by Fermi, from Arno Mignano, to his collaborator Emilio Segrè, then in Hamburg (the original is at Fermi Archives of Chicago): "About the work programs for the coming year, I have none at all: I do not even know whether I will return to play with the Wilson Chamber, or whether I become again a theoretical physicist. Of course the problem of equipping the Institute to work on the nuclear physics is becoming more urgent, if we do not want to reduce us too much in a state of intellectual slumber."

Moreover, in Rome it is not immediately grasped the significance of the discovery of the neutron by Chadwick, announced in February 1932 (after preliminary results of Bothe and Becker, and Joliot and Irene Curie). The Fermi report on the structure of the nuclei, in an important conference in Paris in July 1932, almost does not mention it. It merely exposes the situation at the end of the Congress in Rome, many months before, including the difficulties in the quantum description of the alleged nuclear electrons (which would have required an alleged new theory radically different from quantum mechanics). On the other hand, Heisenberg in Leipzig, immediately realizes the potentialities of the existence of the neutron for a possible quantum-mechanical description of nuclear structure, as it will be explained later. And Majorana, who "follows the scientific movement", snaps the Heisenberg program. Fermi on the other hand, from the theoretical standpoint, continues with his research, at the highest level, on the hyperfine structures, where the magnetic properties of the nucleus are revealed in the change of the spectral lines at the atomic level. While in Rome, experimental research in the field of nuclear physics is oriented towards the study of energy levels through the nuclear spectroscopy of gamma rays emitted in nuclear decays.

This strategic decision was certainly influenced by previous experience in Spectroscopy in Rome (Franco Rasetti, Professor of Spectroscopy, had obtained, among other things, extremely important results on Raman effect). And it is significant that, in the spring of 1933, the National Research Council plans the national research in nuclear physics through a rigid division of tasks, so organized: Rome is responsible for the gamma spectroscopy; physics of the neutron is assigned in Florence to Gilberto Bernardini and collaborators; Padua gets cosmic rays for Bruno Rossi and collaborators. In this context, Fermi and Rasetti also make up a bizarre gamma spectrograph with a crystal of bismuth, described in an article in “La Ricerca Scientifica”, but that will never find application in actual research. In some ways, it is really amazing the Majorana physical intuition, that an effective study of the structure of the nucleus was within the practical possibilities of the moment.

In Leipzig Majorana enters into good relations with Heisenberg. Heisenberg, starting from July 1932, after the discovery of the neutron, was developing a model for the nucleus, assuming that it was composed only of protons and neutrons, held together by exchange forces, very similar to those entering the chemical bond. His results were explained in a series of three papers published, since July 1932, in the eminent German Journal “Zeitschrift für Physik”. Majorana in Leipzig gets the chance to see the third paper before publication and makes two fundamental improvements to Heisenberg theory: Majorana exchange forces change only the position of the two interacting proton and neutron, and have a sign opposite to Heisenberg exchange forces. The advantages of Majorana proposal are very deep. In particular the α particle is recognized as the most stable nuclear structure, and the almost uniform density of nuclei is explained. He immediately publishes a paper in the German Review “Zeitschrift für Physik” [16], as Heisenberg did, and also an “Announcement” (never quoted in Literature), in the Journal “Ricerca Scientifica” [17]. Heisenberg realizes immediately the advantages of Majorana scheme, and begins immediately to advertise these results, in particular in his report at the important Solvay Conference in Bruxelles (October 1933). It is an international triumph for the young Majorana.

After the success of Majorana, research in Nuclear Physics in Rome received by Fermi an energic re-orientation. The proton-neutron nuclear model of Heisenberg-Majorana, based on quantum mechanics, opens the way to the Fermi theory of beta decay (December 1933), based on quantum field theory, where the electron and neutrino are created at the time of beta decay. There is no pre-existing electron in the nucleus, in agreement with Heisenberg-Majorana. It is fully recognized the centrality of the neutron, and Fermi discovers neutron-induced radioactivity in March 1934, led by his theory of beta decay. The discovery of the effects produced by the slowing down of neutrons complete the extraordinary results obtained by Fermi in the period December 1933–October 1934, that will be worth of the Nobel Prize in 1938. These successes take place along lines of research that completely subvert the schedule provided above. Rome actually became the world center of the neutron physics (the bismuth crystal gamma spectrograph no longer exists even in the Museum). Some merit to this success should surely be attributed to Majorana.

3.4 The Silence (1933–1937)

After his return from Leipzig (Majorana comes back to Rome at the beginning of August 1933, at the highest level of his scientific prestige) we have total absence of publications.

His abilities displayed as a leader in the research can find no effective realization.

Moreover, he attempts to attend to his duties for the private professorship, by submitting to the Faculty every year very advanced programs (always different) for the proposed courses. His scientific interests can be reconstructed from the most advanced parts of his proposed programs. He never succeeds to give the free course. He never gets a position at the University.

3.5 Ettore Majorana Full Professor of Theoretical Physics at the University of Naples (1938)

In 1937 Majorana officially reappears in the scientific world, perhaps in connection with the planned national competition for Theoretical Physics (Palermo): he publishes on Nuovo Cimento one of the most important works of his life: "The symmetrical theory of electrons and positrons" [18]. It is a remarkable paper, completely up to date even from the experimental point of view, dealing with the quantum theory of interacting fields, so that the "Dirac sea" (fully occupied negative energy states) is avoided. It foresees the existence of an elementary particle of spin 1/2, which coincides with its anti-particle, "the Majorana neutrino". If the paper was published in order to make stronger his curriculum, then Majorana is moving along well established academic strategies. Nearly one quarter of the manuscript, dated around 1936, is preserved at the Domus Galilaeana in Pisa.

On the relationship with Fermi, we have a reprint of the article with dedication (courtesy of Prof. Giacomo Morpurgo): "To His Excellency Enrico Fermi, with very best regards. Ettore Majorana". Recently we found a reprint with dedication to Gian Carlo Wick: "A Gian Carlo Wick with very best regards. Ettore Majorana." Note the nuance in the dedications: Fermi (His Excellency), expected President of the Evaluating Committee, Wick, a candidate in pole position.

The competition for a full professorship in Theoretical Physics at the University of Palermo (deadline for applications: June 15th, 1937) is a particularly significant example of the academic practice of the time (rich documentation at the Central State Archives in Rome). Ettore diligently submits his application (it is the first occasion for a chance to be officially recognized). Attached to the application are: "Scientific activity" and "List of Publications". Obviously in the list it is missing the communication to Congress in 1928 and the two Announcements on the "Ricerca Scientifica" (not suitable for a Competition).

On July 7, the Committee was appointed directly by the Minister (in a totalitarian regime). Of course Fermi is the President of the Committee, so composed: "Fermi

Enrico; Persico Enrico; Lazzarino Orazio; Polvani Giovanni; Carrelli Antonio". In the first meeting (October 25, 1937, Record 1) the Committee decides to propose to the Minister to appoint directly Majorana as full professor in some University of the Kingdom, outside the procedures of the competition. The scientific Report on Majorana is of course very laudatory, albeit with some important omissions. The Minister accepts the proposal immediately, and on November 2, 1937, Majorana is appointed Full Professor in Theoretical Physics at the Royal University of Naples.

Having removed Majorana from the competition, the work of Committee resumes immediately beginning again with the record N°1! (not number 2!). In November 1937 the Committee, according to the rules, selects the winning "triplet" in the following order: "Wick Giancarlo; Racah Giulio; Gentile Giovanni".

3.6 The Disappearence (1938)

Majorana in Naples attends regularly to his course. The contents can be reconstructed from his notes in Pisa and from the testimony of his students, in particular Prof. Gilda Senatore. From the notes at the Domus, the intent of Majorana is clearly to differentiate his course from that of Fermi, in particular by developing in full detail, some of the important topics, that Fermi had just mentioned (as for example the relativistic corrections to the quantum Bohr atom).

He disappeared in circumstances not yet clarified in late March 1938.

After his disappearance, a presumed intervention by Fermi on Mussolini, to intensify the police researches, is not confirmed in the central archive in Rome (secretariat of the Duce). In fact, the alleged letter of Fermi to Mussolini (27 July 1938), preserved at the Domus (Fig. 3.3), shows puzzling aspects. It contains a central body written with "unknown" handwriting, in a very rhetoric style, certainly completely unrelated to Fermi, as even Emilio Segrè recognized in a letter addressed to Amaldi, who had sent him a copy of the text of the letter. On the other hand, the first and last lines of the letter are written in the handwriting of Ettore's brother, Salvatore. Significant is the attempt to imitate the signature of Fermi. Notice the incredible addressing of "Fermi" toward Mussolini as "Duce", while the correct addressing would have been "Excellency". This letter is a further step in the complex research theme "Majorana". As a matter of fact, a careful comparison of the handwriting of the central body of the letter with the existing documents, at the Domus Galilaeana in Pisa and the Archive of the Department of Physics in Rome, leads to a surprising conclusion. The letter was in reality written by Giovanni Gentile jr, a close friend of Majorana strictly associated with his Family. The style of the letter is in complete agreement with this discovery. A full analysis of this disconcerting episode is contained in a forthcoming paper.

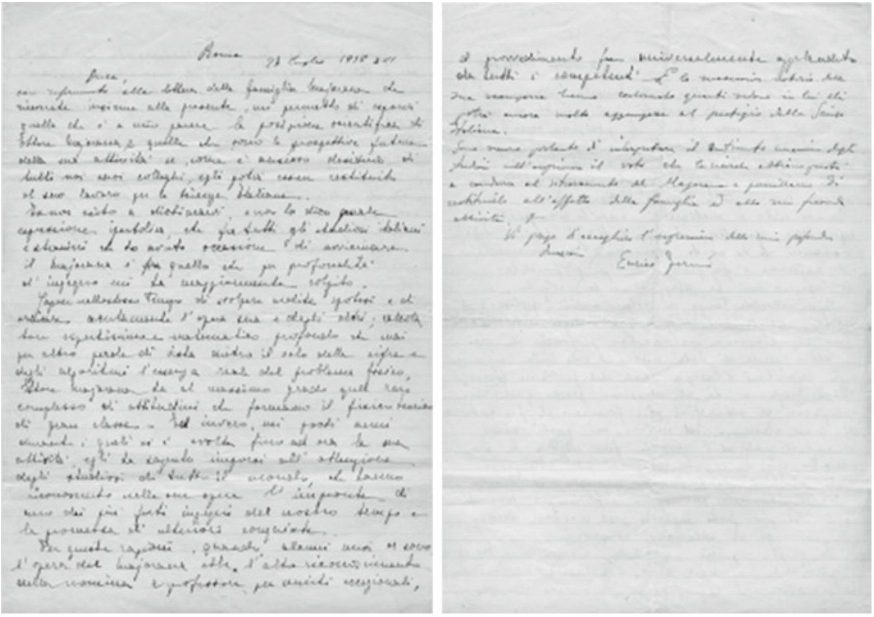


Fig. 3.3 Presumed "Fermi letter" to Mussolini (Domus Galilaeana, Pisa)

3.7 Conclusions

In the confirmatory report (1930) for Fermi Professor of Theoretical Physics, the commission (chairman orso mario corbino) recognized, among other things, the merits of the "construction of a school of young people vigorously trained in the study of the most advanced problems of modern physics". Ettore Majorana was certainly one of them. However, the complex issue of relations between Majorana and Fermi can not be simplistically interpreted in terms of Teacher-Student. Majorana since the early activity shows a wide autonomy and independence. His ability to "follow the scientific movement" allows him to have a significant influence on the development of research in Rome, mainly on the themes of the statistical model of the atom, first, and the physics of the nucleus, later. The international prestige achieved by the development of nuclear models directed him towards an effective role as leader of the research. The situation of apparent separation from active research, created after his return from Leipzig in August 1933, has prevented the occurrence of this event.

Italy lost in 1938, almost simultaneously but in different circumstances, the two leading figures in the field of research in nuclear physics, Majorana and Fermi.

References

1. Fermi E (1927) Un metodo statistico per la determinazione di alcune proprietà dell'atomo. Rendiconti dell'Accademia dei Lincei 6:602–607
2. Fermi E, Amaldi E (1934) Le orbite∞s degli elementi. Memorie dell'Accademia d'Italia 6:119–149
3. Gentile G, Majorana E (1928) Sullo sdoppiamento dei termini Roentgen e ottici a causa dell'elettrone rotante e sull'intensità delle righe del Cesio. Rendiconti dell'Accademia dei Lincei 8:229–233
4. Guerra F, Leone M, Robotti N (2006) Enrico Fermi's discovery of neutron-induced radioactivity: neutrons and neutron sources. Phys Perspect 8:255–281
5. Guerra F, Robotti N (2008) A) Ettore Majorana's forgotten publication on the Thomas-Fermi model. Phys Perspect 10:56–76
6. Guerra F, Robotti N (2008 B) Ettore Majorana. Aspects of his Scientific and Academic Activity. Edizioni della Normale p 243
7. Guerra F, Robotti N (2009) "Enrico Fermi's discovery of neutron-induced artificial radioactivity: the influence of his theory of beta decay". Phys Perspect 11:379–404
8. Majorana E (1929) Ricerca di un'espressione generale delle correzioni di Rydberg, valevole per atomi neutri o ionizzati positivamente. Nuovo Cimento 6:14–16
9. Majorana E (1931) Sulla formazione dello ione molecolare di elio. Nuovo Cimento 8:22–28
10. Majorana E (1931) Reazione pseudopolare fra atomi di idrogeno. Rendiconti dell'Accademia dei Lincei 13:58–61
11. Majorana E (1931) I presunti termini anomali dell'elio. Nuovo Cimento 8:78–83
12. Majorana E (1931) Teoria dei tripletti P' incompleti. Nuovo Cimento 8:107–113
13. Majorana E (1932) Atomi orientati in campo magnetico variabile. Nuovo Cimento 9:43–50
14. Majorana E (1932) Atomi orientati in campo magnetico variabile. La Ricerca Scientifica 5:329
15. Majorana E (1932) Teoria relativistica di particelle con momento intrinseco arbitrario. Nuovo Cimento 9:335–344
16. Majorana E (1933) Über die Kerntheorie. Zeitschrift für Physik 82:137–145
17. Majorana E (1933) Nuove ricerche sulla teoria dei nuclei. La Ricerca Scientifica 4:522
18. Majorana E (1937) Teoria simmetrica dell'elettrone e del positrone. Nuovo Cimento 5:171–184

Chapter 4
Physics Teachers' Education (PTE): Problems and Challenges

Elena Sassi and Marisa Michelini

Abstract A vast majority of the research results acknowledge the crucial role of teacher's education, as a vital tool in enhancing the quality of physics education. The projects like PISA, ROSE and TIMMS showcase the impact of teacher's education as a qualitative improvement in the physics learning environment. In Physics Education Research (PER), the impact of teacher's education had been addressed for the its role in the enhancement of positive interest among the students. The current world-wide state of the art characterizes a large variety of boundary conditions, traditions and practices that are being followed. In our present context, we foucus and discuss on the multidimensional challanges such as competencies needed, degrees required, problems encountered, support to be provided and the basic pre-requirements of Teacher's education for the secondary schools. We present some of the teaching methods and practices followed in coherent with, both, the Student centered and open learning environments along with some of the useful didactical indicators. Also, we potray a couple of research-based examples successfully experimented in Italy. Finally we propose some useful recommendations along with the criteria to be followed in the teachers education for the overall improvement.

4.1 Introduction

This paper discusses different aspects of Physics Teachers' Education (PET), especially the problems highlighted by international survey studies on students' achievements and teachers' characteristics; the links between Physics Education Research

E. Sassi (✉)
Department of Physics,University of Naples "Federico II", Naples, Italy
e-mail: elena.sassi@na.infn.it

M. Michelini
Department of Physics, Chemistry and Environment,University of Udine, Udine, Italy
e-mail: marisa.michelini@uniud.it

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_4,

and Teacher Education; some research-based example interventions; a proposal for a EU Benchmark for physics teaching degrees. The discussion addresses the complex challenge linked to improving the quality of physics teacher preparation, both pre-service and in-service and suggests some recommendations. The strategic role of teachers in the learning processes is well acknowledged; together with students they are key actors of any project aimed at improving scientific education and deeper awareness of the future of our planet. Physics is crucial in describing, modelling, understanding the natural world; teaching and learning physics involves many different dimensions (disciplinary, cultural, historical, social…) and many links with other disciplines. The focus on PTE at school level in the so-called industrialised countries is motivated by at least three different reasons:

- Young people have goals, interests, ways of learning, lifestyles, … that differ in many respects from the people used to refer mainly to printed materials; the same holds for capabilities about Internet, social networks, combination of formal, non-formal and informal education.
- The key concepts, needs, requests of the Knowledge Society are receiving increasing attention in many countries. In this framework, scientific knowledge/education is assuming a growing importance, also as a condition for being aware of and deal with complex socio-political issues, e.g. climate changes, energy, health, …
- Despite the increasing use of technology in education and the growing momentum of informal learning, teachers remain key actors in education. There are many factors and issues related to the profession of teacher (such as the vision of the teaching/learning processes, the increasing number of competences required, the pre-service and in-service education programs, the current teaching practices, the common perception of social role of teachers, ...) which require considerable attention and specific actions.

4.2 Students' Achievements and Teachers' Competencies

Data and analyses come from several studies[1], e.g. in alphabetic order: National Task Force on Teacher Education in Physics (NTFTEP, USA); PISA; ROSE; STEPSTWO; TIMSS and Physics Education Research. The main features of students' achievements and teachers' competences induce to reflect on several aspects of the con-

[1] National Task Force on Teacher Education in Physics (NTFTEP, USA) http://www.ptec.org/taskforce OECD Programme for International Student Assessment PISA PISA www.pisa.oecd.org/ every 3 years 15 years students assessed in Reading, Mathematical and Scientific literacy PISA 2009: 34 OECD members +41 partners countries, *PISA 2009 Results: Executive Summary* ROSE The Relevance of Science Education ROSE http://www.uv.uio.no/ils/english/research/projects/rose/ STEPS TWO Academic Network (2008–2011) http://www.stepstwo.eu/ To support Physics Depts. in post Bologna processes, student-centred/flexible learning, Physics Teacher Education in Universities, to reinforce the study of Physics at Secondary Level Universities from 27 Countries + 7 Associated (Five Universities, EPS, EPSI) TIMSS Trends in International Mathematics and Science Study http://www.timss.bc.edu/TIMSS 2007: 59 countries, six benchmark

struction of a sound scientific knowledge. The achievements in science and physics at secondary school (4, 8° and final year) indicate various problems and difficulties. TIMSS Advanced 2008 indicates that Mathematics and Physics programs vary in duration and intensity (2–5 years, 100–200 hs/y), with generally fewer instructional hours in Physics. A large gap divides the highest and lowest performing countries, with a wide range between the highest and lowest achieving students. In Physics the Netherlands was the top performer; Slovenia and Norway had very similar average achievement. These three countries, together with the Russian Federation, had higher achievement in Physics. The measured change in average achievements (1995–2008) in advanced Mathematics is small in Russian Federation and negative the other three countries . In Physics, Slovenia had essentially no change, some decline for the other three countries. In most countries, the majority of students were males. The PISA 2009 comparison of countries with respect to the OECD average indicate several countries from the East (Hong-Kong, Korea, Shanghai, Singapore, Taipei, Japan) amongst the top performer on both Mathematics and Science scale, together with Canada, Australia and New Zealand. In Europe Finland, the Netherlands, Norway, Germany, Belgium and Denmark do rank well, better than USA; Italy is below the average. The ROSE project (The Relevance of Science Education, 40 Countries in 2010) does not test achievement, it addresses factors of importance to the learning of science and technology (S&T), as perceived by students (about 15 ys), "to contribute to improve curricula, while respecting cultural diversity and gender equity and empowering the learner for democratic participation and citizenship". Results from ROSE show that: students in rich countries, especially girls, have attitudes toward science and scientific careers less positive than those surveyed in so-called developing countries; in Northern EU and Japan they are more ambivalent than adults; girls, in the richest countries, are more negative or sceptical than boys; very many students, in poor countries, want to become scientists and have not this possibility.

On teachers' side, the problems come mainly from three areas: policy and organization, insufficient competencies, inadequate exchange between school and PER; their solutions present interesting challenges. Some are related with institutional aspects, as the features of the educational system (e.g. centrally vs locally decided curricula and syllabuses, teachers as State employees versus recruitment by the school, …); the status of Physics when taught as a single discipline or as part of combined science; the role of Universities and Physics Department in the pre-service education; the different standards for being a certified teacher; the recruitment procedures; the type of the agencies entitled to run programs for in-service teachers education and the contents of PTE programs, etc. Other challenges derive from the shortage of qualified physics teachers (in several countries; the transformation of the indispensable subject matter knowledge (SMK) into a richer pedagogical content knowledge (PCK) that includes applied pedagogy and PER results; the insufficient acquisition of the rapidly

(Footnote 1 continued)
participants; 4 and 8° grades; about 434,000 students; 47,000 teachers, 15,000 school principals TIMSS ADVANCED 2008 (students in last year of secondary school taking or having taken courses in advanced Mathematics and Physics: Mechanics, E&M, Heat&Temperat., Atoms, Nuclei. Ten countries: AM, IR, IT, LB, NL, NO, PH, RU, SI, SE. Changes tracked in 1995–2008: 5 Countries.

increasing number of competences (in physics, physics education, pedagogy, ICT, communication, class management, team-work, etc…) requested to teachers and school system in order to cope with the changes in society and in students' interests and attitudes; the scarcity of resources devoted to programs for continuous teachers professional development (funds, design/implementation capabilities, effective evaluation, …); intrinsic inertia of well-established and ineffective teaching practices; insufficient implementation of validated innovations; … Other challenges are linked to realise concrete ways for: enhancing PTE with knowledge and active experiences of the most significant results of PER (e.g. common and robust learning difficulties, teaching rituals that may result in lack of understanding, …); experimenting the advantages and limits of Technology Enhanced Learning (TEL); constructing multi-faceted supports for both pre and in-service education and for on-field teaching in standard conditions and contexts. The complex problems of the current PTE are well represented in the Report 2010 of the National Task Force on Teacher Education in Physics, by American Association of Physics Teachers, American Physical Society, American Institute of Physics. It states that: "Except for a handful of isolated pockets of excellence, the national system of preparing physics teachers is largely inefficient, mostly incoherent, and massively unprepared to deal with the current and future needs of the nation's students…. Physics departments, schools of education, university administrators, school systems, state agencies, the federal government, as well as business and foundations, have indispensable collaborative roles to play so that *every high school student has the opportunity to learn physics with a qualified teacher*…. .... Science education in the United States lags well behind much of the rest of the world, and in some cases, the gap is growing.... more students than ever before are taking physics from teachers who are inadequately prepared. There is a severe shortage of qualified physics teachers…. many current physics teachers lack the content knowledge and focused pedagogical preparation with which to help their students most effectively: international assessments show time and again that U.S. students lag behind their counterparts in other industrialized nations.... the shortage of qualified teachers is especially severe for those students who take either conceptual physics courses or physics as a gateway to other sciences in high school….“ —a group of students that has experienced the largest increase in size in the last several years".

The 2008 TIMSS ADVANCED results on secondary school Physics teachers (10 countries: AM, IR, IT, LB, NL, NO, PH, RU, SI, SE) indicate a complex scenario. The main area(s) of education are Physics, Mathematics, Chemistry, Engineering, Biology, Education (in Italy: 40 % Phys, 50 % Math, 10 % Eng). The requirements for being a teacher are diverse: Bachelor; Master plus Education course; Certificate Higher Education; Physics studies plus Education plus one year of supervised teaching. The collaboration with teachers of other disciplines varies much, from almost never to 2–3 times/month to once a week (in Italy: about 46, 49, 5 %). The book is still the main educational tool used, in about 100 % of the surveyed countries; in more than half of the time in school the students read "theory" or how to do exercises. The demonstrations of experiments ex-cathedra are common and vary from 11 to 54 %, experiments or investigations done by students from 0 to 30 %, use of calculators and computers from 0 to 50 %. In our experience the difficulties associated to an effective use of ICT and TEL based learning environments are less and less linked with shortage of hardware and software, the insufficient educational competencies for an effective use being the real bottleneck.

Answers to the above challenges can come from the links between PTE and PER; PTE is a research field which has been addressed in depth since long time. The main topics dealt with are: Teaching–Learning Processes (teachers' naïve epistemologies, deeply rooted and ineffective practices, common and robust learning difficulties linked to students' (and sometime teachers') naïve ideas and reasoning that conflict with Subject Matter Knowledge; …); ways to effectively transform Subject Matter Knowledge in Pedagogical Content Knowledge, focusing on its construction in PTE programs; the role of lab-work and the associated approaches, proposals and materials; modelling and simulation activities, support by multi-media; validated strategies to encourage/implement active and critical learning (e.g. Prediction–Experiment–Comparison learning/teaching cycle); experiential modality of a PTE activity (to do personally and in detail what will be proposed to the students); critical analysis of transformations of research-based proposals made by teachers and students in standard contexts and conditions; models and experimentation of prototypes of PTE programs (pre and in-service), etc… For sake of brevity it is not possible to discuss at length all these aspects.

4.3 Research-Based PTE Interventions

The rationale of a research-based PTE program is multi-dimensional. A not exhaustive list of key points has two levels. (A) the integration of different knowledge domains such as *(i) topical knowledge* about specific topics: crucial concepts about a phenomenon, its regularities, aspects, interpreting model(s), laws, applications, design/run of experiments; *(ii) net-worked knowledge* that links various types of concepts and skills; *(iii) meta-knowledge* i.e. the capability to build new theoretical and experimental knowledge); *(iv)* the acquisition of multi-competences (e.g. application of knowledge, independent learning, analytical and computational capabilities; ability for criticism, synthesis, communication and teamwork). (B) Focus on: -not-yet-much-common teaching methods (e.g. student centred and open learning environments, problem and project based procedures, peer instruction, …); -experimental activities via various types of lab-work (in presence with real-time sensor-based experiments and ready-to-go apparatuses, remote-controlled-experiments, virtual lab); -support by multi-media (extensively interpreted in); -links amongst phenomenological observations, data, abstract formal representations, modelling activities and theoretical reflections; -strategies to feed and enhance students' interest and motivation.

Hereafter we briefly discuss four emblematic examples of TE in Physics. The first two have been designed and implemented at University of Udine on energy for perspective primary teacher education ([2]) and on quantum mechanics basic concepts for in-service teacher in upper secondary school ([3]). The others are Projects and initiatives taken by international entities: Physwere by ICPE and Muse by EPS-PED.

4.3.1 Energy Intervention Module for Prospective Primary Teachers Education (FIME)

The Formative research based Intervention Module about Energy (FIME) was proposed in two different groups of 250 Prospective Primary Teachers (PPT), 21–22 years old, in two years at University of Udine. FIME include a preliminary subject centered part (CK) and an innovative proposal about energy for primary school, based on simple qualitative exploration and inquiry strategy by means of tutorials. To educate PPT to the Energy concept two different kinds of problems have to be overcome: (a) the lacks in the disciplinary knowledge [24, 27, 47] and in particular the identification of energy as a state property of a system; (b) the way of thinking to the pedagogical approaches only related to forms of Energy and to Energy sources, typically adopted in the textbook.

The sample was composed by N = 101 PPT in the first experimentation and by 143 PPT in the second one. An additional Conceptual Lab activity involve 37 of the PPT of the second group.

The first year FIME was organized in the following parts: (1) Pre/questionnaire (1 h), (2) Discussion on the foundation in physics of the concept of energy in traditional way and analysis of the main concepts and consequences related to: kinetic energy theorem, energy conservation principle; the first thermodynamic principle (4 h), (3) Collection of the questions posed by PPT on energy and relative discussion (1 h), (4) Presentation and discussion on the rationale of the research based proposal on energy developed [21], with illustration of the simple everyday experimental apparatus and explorative activity (4 h), (5) Post/questionnaire (1 h).

The post-questionnaire composed by 15 open ended questions was proposed to the PPT after the instruction to evaluate the PCK, during the final examination. The questionnaire was designed on the following main conceptual knots emerging from literature: energy associated to human or living being, as fuel-like substance which is possessed by living things; energy possesses only by moving objects (Stead 1980. Watts 1983) or as product of some process and existing only during this process ([35]; Watts 1983; [11]); energy; energy as force or power (Trumper 1983, [10]); different forms of energy and recognition of the form associated to standing objects (Brook and Wells 1988; Carr and Kirkwood 1998); conservation of energy (Duit 1981; Watts 1983; Black and Solomon 1983. Brook and Driver 1984. Driver and

[2] Heron et al.[20]

[3] Michelini M, Santi L, Stefanel A (2011), Teacher discussion of crucial aspects, cardinal concepts and elements peculiar to Quantum Mechanics starting from an educational proposal, in Battaglia et al. [3]

Warrington 1984, [46]; transformation of energy and process (Carr and Kirkwood 1998; Gilbert and Watts 1983; Duit [11]; Trumper [49] Dawson-Tunik 2005).

The format was a set up in two parts: (I) a CK question; (II) the typical answers of 4–5 students to the CK question posed and the request to describe the characteristics of the students' ideas and "what teacher have to do" for each student.

From data analysis it emerged that 87 % of PPT use the types of energy to give an appropriate description of simple processes in terms of kinetic, potential, internal energy; only in few cases 30–35 % are present difficulties in distinguishing potential energy and internal energy, in some case the energy associated with light. Analogous percentage we obtain for what concern the identification of energy as quantity that is transformed from a form to another and that is conserved. Concept of transformation and conservation are often associated ("because it is transformed"). For a group of PPT (about 40 %) the transformation is in any case associated to a dispersion or a loss of energy. This results evidence an important modification in the initial conception about energy with respect the results of the pre- questionnaire. Another picture emerge, when the Pedagogical Content Knowledge (PCK) is considered, analyzing in particular the teacher plan on how propose energy to the pupils, how they organize the topic for school, and ideas expressed in oral examinations. A large majority of PPT recover the initial ideas and conception when they have to think educational activity and paths for pupils. For instance about 72 % mentioned as a first goal the wrong definition of energy, frequently proposed in the textbooks: "Energy is the capacity to do work". At the same time one of the most diffused aim was to teach to pupils that "It exists in different forms: nuclear, kinetic, thermal……", without any distinction between type and forms. About this point the more evident change was that a large majority includes forms of energy: kinetic, potential, internal and the usually quoted energy forms related to sources solar, hydroelectric, nuclear, wind energy. Also an ambiguous language was used in some case (es example "it is transformed (for instance in the movement of a turbine)") or an assertive approach was used (for instance: "It is conserved <<nothing is destroyed, nothing is conserved>>"). This results confirm that a reconstruction of the concepts and a proved CK do not produced effective changes in the pedagogical organization of the educational proposals..

For this scope, in the second year the FIME was restructured to include personal reflection and successive group discussion on the main conceptual nuclei and knots about energy. For a group of 37/143 PPT a PCK lab was carried out using papers on learning problems taken by literature for design based educational path proposals by PPT. PCK-lab imply a personal reflection activity on CK aimed to discuss pedagogical aspects.

The main Research Questions in second year of FIME study were how sort of contribute to the PCK formation on Energy produce in perspective primary teachers a strategy based on:

1. exploration of an innovative teaching/learning proposal on energy?
2. personal involvement in the analysis of conceptual knots and learning questions combined with a pear to pear discussion?
3. What kind of PCK?

From an intermediate questionnaire emerge that improvement was obtained in the CK: Energy is identified as a state properties of a systems (83 %), that can be transformed (87 %) and the process is described in terms of actions and in terms of energy (75 %). A great personal involvement was needed to transform the content competences in professional attitudes. In this process the role of peer collaboration and of idea comparison were relevant. In particular a personal involvement in the analysis of PCK questions combined with a peer to peer discussion build gradually effective PCK on energy for the small group of PPT (37). A relevant contribution comes from the knowledge of typical students learning problems on the topic and how it is possible to face in school classroom (32/37).

Data emerging from the questionnaire answers are crossed with those obtained by tutorial worksheets filled during the PCK lab, the portfolios of the prospective teachers and the discussion in large group about the educational path. Results evidenced a relevant and generalized increasing about the CK, as well the PCK. The way to build and monitor the PCK competences and to act for their improvement for PPT appear to be fruitful in the identification of the way of thinking (34/37). The integration of research results in the FIME offers the opportunity to enrich the formative module not only with respect to the CK competences, but also for those of PCK.

4.3.2 Research Based Quantum Mechanics Formation for in Service Teachers

This second example of research based intervention module is on in-service teacher (IT) education on quantum mechanics in the framework of the Master on Didactic Innovation in Physics Education and Guidance (Master IDIFO) for in-service teacher education, now at 4th edition from 2006. Master IDIFO4 is instituted at the University of Udine with the cooperation of more than 20 Italian Universities as a two years activity for 60 cts organized in blended modality, being the main part in e-learning with intensive workshops on campus. For e-learning activities a specific web environment was developed. The formative activities are structured in four Training Areas (general, characterizing, project-oriented and on site) which are set out in five thematic Modules: (A) quantum physics (18 cts); (B) Relativity (12 cts); (C) statistical physics and material sciences (15 cts); (D) nuclear physics, particle physics and cosmology (2 cts); (E) Formative guidance and problem-solving as an operative challenge for guidance (6 ctsfu). Each Module include: (a) e-learning formation done by a responsible of the specific course included in a Module, by means of the material that has been selected and assessed according to research outcomes (30 cfu); (b) experimental laboratory activities (4 cfu); (c) three intensive on campus Workshops of 6 cts (approximately 60 hours) at the University of Udine; (d) planning activities for teaching / learning intervention on didactic innovation (7 cfu); (e) teaching–apprentiship: didactic experimentation activities: four activities of at least 6 h on Modules A, B, C & D, E respectively (7–11 cfu).

A particular focus in each course is on the discussion of didactic proposals, the analysis and evaluation of results related to research questions brought to light by didactic research into the various themes under investigation: individual and group discussion has been favoured.

The formative model individuated with the Master IDIFO integrate cultural, disciplinary, didactic and professional aspects. It is considered by the teacher-students participating as the most efficacy and corresponding to their needs. In particular it combines 'metacultural' approach with an experience-based, on situated ([4]) training method, offering each person the chance to develop a project according to his or her needs and motivation.

The rationale for teaching /learning paths in quantum physics for upper secondary school is widely discussed. The Dirac approach to quantum mechanics (QM) was discussed with teachers starting from an educational proposal for secondary school. The QM way of thinking was analyzed in a community of teachers and researchers. The research on teacher education carried out in this framework was focused on the following questions:

RQ1. Which are the most problematic knots regarding QM in a group of high level teachers?
RQ2. Which difficulties one faces in a teacher education based on a new proposal of QM teaching?
RQ3. Which learning paths results more effective for a real improvement of PCK?
RQ4. How do teachers modify the proposal of reference when asked to design a didactical path?

The 22 teacher involved had a long teaching experience, except for one, that is employed in an optic industry. The educational module is subdivided into three main steps:

1. Course A focused on the presentation and discussion in web forum of the knots on which the proposal of reference is developed and of the working sheets which are integral part of it.
2. In-person meeting for discussing with the teachers on the rationale of the proposal itself and the unsolved knots remaining after the web forum discussion.
3. Course B constituted by a web didactical laboratory, aimed at designing a micro module focused on the reference proposal analyzed the previous steps.

In a pre-questionnaire teachers have to list three topics on QM of particular interest for high school students, explaining the reasons of the choice (Q1) and to two elements characterizing quantum mechanics behavior with respect to classic mechanics, explaining the reasons of the choice (Q2). The maps produced by teachers for the educational design at the end of the activities were analyzed on the light of the aspects emerged in Q1 and Q2 and the discussed aspect in the community.

In the data of Fig. 4.1 the radical change in contents considered important by IT for the developing of the didactical proposal is clear. In the lists of tasks Q1 and Q2

[4] Michelini [30]

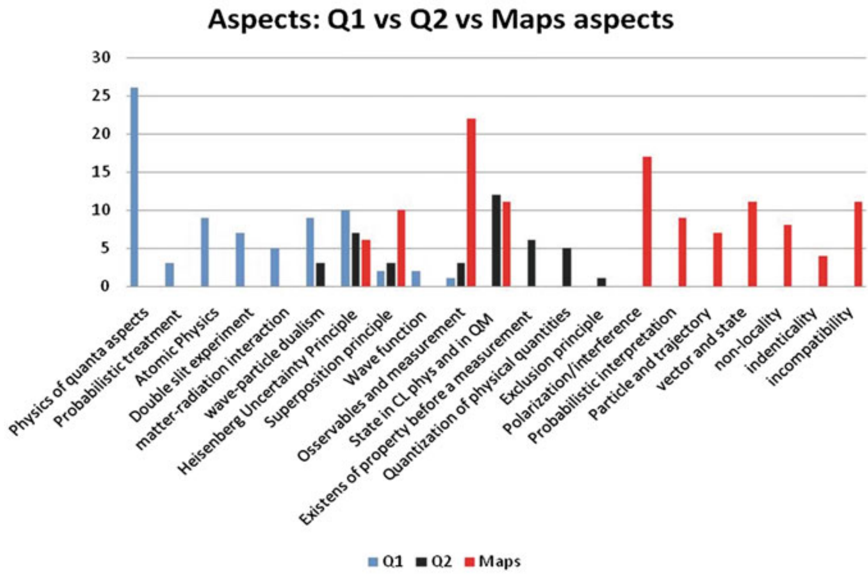


Fig. 4.1 Aspects emerged by Q1 and Q2 inquiry and from Maps prepared by IT to describe the rationale of the planned paths for a teaching/learning intervention on QM

the predominant category refers to aspects of quantum physics but few of these are characterizing elements of QM as a theory; in the categories emerged in the final maps only the basic aspects of the theory are emerged. What can be clearly seen is the pre-eminence of the polarization context, underlining that for 12 teachers the reference context stayed the one of the polarization, for other five the context of diffraction and spin (not present in the diagram because its frequency is two) are also introduced.

From data it emerged that even well prepared teachers have a vision of the teaching of QM physics of quanta oriented (RQ1). The indetermination principle is considered a key one in QM, likewise, for many other student-teachers, the quantization of physical observables (discrete spectrum).

The main difficulties encountered (RQ2) on the learning path are about leaving context usually explored with high school students, in particular the indetermination principle for position and impulse, the context of free propagation, rather that contexts of two-state systems which are simpler.

The elements of the learning path that led to the main changes in PCK (RQ3) are: (a) the didactical proposal offered for discussion focused on facing the basic concepts of QM and on the detailed analysis of the instruments, (b) the rich exchange developed on the web between the student-teachers and between them and the tutors about the different basic concepts of QM followed in the reference didactical path, (c) the direct involvement in the construction of conceptual and organization maps of contents and work modalities. In particular, the main changes are the passing from a vision focused on the physics of quanta, to a vision in which a central role in the quantum theory is covered by the superposition principle.

In the designing of personal didactical paths, even if the superposition principle plays a basic role for most of the student-teachers, the attention is mainly focused on the measurement theory and on the concept of state. In the specific case of high level competence student-teachers, the main integrations are about other phenomenological contexts with a similar approach to the one proposed, like the one of the spin and the diffraction phenomenology (RQ4).

4.3.3 PHYSWARE Model Workshops

The third example is PHYSWARE, by International Commission of Physics Education (ICPE), held at International Centre for Theoretical Physics, Trieste Italy (2009, 16–27 Feb) http://cdsagenda5.ictp.trieste.it/full_display.php?ida=a07137. It has been "designed to enhance the quality of physics education at the tertiary level, especially in the developing countries, conceptualized as a series of model workshops and resource materials for physics teachers and teacher trainers that exemplify how active learning methods can be adapted to meet the needs of students in developing countries , to provide materials at the undergraduate level using affordable hands-on equipment that can be locally adapted by teachers and their students, to offer exposure to appropriate technologies and computer-based tools for enhancing conceptual understanding ..., to provide a forum to the teacher-leaders to share experiences and exchange ideas about dissemination of active learning methods". Newtonian Mechanics was the theme for PHYSWARE 2009. Out of more than 200 applications from 48 countries, 35 participants were selected from 27 countries in Africa, Asia, Latin America and Europe. The ten working days workshop has had four session of 1.45 h on each day plus seven 2 h post dinner sessions for posters and discussions, involving the participants in research-based conceptual tests , diagnostic tools and learning cycles promoting active engagement. The first week activities focused on lab-work and class activities using no-cost, locally available materials (e.g. pendula of different lengths as clocks to measure time in arbitrary units, a mahogany flower pendulum to study damping). In the second week the participants worked collaboratively on didactic projects using motion and force sensors, photo-gates, video clips and simulations. All materials used are freely downloadable at the above site. Issues of multicultural and multiethnic classroom were also discussed. A Discussion Group and a Blog have been realised by the participants in addition to the PHYSWARE Workshop site at the ICTP portal and the Wiki created by the directors. Feedback from participants' evaluation has been extremely positive. A five year action plan with ICTP has been agreed, for workshops to be held in Trieste and in developing countries.

4.3.4 MUSE Project

The last example is MUSE (More Understanding with Simple Experiments), a project started in 2008 in the framework of European Physical Society—Physics Education Division (EPS-PED) http://education.epsdivisions.org/muse . The purpose of MUSE is to contribute to: awareness of relevance of physics in nowadays culture; interaction amongst school and university; better quality of physics teaching/learning by addressing physics and teacher education, new methods/ practices, differences/similarities in the European educational systems; ... MUSE offers research-based, free-downloadable materials (nine up-to-now). The Added Value in Education (AVE) addresses: cognitively dense and easy-to-assemble experiments using low-cost and easily available materials; Prediction Experiment Comparison learning cycle; variation approaches (what happens if … is changed?); identification/analysis of diverse viewpoints; interactive cognitive dynamics via peer learning; naïve ideas/reasoning conflicting with physics knowledge; learning difficulties studied by PER and plausible underlying reasons; teaching rituals resulting in misleading argumentation. The audience aimed at are teachers, the communities of Physics Education, Physics Teacher Education, Physics Education Research, Educational Authorities et al. To present the MUSE approaches with in-presence activities, two workshops held at the GIREP-EPEC 2011 Conference in Finland have involved about 35 participants with success and interest.

4.4 Some Recommendations for PTE

They can be grouped in three main areas. The first area deals with "experiential modality", i.e. teachers have a personal experience of the situations they will propose to the students as: strategies, approaches, methods, activities, tools, assessment. The goal is to experience in terms of "hands and minds on" with the most didactically effective experimental methods and techniques, modelling and simulation activities, student-centred learning environments, structured collaborative projects. The second area has to do with the time-scale of PTE that most often is concentrated in a small numbers of (isolated) training episodes rather than being designed as a continuous process lasting for the duration of the teaching activity thanks also to the opportunities offered by ICT and TEL. Communities of practice foster the process of sharing common problems and their solutions. Commented repositories of best practices, developed in standard contexts, allow sharing patrimonies of knowledge, expertise and innovation. This process of autonomous education, together with a series of in-presence episodes, can realise a continuous professional development program. The key words are: cooperation, collaboration, synergy amongst school, educational agencies, Universities.

The third area refers to the many contributions offered by PER to acquire crucial competences, as. e.g., to re-build zones of SMK, to construct PCK, to address naive ideas/reasoning conflicting with disciplinary knowledge, to integrate diverse types of knowledge, to effectively and wisely use the increasingly research-based and technology-based proposals. Modern physics in secondary school require innovation in contents, strategies and methods and PTE imply a change in the way of thinking professional work by teachers: a long process is necessary and teachers have to be supported. The challenges are many, great and complex, deriving from the various problems/difficulties affecting PTE; therefore they are appealing and call for a great, focused effort by the communities of Physics Education, Physics Teacher Education, Physics Education Research and Educational Authorities.

References

1. AJP (2002) Am J Phys 70 (3).
2. Appleton K (2006) Science pedagogical content knowledge and elementary school teachers. In: Appleton K (ed) Elemetary science teacher education: international perspective on contemporary issues and practice. Elbaum, Newyork, pp 31–54
3. Battaglia RO, Cazzaniga L, Corni F, De Ambrosis A, Fazio C, Giliberti M, Levrini O, Michelini M, Mossenta A, Santi L, Sperandeo RM, Stefanel A (2011) Master IDIFO (Innovazione Didattica in Fisica e Orientamento): a community of Italian physics education researchers for a community of teachers on modern physics. In: Raine D, Hurkett C, Rogers L (eds) Physics community and cooperation, vol 2, selected papers in GIREP-EPEC & PHEC Book, Leicester, Lulu [ISBN: 978-1-4466-1139-5], pp 97-136.
4. Bransford J, Brown A, Cocking R (2002) How people learn. National Research Council, Washington DC
5. Brook AJ, Wells P (1998) Conserving the circus? An alternative approach to teaching and learning about energy. Phys Educ 23(1):80–85
6. Carroll TG (2000) If we didn't have the schools we have today, would we create the schools we have today? Contemp Issues Technol Teach Educ 1(1):117–140
7. Colonnese D (2009) Energy in primary, Master thesis of DidSciUD-Michelini M responsible. University of Udine, Udine.
8. Davis EA, Petish D, Smithey J (2006) Challenges new science teachers face. Rev Educ Res 76(4):607–651
9. Dawson-Tunik T, Stein Z (2004) It has bounciness inside! developing conceptions of energy. http://devtestservice.com/PDF/Bounciness.pdf. Accessed Dec 2008
10. Driver R, Warrington L (1985) Students' use of the principle of energy conservation in problem situations. Phys Educ 20:171–175
11. Duit R (1984) Learning the energy concept in school-empirical results from The Philippines and West Germany. Phys Educ 19:59–66
12. Duit R, Gropengieβer H, Kattmann U (2012) Toward science Education research that is relevant for improving practice: the model of educational reconstruction.
13. Enghag M, Niedderer H (2008) Two dimensions of student ownership of learning during small-group work in physics. Int J Sci Math Educ 6(4):629–653
14. Eylon BS, Bagno E (2006) Research-design model for professional development of teachers: designing lessons with physics education research. Phys Rev Special Topics, Phys Educ Res 2:02016-1–14
15. Fisher HE (ed) (2005) Developing standard in research on science education. Taylor and Francis, London, pp 1–9
16. Ghirardi GC, Grassi R, Michelini M et al (1995) A fundamental concept in quantum theory: the superposition principle. In: Bernardini Carlo (ed) Thinking physics for teaching. Plenum Publ, Corp, Aster, pp 329–334

17. Gilbert J (2005) Constructing worlds through science education: the selected works of John K Gilbert. Routledge, Softcover
18. Osborne J, Goldring H (1994) Students' difficulties with energy and related concepts. Phys Educ 29.
19. Griffiths T, Guile D (2003) A connective model of learning: the implications for work process knowledge. Eur Educ Res J 2(1):56–73
20. Heron P, Michelini M, Stefanel A (2011) Evaluating pedagogical content knowledge of energy of prospective teachers. In: Raine D, Hurkett C & Rogers L (eds) Physics community and cooperation, vol 2, selected papers in GIREP-EPEC & PHEC Book, Leicester, Lulu [ISBN: 978-1-4466-1139-5], pp 147-161.
21. Heron P, Michelini M, Stefanel A (2008) Teaching and learning the concept of energy in primary school. In: Costantinou C (ed) Physics curriculum design, development and validation.
22. Holbrook J, Rannikmäe M (2001) Introducing STL: a philosophy and teaching approach for science education: ICASE, pp 1–19.
23. IDIFO (2006) http://idifo.fisica.uniud.it/uPortal/render.userLayoutRootNode.uP
24. Kirkwood VM, Bearlin M, Hardy T (1989) New approaches to the in-service education in science and technology of primary and early childhood teachers. Res Sci Educ 19:174–186
25. Kirkwood VM, Carr M (1989) A valuable teaching approach: some insights from LISP.
26. Kouhila M (2000) Formation en épistémologie de la physique à l'ENS. Didaskalia 17:173–202
27. Kruger C (1990) Some primary teachers' ideas about energy. Phys Educ 25:85–91
28. Magnusson S, Krajcik J, Borko H (1999) Nature, sources and development of pedagogical content knowledge. In, Gess-Newsome J, Lederman NG (eds).
29. McDermott L (1991) Millikan lecture 1990: what we teach and what is learned. AJP 59:301–315
30. Michelini M (2004) Quality development in the teacher education and training. Girep book of selected papers, RT1, Forum, Udine, pp 127–132 [ISBN: 88-8420-225-6].
31. Michelini M, Ragazzon R, Santi L, Stefanel A (2000) Proposal for quantum physics in secondary school. Phys Educ 35(6):406–410
32. Michelini M (2003) New approach in physics education for primary school teachers: experimenting innovative approach in Udine University. In: Ferdinande H, Valcke E, Formesyn T (eds) Inquiries into European higher education in physics, European Physics Education Network (EUPEN), vol 7, p 180.
33. Mikeska JN, Anderson CW, Schwarz CV (2009) Principled reasoning about problems of practice. Sci Educ 93:678–686
34. Millar R (2005) Teaching about energy. Department of Educational Studies Research Paper 2005/11. York University.
35. Nicholls G, Ogborn J (1993) Dimensions of children's conceptions of energy. Int J Sci Educ 15:73–81
36. Petri F, Niedderer H (1998) Int J Sci Educ 20(9):1075–1088
37. Physed (2000) Special issue, Phys Educ 35(6) IOP.
38. Pinto' R (2005) Introducing curriculum innovation in science: identifying teachers' transformations and the design of related teacher education. Sci Educ 89(1):1–12
39. Posner GJ, Strike KA, Hewson PW, Gerzog WA (1982) Accommodation of a scientific conception: toward a theory of conceptual change. Science Education 66(1982):211
40. Pospiech G (2000) A modern course in quantum physics for teacher education. In: Xingkai L, Kaihua Z (eds) Turning the challenge into opportunities. Guangxi Normal University Press, Guilin, pp 244–248. ISBN 7-5633-3054-2
41. Psillos D, Niedderer H (eds) (2002) Teaching and learning in the science laboratory. Kluwer Academic Publishers, Dordrecht, pp 21–30
42. Shön DA (1988) Educating the reflective practitioner. Jossey-Bass, San Francisco
43. Shulman LS (1986) Those who understand: knowledge growth in teaching. Educ Res 15(2):4–14
44. Shulman L (1986) Those who understand: knowledge growth in teaching. Educ Res 15(2):4–14

45. Simons PRJ (1996) Metacognitive strategies: teaching and assessing. In: Anderson LW (ed) International encyclopaedia of teaching and teacher education. Elsevier Science, Oxford, pp 325–342
46. Solomon J (1985) Teaching the conservation of energy. Phys Educ 20(2):165–170
47. Spirtou A, Koumaras P (1994) Educating primary teachers in science: the case of energy. In: Lijnse PL (ed) European research in science education-proceedings of the first Ph. D. Summerschool, pp 259–265.
48. Tiberghien A, Jossem EL, Barojas J (eds) Connecting research in physics education with teacher education, vol 1, An I.C.P.E. Book International Commission on Physics Education 1997 1998 http://www.physics.ohio-state.edu/ jossem/ICPE/TOC.html
49. Vicentini M, Sassi E (eds) Connecting research in physics education with teacher education, vol 2, An I.C.P.E. Book International Commission on Physics Education 2008 http://web.phys.ksu.edu/ICPE/Publications/teach2/index.html
50. Viennot L (1995) The contents of physics: essential elements, common view. In: Bernardini C, Tarsitani C, Vicentini M (eds) Thinking physics for teaching. Plenum Press, New York
51. Viennot L, Francoise Chauvet P, Colin G, Rebmann (2005) Designing strategies and tools for teacher training: the role of critical details. Examples Optics Sci Educ 89(1):13–27
52. Trumper R (1993) Children's energy concepts: a cross-age study. IJSE 15(2):139–148

Part II
Astronomy and Astrophysics

Chapter 5
Simulation of High Energy Emission from Gamma-Ray Bursts

Houri Ziaeepour

Abstract Gamma-Ray Bursts (GRBs) are the must violent explosions after the Big-Bang. Their high energy radiation can potentially carry information about the most inner part of the accretion disk of a collapsing star, ionize the surrounding material in the host galaxy, and thereby influence the process of star formation specially in the dense environment at high redshifts. They can also have a significant contribution in the formation of high energy cosmic-rays. Here we present new simulations of GRBs according to a dynamically consistent relativistic shock model for the prompt emission, with or without the presence of an magnetic field. They show that the properties of observed bursts are well reproduced by this model up to GeV energies. They help to better understand GRB phenomenon, and provide an insight into characteristics of relativistic jets and particle acceleration which cannot yet be simulated with enough precision from first principles.

5.1 Introduction

The history of observation of exploding stars goes back quite a long time to 185 AD [1]. From this observations we have learned that the life of massive and intermediate mass stars - with a mass close or slightly higher than the Sun - ends with violent explosions, generally called supernovae. The progenitor of supernovae are divided to two main groups [2]: Old white dwarfs which arrive to a critical mass - Chandrasekhar limit about $1.38 M_{\odot}$ - by accretion of material from a companion (type Ia), and very massive young stars that collapse on themselves and depending on absence or presence of hydrogen line in their spectrum, are classified as type Ib/c or type II.

H. Ziaeepour (✉)
Max Planck Institut für Extraterrestrische Physik (MPE), Giessenbachstraße 1,
85748 Garching, Germany

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_5,

In 1960s spy satellites called *Vela* designed to detect x-ray, gamma-ray, and neutron from space and atmospheric nuclear tests observed flashes of gamma-ray of extra-solar system origin [3, 4]. Distribution of their duration shows a clear grouping of bursts to short with duration $\lesssim 2$ sec and long with duration $\gtrsim 2$ sec. Their occurrence at cosmological distance and their association to supernovae and explosion of stars was first suggested in 1986 by B. Paczynski [5]. The short bursts are believed to have been generated in the collision of compact objects such as two neutron stars or a neutron star and a black holes, and long bursts in the core collapse of massive stars.

Motivated by the absence of detection in other wavelengths and by compactness of the source (see e.g. [7]), a *fireball* of strongly interacting $e^{\pm}$ plasma ejected during the explosion has been suggested as the origin of these Gamma-Ray Bursts (GRBs) [5, 6]. In this model the annihilation of $e^{\pm}$ to photons is assumed to be the origin of detected gamma-ray emission. But this model has various problems. For instance, it is difficult if not impossible to explain the Fast Rise Exponential Decline (FRED) shape of the peaks, their randomness, and long-lasting afterglow which has been observed since 1998 for majority of bursts, thanks to angular resolution new gamma-ray telescopes such as BATSE, Swift, and Fermi, multi-wavelength detectors on board of the Swift and Fermi satellites, and fast slew ability of ground based telescopes. Also it cannot explain the power-law spectrum of observed bursts and the lack of a thermal emission with a temperature $\sim$1 MeV.

In the internal shock model, Synchrotron Self-Compton (SSC) emission produced by collisions between shells inside a relativistic ejecta are considered to be the origin of observed prompt gamma-ray [8]. Similarly, the afterglow in lower energies is assumed to be produced by the collision of the remnant of the jet with circum-burst material or the Inter-Stellar Material (ISM). Other models such as a flow of magnetized plasma - a Poynting flow - is another popular model for GRBs [15]. In this model the gamma-ray is emitted by electrons accelerated by reconnection of magnetic field lines. Variants and combination of these models are also suggested by various authors to solve some of the short comings of these models.

None of these models is completely flawless. As mentioned above the spectrum of GRBs is not consistent with a close to thermal spectrum predicted by a standard fireball model. The Poynting flow model cannot explain in a natural way the fast variation of GRB emission because the frequency of reconnection is expected to be very low. SSC that is the most favorite model of GRB emission has also various issues: To have a sufficiently hard emission the magnetic field must be significant such that the emission from most popular electrons with a Lorentz factor close to the minimum γ_m that make the peak of spectrum be enough hard. This makes the duration of emission of single electrons very short and is known as fast cooling problem. Therefore it seems that SSC is not able to sustain long bursts. More seriously, synchrotron theory predicts a spectrum index $\alpha \sim -4/3$ at $E \ll E_{peak}$ [16], but observations show softer distribution with $\alpha \gtrsim -1$ at lower wing is observed. [17]. Recently, observations by the Fermi satellite up to energies $\sim$100 GeV have detected a high energy component in both short and long bursts that is delayed by up to few tens of seconds in long bursts from $E \sim 100$ MeV component. It fades much slower than lower energies. Finally, SSC has a small efficiency. Particle In Cell (PIC) simulations show that only $\lesssim 10\,\%$ of the total kinetic energy is transfer to electrons [18]. At present PIC simulations are not yet able to simulate GRB emission from first principles. In this proceedings we review an approximate but realistic formulation of SSC in the context of relativistic shocks model [9–11]. The aim of this exercise has been to see if despite issues discussed above internal shock-SSC model can explain observations. We also extend the model by considering an external precessing magnetic field to explain coherent oscillations observed in GRB 090709A and with less significance in other bursts. Then we present light curves and spectra of a number of simulated bursts according to this approximation. We show that due to rapid variation of physical quantities, even in presence of a precessing field, little evidence of coherent oscillation is imprinted in the emission. This explains the lack of observation of a significant oscillatory component in the light curves of GRBs. In this proceedings we present a summary of physical processes and motivations of the approximations and parameters used in our model, as well as its formulation. Details can be found in [9, 11].

5.2 Synchrotron Emission by Relativistic Shocks

In the framework of internal shock model, collisions between shells of material with different densities and velocities ejected by a central source produce mildly relativistic shocks. They are assumed to be cold and baryon dominated. Apriori there is no reason why faster shells should be ejected later, nonetheless velocity segregation can be automatically generated by deceleration of the front shells when they interact with surrounding material specially in Wolf-Rayet (WR) stars - the candidate progenitor of long GRBs [12]. Weak precursors observed in many bursts can be due to this process [13].

During a collision compression of the particles behind the shock front and turbulence create transverse electric and magnetic fields and produce what is called Electromagnetic Energy Structure (EES) - a solitinic electromagnetic wave across the shock front. Particles of the slow shell fall along a helical path into the shock front and are accelerated by this field and by the short range random fields through Fermi processes. However, their penetration distance in the fast shell (upstream) is very short and they are reflected to down stream. During this deceleration they emit a fraction of their kinetic energy due to the presence of the shock induced magnetic field. The presence of an external magnetic field both helps the acceleration of electrons [19] and as we see below the emission of synchrotron radiation. This process is continuous i.e. electrons move back and forth across the shocked zone and in this way, dissipate the kinetic energy of the fast shell through synchrotron emission in places where the induced magnetic field is strong and transversal. There is phase shift in the EES between electric and magnetic field. Its presence is crucial for SSC process in general, and for understanding the origin of high energy delayed tail in particular. The lifetime this acceleration-dissipation process is short because in a neutral plasma for each electron that falls to the shock front, one or more baryons—

protons and neutrons—which are more massive fall too. For an observer in the rest frame of the slow shell the absorption of these baryons by the fast shell slow it down, reduces the discontinuity—the shock—and the strength of EES. We do not consider a significant internal energy for the shells and assume that the turbulence and mixing transfer the energy from fast shell to slow shell by elastic scattering. In this sense the collision is radiative, i.e. all the energy excess is radiated out.

One can distinguish two *shocked zone* in the opposite side of the initial discontinuity. If the velocity of massive particles—presumably baryons—are reduced to relativistic sound speed in the upstream, a secondary *reverse shock* front will form which propagates in the opposite direction of the main *forward shock*. Although it was expected that the difference between the dominant synchrotron frequency and time evolution of emission from forward and reverse shocks make their separation possible, multi-band and early observations of GRBs have shown the contrary. Therefore in the present approximation we only consider one radiation emitting region and call it *the active region*. Note that what we call active region does not correspond to shocked material. In particular, its width initially increases to a peak value, then declines and at the end of the collision i.e. when the two shells are coalesced, it disappears. This is in contrast of shocked region which increases monotonically until shells are completely mixed.

To simplify the model further, we also assume that the thickness of this emitting region is small, i.e. the propagation time of photons in this region is smaller than time resolution of this model. In fact for objects moving with ultra relativistic speeds with respect to a far observer, time and distance are approximately proportional: $r'(t') = \beta'(t)ct' \approx ct'$.[1] Under these approximations evolving quantities only depend on the average distance of the active region from central engine. Mathematically, this approximation is equivalent to assuming a wavelike behaviour for dynamical quantities i.e they depend on $r' - c\beta' t'$ rather than r' and t' separately. When $\beta' = const$, i.e. when there is no collision or dissipation, this is an exact solution. In this case the solution at every point can be obtained from the solution of one point.

In the standard treatment of SSC models [8, 16] a simple power-law distribution is assumed for the Lorentz factor of accelerated electrons $n'_e(\gamma_e) = N_e(\gamma_e/\gamma_m)^{-(p+1)}$ for $\gamma_e \geqslant \gamma_m$. The parameters used for the phenomenological modeling of the shock and SSC emission such as fractions of total kinetic energy transferred to accelerated electrons ϵ_e and to a transversal magnetic field ϵ_B are also considered to be fixed. However, in a phenomenon as fast evolving as a GRB these assumptions do not seem realistic. For this reason in our formulation it is assumed that ϵ_e, ϵ_B, and densities evolve with time. We also consider shorter distances for the collision between shells in the range of $\sim 10^{10} - 10^{12}$ cm rather than $\gtrsim 10^{14}$ cm considered in the literature. This leads to short lags between energy bands consistent with observations. The motivation for this choice is the detection of variabilities up to the shortest time

[1] Through this work quantities with a prime are measured with respect to the rest frame of the slow shell and without prime with respect to a far observer at the redshift of the central engine. Parameters do not have a prime even when the parametrization is in the slow shell frame.

resolution of present instruments—$\sim 10^{-3}$ sec—and the association of anisotropies to the accretion disk around the forming compact object, presumably a black hole or neutron star, in the center of the collapsing star.

As for the external magnetic field, it can have various origins: a precessing Poynting flow, magnetic field frozen in the plasma, the magnetic field of the central engine or its accretion disk, and the dynamo field of the envelop if it is not completely interrupted by the explosion. Evidently a combination of all these cases can be present. For this reason and also because our simple model cannot distinguish between these field, we do not specify the origin of the magnetic field in the mathematical formulation or in the simulations, and simply consider a precessing external field i.e. a field with an origin other than the shock. We have also inspired by findings of PIC simulations and use them as input and/or motivation in the choices of parameters and distributions. For instance, simulations of relativistic shocks of $e^{\pm}$ plasma show that the distribution of accelerated electrons is close to a power-law with exponential cutoff $n'_e(\gamma_e) = N_e(\gamma_e/\gamma_m)^{-(p+1)} exp(-\gamma_e/\gamma_{cut})$ or a broken power-law [18]. We also use the penetration distance of accelerated electrons in the slow shell $\lesssim \mathcal{O}(1) \times 10^3 \lambda_{ep}$, where λ_{ep} is plasma wavelength of electrons, in the calculation of Inverse Compton (IC) scattering of photons. This quantity is crucial because if we assume the same density in the whole slow shell, most of the photons would be scattered before they can leave colliding shells, leading to significant deviation of their spectrum from power-law with break expected from synchrotron emission and creates a double peak spectrum, in contradiction with observations.

5.3 Formulation

Conservation of energy and momentum determines the evolution of the shock. The velocity β' of the fast shell/active region decreases due to absorption of particles from slow shell and dissipation of kinetic energy as radiation due to synchrotron and self-Compton interactions. After a variable change the dynamic equations - energy-momentum conservation equations - for the active region can be written as:

$$\frac{d(r'^2 n' \Delta r' \gamma')}{dr'} = \gamma' \left(r'^2 \frac{d(n' \Delta r')}{dr'} + 2r'(n' \Delta r') \right) + r'^2 (n' \Delta r') \frac{d\gamma'}{dr'} = n'_0(r) r'^2 - \frac{dE'_{sy}}{4\pi mc^2 dr'} \tag{5.1}$$

$$\frac{d(r'^2 n' \Delta r' \gamma' \beta')}{dr'} = \beta' \gamma' (r'^2 \frac{d(n' \Delta r')}{dr'} + 2r'(n' \Delta r')) + r'^2 (n' \Delta r') \frac{d(\beta' \gamma')}{dr'} = -\frac{dE'_{sy}}{4\pi mc^2 dr'} \tag{5.2}$$

where r' is the average distance of the active region from central engine, n' is the baryon number density of the fast shell measured in the slow shell frame, n'_0 is the baryon number density of the slow shell in its rest frame and in general it depends on r'. Here we assume that $n'_0(r') = N'_0(r'/r'_0)^{-\kappa}$. For the ISM or thin shells where density difference across them is negligible $\kappa = 0$, i.e. no radial dependence. For a wind surrounding the central engine $\kappa = 2$. If we neglect the transverse expansion of the jet, for a thin shell/jet expanding adiabatically $\kappa = 2$ too. If the lifetime of the collision is short we can neglect the density change due to expansion during the collision and assume $\kappa = 0$. $\Delta r'$ is the thickness of the active region, γ' is the Lorentz factor of the fast shell with respect to the slow shell, $\beta' = \sqrt{\gamma'^2 - 1}/\gamma'$, $m = m_p + m_e \approx m_p$, E'_{sy} is the total emitted energy, and c is the speed of light. The evolution of the average radius of the shell is:

$$r'(t') - r'(t'_0) = c \int_{t'_0}^{t'} \beta'(t'')dt'' \tag{5.3}$$

where the initial time t'_0 is considered to be the beginning of the collision. Equations (5.1) and (5.2) can be solved exactly for the column density of the active region $n'(r')\Delta(r')$, but $\beta'(r')$ does not have an analytical solution. We use an iterative technique which allow to determine the solution as a function of the *coupling* defined as:

$$\mathcal{A} \equiv \frac{4\alpha m_p^2 \sigma_T n'_0 \Delta r'(r'_0) \epsilon_e^2(r'_0) \epsilon_B(r'_0)}{3m_e^2} \tag{5.4}$$

In presence of an external magnetic field a second interaction term is also present which its *coupling* can be defined as:

$$\mathcal{A}_1 \equiv \frac{\alpha m_p \sigma_T \Delta r'(r'_0) \epsilon_e^2(r'_0) B'^2_{ex\perp}(r'_0)}{6\pi c^2 m_e^2} \tag{5.5}$$

Details of this calculation is described in [9, 11]

For the determination of synchrotron flux we use textbook formulations. Nonetheless, we have to integrate over angular distribution of emission for the observer. Notably, we must take into account the fact that due to relativistic effects even the emission from a spherical shell seems highly collimated at far distances. The angle of collimation is $\sim 1/2\Gamma(r)$ where $\Gamma(r)$ is the Lorentz factor of the active region. The details of this calculation is discussed [9]. Finally, after integration over emission angle the expression for the received synchrotron flux is:

$$\frac{dP}{\omega d\omega} = \frac{4\sqrt{3}e^2}{3\pi} r^2 \frac{\Delta r}{\Gamma(r)} \int_{\gamma_m}^{\infty} d\gamma_e n'_e(\gamma_e) \gamma_e^{-2} K_{2/3}(\frac{\omega'}{\omega'_c}) + \mathcal{F}(\omega, r) \tag{5.6}$$

where $\mathcal{F}(\omega, r)$ includes subdominant terms and terms depending on the curvature of the emission surface. In [9] it is argued that these terms are much smaller than the dominant term, thus we neglect them in the simulations.

The advantage of the approximation presented here is that one can use the approximate analytical solutions to study the effect of various parameters and quantities on the evolution of dynamics of the ejecta and its synchrotron emission. However, the price to pay for this simplification is that we cannot determine the evolution of $\Delta r'(r')$ from first principles and must consider a phenomenological model for it. In our simulations we have used following phenomenological expressions:

$$\Delta r' = \Delta r'_{\infty}\left[1 - \left(\frac{r'}{r'_0}\right)^{-\delta}\right]\Theta(r' - r'_0) \quad \text{Steady state model} \tag{5.7}$$

$$\Delta r' = \Delta r_{\infty}\left[1 - \exp(-\delta' \frac{r'}{r'_0})\right]\Theta(r - r'_0) \quad \text{Exponential model} \tag{5.8}$$

$$\Delta r' = \Delta r'_0\left(\frac{\gamma'_0 \beta'}{\beta'_0 \gamma'}\right)^{\tau}\Theta(r' - r'_0) \quad \text{dynamical model} \tag{5.9}$$

$$\Delta r' = \Delta r'_0\left(\frac{r'}{r'_0}\right)^{-\delta}\Theta(r' - r'_0) \quad \text{Power-law model} \tag{5.10}$$

$$\Delta r' = \Delta r'_0 \exp\left(-\delta' \frac{r'}{r'_0}\right)\Theta(r' - r'_0) \quad \text{Exponential decay model} \tag{5.11}$$

The initial width $\Delta r'(r'_0)$ in the first two models is zero, therefore they are suitable for the description of formation of an active region at the beginning of internal or external shocks. The last three models are suitable for describing more moderate growth or decline of the active region.

5.4 Simulations

Table 5.1 shows the list of parameters of this model. Due to their large number it is not possible to explore the totality of parameter space. Therefore, in this section we show a few examples of simulated light curves and spectra of GRBs. Each simulation consists of at least 3 time-intervals (regimes) during which exponents are kept constant. Moreover, each time interval corresponds to a given model for the evolution of the width of the active region. The first regime must be either steady state or exponential for in which the initial width of the active region is zero. Following regimes can be either dynamical or power-law. Matching between values of evolving quantities at the boundary of regimes assures the continuity of physical quantities.

Figure 5.1 show few examples of simulated bursts without and with an external magnetic fields. It is clear that in presence of an external field the bursts are usually brighter, harder, and lags between the light curves of the various energy bands are in general smaller and more consistent with observations. Nonetheless, some simulated bursts have small lags even in absence of an external field. Therefore, its presence is not a necessary condition for the formation of a GRB. The light curves of these

Table 5.1 Parameter set of the phenomenological shock model

model	r_0 (cm)	$\frac{\Delta r'_0}{r'_0}$	p	α_p	γ_{cut}	κ	γ'_0	τ	δ	ϵ_B	α_B
ϵ_e	α_e	N'_0 (cm^{-3})	$n'\Delta r'(r'_0)$ (cm^{-2})	Γ_f	$\lvert B\rvert$ (Gauss)	f (Hz)	α_x	phase (rad.)	$(\frac{r'}{r'_0})_{max}$		

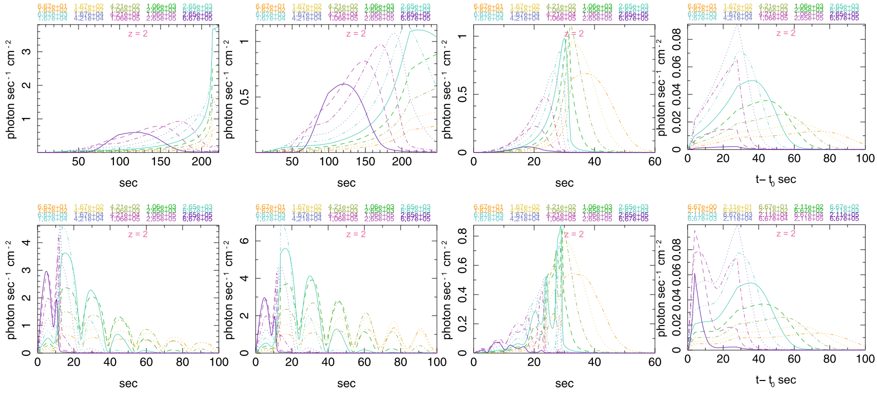


Fig. 5.1 *Top row* GRB simulations without an external magnetic field. Each simulation includes three regimes. The lags in the first two models from *left* is too large to be consistent with observations. *Bottom row* Simulations with a precessing external magnetic field. From *left* to *right* (1) $|B| = 100$ kG, $f = 0.2$ Hz (2) $|B| = 12$ kG, $f = 0.2$ Hz, and middle regime dynamical, (3) $|B| = 100$ kG, $f = 0.2$ Hz, and middle regime steady state, (4) $|B| = 2.5$ kG, $f = 0.1$ Hz. Other parameters are the same as simulations in the *top row*. In all plots of light curves energy bands are written on the top of the plot and the color of their font corresponds to the color of light curves of the band

examples show that despite the presence of a precessing field - similar to what is seen in pulsars and magnetars - there is barely any signature of oscillation in these light curves, specially in high energy bands. Only when the duration of the burst is a few times longer than oscillation period, the precession of the field creates a detectable periodic component in the light curves, see Fig. 5.1. We also note that oscillations are more visible in soft X-ray and UV/optical bands. However, only in the case of the detection of a well separated precursor low energy data are available during the prompt emission. Moreover, the necessity of binning of optical data to reduce the noise can smear fast oscillations, see Fig. 5.1. Therefore such oscillation can be hardly observed in real GRBs. On the other hand, in many burst e.g. GRB 070129 semi-periodic variations of the early time X-ray light curves, similar to two of the examples in Fig. 5.1, have been observed for a few hundreds of seconds.

Our model can be applied to both long and short bursts. An example of simulation of short bursts is shown in Fig. 5.2. The lags in this example is very short - consistent with observed zero lags - both in presence and absence of an external magnetic field. It also shows that very fast precession of the field can be confused with the shot noise. A remarkable properties of simulation with an external magnetic field is the presence of non-periodic substructures is the light curves from nonlinearity and fast

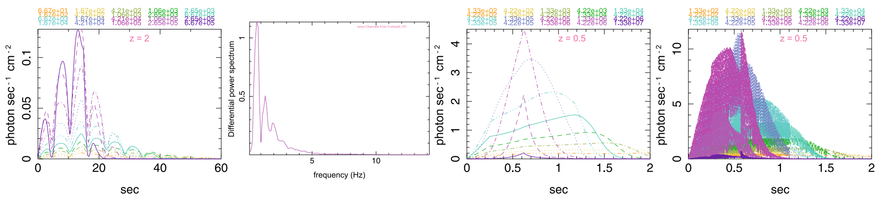


Fig. 5.2 From *left* to *right* (1) Simulation of a long burst with a precessing magnetic field and (2) its the power distribution function (PDF) of the total light curve. Oscillatory component in high energy bands is visible by eye and in the PDF. (3, 4) Simulations of a short burst without and with a fast precessing $f = 500$ Hz external magnetic field. The comparison of two plots show that due to nonlinearity of the dynamics the magnetic field induces substructures at time scales much longer than the precession period.

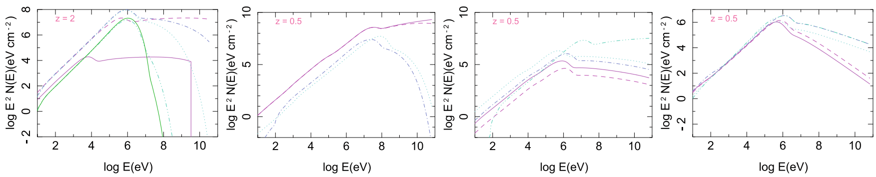


Fig. 5.3 From *left* to *right* (1) Electron distribution: power-law with exponential cutoff, $p = 2.5$ and $\omega_{cut}/\omega_m = 0.5$ (*full line*), 1 (*dash*-three *dots*), 10 (*dot*), 100 (*dash-dot*); $p = 1.9$ and $\omega_{cut}/\omega_m = 1000$ (*dash*). The external magnetic field in these simulations is 10 kG. The low amplitude *full line* has $p = 1.9$ and $\omega_{cut}/\omega_m = 1000$ but no external magnetic field. (2) Electron distribution: power-law with exponential cutoff for electrons, $|B_{ext}| = 70$ kGauss: $\omega_{cut}/\omega_m = 1000$, $p = 1.5$ (*full line*); $\omega_{cut}/\omega_m = 100$, $p = 1.5$ (*dash*); $\omega_{cut}/\omega_m = 3$, $p = 2$ (*dot-dash*); $\omega_{cut}/\omega_m = 3$, $p = 2.5$ (*dot*). (3) Electron distribution: power-law with exponential cutoff, $|B_{ext}| = 100$ kGauss: $\omega_{cut}/\omega_m = 1000$, $p = 2.5$, $\epsilon_e = 0.002$, $\Gamma = 500$ (*full line*); ω_{cut}/ω_m and p as previous case and $\epsilon_e = 0.02$, $\Gamma = 50$ (*dash*); $n'_0 = 5 \times 10^{15}$ cm^{-3} and other parameters as the previous case (*dot-dash*); varying p with index $-0.2,\ 0,\ 0.5$, initial $p = 2.5$ and $\epsilon_e = 0.002$ (*dot*); initial $p = 1.8$, the same indices as previous, and $\omega_{cut}/\omega_m = 0.5,\ 1000,\ 100$ (dash-3 dots). 4) Electron distribution: broken power-law a broken slope at $\omega_{cut}/\omega_m = 3$ and $p_1 = 2.5$, $p_2 = 4$, $|B_{ext}| = 17$ kGauss (*full line*); $p_1 = 2.1$, $p_2 = 4$ and same $|B_{ext}|$ as previous (*dash*); $p_1 = 2.1$, $p_2 = 3$, $|B_{ext}| = 26$ kGauss (*dot-dash*); same slope and $|B_{ext}| = 35$ kGauss (*dot*); same slope and $|B_{ext}| = 70$ kGauss (*dash*-three *dots*)

variation of physical quantities. Figure 5.3 shows some examples of spectra obtained for simulated bursts. They have a variety of behaviour at high energies. Notably, when the cutoff energy is high and the spectrum is flatter than what is possible for a simple power-law i.e. $p \leqslant 2$, the slope of the fluence at very high energies is positive, i.e. flux increases. Evidently, even in this case at very high energies the spectrum bends and the slope becomes negative. An example of such cases is the dash-dot spectrum in the first plot of Fig. 5.3.

5.4.1 *Compton Scattering, Delayed High Energy Component, and Other Issues*

We have also simulated the effect of the inverse Compton scattering of photons by high energy electrons, but due to the limited length of this proceeding we do not show them here. As we mentioned in Sect. 5.2 it is crucial to consider a realistic distribution of accelerated electrons along the photons path, otherwise we overestimate the scattering rate and obtain a spectrum that is not consistent with observations. Nonetheless, the light curves of IC scattered photons are more extended in time. Thus, apriori they should explain the detected delayed emission at high energies. However their flux is much smaller then what is observed.

On the other hand, the similarity of spectra shown in Fig. 5.3 which include only the synchrotron emission to observations is the evidence that the origin of delayed high energy emission is the same as lower energies. In [11] we have given detailed arguments that the reason for the delay of high energy emission is that the most energetic electrons are trapped in the EES and follow its propagation. In fact this is a self-organizing processes: energetic electrons can follow the propagation of EES. In this way they stay in the region where electric field is strong and do not get a lag that brings them to the region where magnetic field is strong. But, because they stay longer in the high electric field region, they are accelerated more and can better follow the propagation of EES. Simple calculations show that for nominal value of parameters the delay can be tens of seconds consistent with observations. We leave a quantitative study of this process for a future work. The issues of the slope at low energies and efficiency are also discussed in [11] and we show that they can be explained in the context of SSC model.

References

1. Zhao FY, Strom RG, Jiang SY (2006) The guest star of AD185 must have been a supernova", chin. J Astron Astrophys 6(5):635640
2. da Silva LAL (1993) The classification of supernovae. Astrophys Space Sci 202(2):215236
3. Klebesadel RW, Strong IB, Olson RA (1973) Observations of gamma-ray bursts of cosmic origin. ApJ 182L85.
4. Bonnell J (1995) http://apod.nasa.gov/htmltest/jbonnell/www/grbhist.html
5. Paczynski B (1986) Gamma-ray bursters at cosmological distances. ApJ Lett 308(2):L43–L46
6. Goodman J (1986) Are gamma-ray bursts optically thick? ApJ Lett 308(2):L47
7. Piran T (1999), Gamma-ray bursts and the fireball model. Phys Rep 314:575 [astro-ph/9810256].
8. Rees MJ, Mészáros P (1994) Unsteady outflow models for cosmological gamma-ray bursts. ApJ 430:L93 [astro-ph/9404038].
9. Ziaeepour H (2009a), Gamma ray bursts cook book I: formulation. MNRAS 397:361 [arXiv:0812.3277].
10. Ziaeepour H (2009b), Gamma ray bursts cook book II: simulation. MNRAS 397:386 [arXiv:0812.3279].
11. Ziaeepour H, Gardner B (2011) Broad band simulation of gamma ray bursts (GRB) prompt emission in presence of an external magnetic field. J Cosmol Astrop Phys 12:001

12. Fenimore EE, Ramirez-Ruiz E (1999) Gamma-ray bursts as internal shocks caused by deceleration. [astro-ph/9909299].
13. Umeda H, Tominaga N, Maeda K, Nomoto K (2005) Precursors and main bursts of gamma-ray bursts in a hypernova, scenario. ApJ 633:L17 astro-ph/0509750.
14. Medvedev M, Loeb A (1999) Generation of magnetic fields in the relativistic shock of gamma-ray-burst sources. ApJ 526:697 [astro-ph/9904363].
15. Lyutikov M, Blandford RD (2003) Gamma ray bursts as electromagnetic outflows. [astro-ph/0312347].
16. Sari R, Narayan R, Piran T (1996) Cooling time scales and temporal structure of gamma-ray bursts. ApJ 473:204 [astro-ph/9605005].
17. Sakamoto T et al. (2008) The first swift BAT gamma-ray burst catalog, ApJ Suppl 175:179 [arXiv:0707.4626].
18. Spitkovsky A (2008) Particle acceleration in relativistic collisionless shocks: fermi process at last? ApJ 682:5 [arXiv:0802.3216].
19. Murphy GC, Dieckmann ME, Drury L, O'C (2010) Multidimensional simulations of magnetic field amplification and electron acceleration to near-energy equipartition with ions by a mildly relativistic quasi-parallel plasma collision. IEEE Trans Plasma sci 38:2985 [arXiv:1011.4406].

Chapter 6
Effects of Modified Dispersion Relations and Noncommutative Geometry on the Cosmological Constant Computation

Remo Garattini

Abstract We compute Zero Point Energy in a spherically symmetric background with the help of the Wheeler-DeWitt equation. This last one is regarded as a Sturm-Liouville problem with the cosmological constant considered as the associated eigenvalue. The graviton contribution, at one loop is extracted with the help of a variational approach together with Gaussian trial functionals. The divergences handled with a zeta function regularization are compared with the results obtained using a Noncommutative Geometry (NCG) and Modified Dispersion Relations (MDR). In both NCG and MDR no renormalization scheme is necessary to remove infinities in contrast to what happens in conventional approaches. Effects on photon propagation are briefly discussed.

6.1 Introduction

The Cosmological Constant problem is certainly one of the most fascinating challenges of our days. A challenge because all the attempts that try to explain the 10^{120} orders of magnitude of discrepancy between the theory and observation have produced unsatisfying results. If we believe that Quantum Field Theory (QFT) is a part of the real world, then the theoretical predictive power to compute the Cosmological Constant must be entrusted to the methods of QFT at the Planck scale. Indeed, calculating the Zero Point Energy (ZPE) of some field of mass m with a cutoff at the Planck scale, we obtain

R. Garattini (✉)
Facoltà di Ingegneria, Università degli Studi di Bergamo, Viale Marconi 5, 24044 Dalmine (Bergamo), Italy
e-mail: remo.garattini@unibg.it

R. Garattini
I.N.F.N.–sezione di Milano, Milan, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_6,

$$E_{ZPE} = \frac{1}{2}\int_0^{\Lambda_p} \frac{d^3k}{(2\pi)^3}\sqrt{k^2+m^2} \simeq \frac{\Lambda_p^4}{16\pi^2} \approx 10^{71}GeV^4. \tag{6.1}$$

while the observation leads to a ZPE of the order 10^{-47}GeV^4. This surely represents one of the worst predictions of QFT. However if one insists to use the methods of QFT applied to General Relativity, one necessarily meets one of the most famous equations appeared in the literature of Cosmology and Gravity: the Wheeler-DeWitt (WDW) equation [1]. The WDWequation was originally introduced by Bryce DeWitt as an attempt to quantize General Relativity in a Hamiltonian formulation. It is described by

$$\mathscr{H}\Psi = \left[(2K)\, G_{ijkl}\pi^{ij}\pi^{kl} - \frac{\sqrt{g}}{2\kappa}\left({}^3R - 2\Lambda\right)\right]\Psi = 0 \tag{6.2}$$

and it represents the quantum version of the classical constraint which guarantees the invariance under time reparametrization. G_{ijkl} is the super-metric, π^{ij} is the super-momentum, ^{3}R is the scalar curvature in three dimensions and Λ is the cosmological constant, while $\kappa = 8\pi G$ with G the Newton's constant. An immediate application of the WDW Equation is given in terms of the Friedmann-Robertson-Walker (FRW) mini superspace, where all the degrees of freedom but the scale factor are frozen. The FRW metric is described by the following line element

$$ds^2 = -N^2dt^2 + a^2(t)\, d\Omega_3^2, \tag{6.3}$$

where $d\Omega_3^2$ is the usual line element on the three sphere, namely

$$d\Omega_3^2 = \gamma_{ij}dx^i dx^i dx^j. \tag{6.4}$$

In this background, we have simply

$$R_{ij} = \frac{2}{a^2(t)}\gamma_{ij} \quad \text{and} \quad R = \frac{6}{a^2(t)} \tag{6.5}$$

and the WDW equation $H\Psi(a) = 0$, becomes

$$\left[-a^{-q}\left[\frac{\partial}{\partial a}a^q\frac{\partial}{\partial a}\right] + \frac{9\pi^2}{4G^2}\left(a^2 - \frac{\Lambda}{3}a^4\right)\right]\Psi(a) = 0. \tag{6.6}$$

Eq. (6.6) assumes the familiar form of a one-dimensional Schrödinger equation for a particle moving in the potential

$$U(a) = \frac{9\pi^2}{4G^2}a^2\left(1 - \frac{a^2}{a_0^2}\right) \tag{6.7}$$

with total zero energy. The parameter q represents the factor-ordering ambiguity and $a_0 = \sqrt{\frac{3}{\Lambda}}$ is a reference length. The WDW Eq. (6.6) has been solved exactly in terms of Airy functions by Vilenkin [2] for the special case of operator ordering $q = -1$. The Cosmological Constant here appears as a parameter. Nevertheless, except the FRW case and other few examples, the WDW equation is very difficult to solve. This difficulty increases considerably when the mini superspace approach is avoided. However some information can be gained if one changes the point of view. Indeed, instead of treating Λ in Eq. (6.2) as a parameter, one can formally rewrite the WDW equation as an expectation value computation[1] [3, 4]. Indeed, if we multiply Eq. (6.2) by $\Psi^*\left[g_{ij}\right]$ and functionally integrate over the three spatial metric g_{ij} we find

$$\frac{1}{V}\frac{\int \mathcal{D}\left[g_{ij}\right]\Psi^*\left[g_{ij}\right]\int_\Sigma d^3x\hat{\Lambda}\Psi\left[g_{ij}\right]}{\int \mathcal{D}\left[g_{ij}\right]\Psi^*\left[g_{ij}\right]\Psi\left[g_{ij}\right]} = \frac{1}{V}\frac{\left\langle\Psi\left|\int_\Sigma d^3x\hat{\Lambda}_\Sigma\right|\Psi\right\rangle}{\langle\Psi|\Psi\rangle} = -\frac{\Lambda}{\kappa}. \quad (6.8)$$

In Eq. (6.8) we have also integrated over the hypersurface Σ and we have defined

$$V = \int_\Sigma d^3x\sqrt{g} \quad (6.9)$$

as the volume of the hypersurface Σ with

$$\hat{\Lambda}_\Sigma = (2\kappa)\, G_{ijkl}\pi^{ij}\pi^{kl} - \sqrt{g}^3R\Big/(2\kappa)\,. \quad (6.10)$$

In this form, Eq. (6.8) can be used to compute ZPE provided that Λ/κ be considered as an eigenvalue of $\hat{\Lambda}\Sigma$. In particular, Eq. (6.8) represents the Sturm-Liouville problem associated with the cosmological constant. To solve Eq. (6.8) is a quite impossible task. Therefore, we are oriented to use a variational approach with trial wave functionals. The related boundary conditions are dictated by the choice of the trial wave functionals which, in our case are of the Gaussian type. Different types of wave functionals correspond to different boundary conditions. The choice of a Gaussian wave functional is justified by the fact that ZPE should be described by a good candidate of the "*vacuum state*". In the next section we give the general guidelines in ordinary gravity and in presence of Modified Dispersion Relations and the Non Commutative approach to QFT. Units in which $\hbar = c = k = 1$ are used throughout the paper.

[1] See also Ref. [5] for an application of the method to a $f(R)$ theory.

6.2 High Energy Gravity Modification: the Example of Non Commutative Theories and Gravity's Rainbow

As an application of the Eq. (6.8), we consider the simple example of the Mini Superspace described by a FRW cosmology (6.3). We find the following simple expectation value

$$\frac{\int \mathscr{D}a\Psi^*(a)\left[-\frac{\partial^2}{\partial a^2}+\frac{9\pi^2}{4G^2}a^2\right]\Psi(a)}{\int \mathscr{D}a\Psi^*(a)\left[a^4\right]\Psi(a)}=\frac{3\Lambda\pi^2}{4G^2} \tag{6.11}$$

where the normalization is modified by a weight factor. The application of a variational procedure with a trial wave functional of the form

$$\Psi=\exp\left(-\beta a^2\right) \tag{6.12}$$

shows that there is no real solution of the parameter β compatible with the procedure. Nevertheless, a couple of imaginary solutions of the variational parameter

$$\beta=\pm i\frac{3\pi}{4G} \tag{6.13}$$

can be found. These solutions could be related to the tunneling wave function of Vilenkin [2]. Even if the eigenvalue procedure of Eq. (6.11) leads to an imaginary cosmological constant, which does not correspond to the measured observable, it does not mean that the procedure is useless. Indeed, Eq. (6.8) can be used to calculate the cosmological constant induced by quantum fluctuations of the gravitational field. To fix ideas, we choose the following form of the metric

$$ds^2=-N^2(r)\,dt^2+\frac{dr^2}{1-\frac{b(r)}{r}}+r^2\left(d\theta^2+\sin^2\theta d\phi^2\right). \tag{6.14}$$

where $b(r)$ is subject to the only condition $b(r_t)=r_t$. We consider $g_{ij}=\bar{g}_{ij}+\hbar_{ij}$; where $\bar{g}_{ij}$ is the background metric and $\hbar_{ij}$ is a quantum fluctuation around the background. Then we expand Eq. (6.8) in terms of h_{ij}. Since the kinetic part of $\hat{\Lambda}_\Sigma$ is quadratic in the momenta, we only need to expand the three-scalar curvature $\int d^3x\sqrt{g}{}^3R$ up to second order in $\hbar_{ij}$ [2] As shown in Ref. [7], the final contribution does not include ghosts and simply becomes

$$\frac{1}{V}\frac{\left\langle\Psi^\perp\left|\int_\Sigma d^3x\left[\hat{\Lambda}_\Sigma^\perp\right]^{(2)}\right|\Psi^\perp\right\rangle}{\left\langle\Psi^\perp|\Psi^\perp\right\rangle}=-\frac{\Lambda^\perp}{K} \tag{6.15}$$

[2] See Refs. [5–7] for technical details.

The integration over Gaussian Wave functionals leads to

$$\hat{\Lambda}_{\Sigma}^{\perp} = \frac{1}{4V}\int_{\Sigma} d^3x\sqrt{g}G^{ijkl}\left[(2\kappa)\,K^{-1\perp}(x,x)_{ijkl} + \frac{1}{(2\kappa)}\left(\tilde{\Delta}_L\right)^a_j K^{\perp}(x,x)_{iakl}\right], \tag{6.16}$$

where

$$(\Delta_L h)_{ij} = \left(\Delta_L h^{\perp}\right)_{ij} - 4R_i^k h_{kj}^{\perp} +^3 Rh_{ij}^{\perp} \tag{6.17}$$

is the modified Lichnerowicz operator and Δ_{L} is the Lichnerowicz operator defined by

$$\left(\tilde{\Delta}_L h\right)_{ij} = \Delta h_{ij} - 2R_{ikjl}h_j^k + R_{jk}h_i^k \quad \Delta = -\nabla^a\nabla_a. \tag{6.18}$$

G^{ijkl} represents the inverse DeWitt metric and all indices run from one to three. Note that the term $-4R_i^k\hbar_{kj}^{\perp} + {}^3R\hbar_{ij}^{\perp}$ disappears in four dimensions. The ropagator $K^{\perp}(x,x)_{iakl}$ can be represented as

$$K^{\perp}\left(\overrightarrow{x},\overrightarrow{y}\right)_{iakl} = \sum_{\tau}\frac{h_{ia}^{(\tau)\perp}\left(\overrightarrow{x}\right)h_{kl}^{(\tau)\perp}\left(\overrightarrow{y}\right)}{2\lambda(\tau)}, \tag{6.19}$$

where $\hbar_{ia}^{(\tau)\perp}(\vec{x})$ are the eigenfunctions of $\hat{\Delta}_L.\tau$ denotes a complete set of indices $\lambda(t)$ and are a set of variational parameters to be determined by the minimization of Eq. (6.16). The expectation value of $\widehat{\Delta}_{\Sigma}^{\perp}$ is easily obtained by inserting the form of the propagator into Eq. (6.16) and minimizing with respect to the variational function $\lambda(t)$. Thus the total one loop energy density for TT tensors becomes

$$\frac{\Lambda}{8\pi G} = -\frac{1}{2}\sum_{\tau}\left[\sqrt{\omega_1^2(\tau)} + \sqrt{\omega_2^2(\tau)}\right]. \tag{6.20}$$

The above expression makes sense only for $\omega_i^2(\tau) > 0$, where ω_i are the eigenvalues of $\tilde{\Delta}_L$. Following Refs. [5–7], we find that the final evaluation of expression (6.20) is

$$\frac{\Lambda}{8\pi G} = -\frac{1}{\pi}\sum_{i=1}^{2}\int_0^{+\infty}\omega_i\frac{d\tilde{g}(\omega_i)}{d\omega_i}d\omega_i = -\frac{1}{4\pi^2}\sum_{i=1}^{2}\int_{\sqrt{m_i^2(r)}}^{+\infty}\omega_i^2\sqrt{\omega_i^2 - m_i^2(r)}d\omega_i, \tag{6.21}$$

where we have included an additional 4π coming from the angular integration and where we have defined two r-dependent effective $m_1^2(r)$ masses $m_2^2(r)$ and

$$\begin{cases} m_1^2(r) = \frac{6}{r^2}\left(1 - \frac{b(r)}{r}\right) + \frac{3}{2r^2}b'(r) - \frac{3}{2r^3}b(r) \\ m_2^2(r) = \frac{6}{r^2}\left(1 - \frac{b(r)}{r}\right) + \frac{1}{2r^2}b'(r) - \frac{3}{2r^3}b(r) \end{cases} \quad (r \equiv r(x)). \tag{6.22}$$

The effective masses have different expression from case to case. For example, in the Schwarzschild case, b(r) = 2MG, we find

$$\begin{cases} m_1^2(r) = \frac{6}{r^2}\left(1 - \frac{2MG}{r}\right) - \frac{3MG}{r^3} \\ m_2^2(r) = \frac{6}{r^2}\left(1 - \frac{2MG}{r}\right) + \frac{3MG}{r^3} \end{cases} \qquad (r \equiv r(x)). \tag{6.23}$$

The expression in Eq. (6.21) is divergent and must be regularized. For example, the zeta function regularization method leads to

$$\rho_i(\varepsilon) = \frac{m_i^4(r)}{64\pi^2}\left[\frac{1}{\varepsilon} + \ln\left(\frac{4\mu^2}{m_i^2(r)\sqrt{e}}\right)\right] \qquad i = 1, 2 \ldots, \tag{6.24}$$

where an additional mass parameter μ has been introduced in order to restore the correct dimension for the regularized quantities. Such an arbitrary mass scale emerges unavoidably in any regularization scheme. The renormalization is performed via the absorption of the divergent part into the re-definition of a bare classical quantity and the final result is given by

$$\frac{\Lambda_0}{8\pi G} = \frac{m_1^4(r)}{64\pi^2}\ln\left(\frac{4\mu^2}{m_1^2(r)\sqrt{2}}\right) + \frac{m_2^4(r)}{64\pi^2}\ln\left(\frac{4\mu^2}{m_2^2(r)\sqrt{e}}\right) \tag{6.25}$$

Of course, one can follow other methods to obtain finite results: for instance, the use of a UV-cut off. Nevertheless it is possible to obtain finite results introducing a distortion in the space-time from the beginning. This can be realized with the help of the Non Commutative Approach to QFT developed in Ref. [8] or with the help of Gravity's Rainbow developed in Ref. [10]. Noncommutative theories provide a powerful method to naturally regularize divergent integrals appearing in QFT. Eq. (6.21) is a typical example of a divergent integral. The noncommutativity of spacetime is encoded in the commutator $[x^\mu, x^v] = i\theta^{\mu v}$, where $\theta^{\mu v}$ is an antisymmetric matrix which determines the fundamental discretization of spacetime. In even dimensional space-time, $\theta^{\mu v}$ can be brought to a block-diagonal form by a suitable Lorentz rotation leading to

$$\theta^{\mu v} = diag\left[\theta_1 \varepsilon^{ab} \theta_2 \varepsilon^{ab} \ldots \theta_{d/2} \varepsilon^{ab}\right] \tag{6.26}$$

with ε^{ab} a 2×2 antisymmetric Ricci Levi-Civita tensor. If $\theta_i \equiv \theta \forall i = 1 \ldots d/2$, then the space-time is homogeneous and preserves isotropy. The effect of the θ length on ZPE calculation is basically the following: the classical Liouville counting number of nodes

$$dn = \frac{d^3\vec{x}d^3\vec{k}}{(2\pi)^3}, \tag{6.27}$$

is modified by distorting the counting of nodes in the following way [8]

$$dn = \frac{d^3xd^3k}{(2\pi)^3} \Rightarrow dn_i = \frac{d^3xd^3k}{(2\pi)^3} \exp\left(-\frac{\theta}{4}\left(\omega_{i,nl}^2 - m_i^2(r)\right)\right), \quad i = 1, 2, \tag{6.28}$$

This deformation corresponds to an effective cut off on the background geometry (6.14). The UV cut off is triggered only by higher momenta modes $\gtrsim 1/\sqrt{\theta}$ which propagate over the background geometry. As an effect the final induced cosmological constant becomes

$$\begin{aligned}\tfrac{\Lambda}{8\pi G} = \tfrac{1}{6\pi^2}\Bigg[\int_{\sqrt{m_1^2(r)}}^{+\infty}\sqrt{\left(\omega^2 - m_1^2(r)\right)^3}e^{-\frac{\theta}{4}\left(\omega^2 - m_1^2(r)\right)}d\omega \\ + \int_{\sqrt{m_2^2(r)}}^{+\infty}\sqrt{\left(\omega^2 - m_2^2(r)\right)^3}e^{-\frac{\theta}{4}\left(\omega^2 - m_2^2(r)\right)}d\omega\Bigg],\end{aligned} \tag{6.29}$$

where an integration by parts in Eq. (6.21) has been done. We recover the usual *divergent* integral when $\theta \to 0$. The result is finite and we have an induced cosmological constant which is regular. We can obtain enough information in the asymptotic régimes when the background satisfies the relation

$$m_0^2(r) = m_1^2(r) = -m_2^2(r), \tag{6.30}$$

which is valid for the Schwarzschild, Schwarzschild-de Sitter (SdS) and Schwarzschild-Anti de Sitter (SAdS) metric close to the throat. Indeed, defining

$$x = \frac{m_0^2(r)\,\theta}{4}, \tag{6.31}$$

we find that when $x \to +\infty$,

$$\frac{\Lambda}{8\pi G} \simeq \frac{1}{6\pi^2\theta^2}\sqrt{\frac{\pi}{x}}\left[3 + \left(8x^2 + 6x + 3\right)\exp(-x)\right] \to 0. \tag{6.32}$$

Conversely, when $x \to 0$, we obtain

$$\frac{\Lambda}{8\pi G} \simeq \frac{4}{3\pi^2\theta^2}\left[2 - \left(\frac{7}{8} + \frac{3}{4}\ln\left(\frac{x}{4}\right) + \frac{3}{4}\gamma\right)x^2\right] \to \frac{8}{3\pi^2\theta^2}. \tag{6.33}$$

The other interesting cases, namely de Sitter and Anti-de Sitter and Minkowski are described by

$$m_1^2(r) = m_2^2(r) = m_0^2(r), \tag{6.34}$$

leading to

$$\frac{\Lambda}{8\pi G} \simeq \frac{1}{6\pi^2}\left(\frac{4}{\theta}\right)^2\frac{3}{8}\sqrt{\frac{\pi}{x}} \to 0, \tag{6.35}$$

when $x \to \infty$ and

$$\frac{\Lambda}{8\pi G} \simeq \frac{1}{6\pi^2}\left(\frac{4}{\theta}\right)^2\left[1-\frac{x}{2}+\left(-\frac{7}{16}-\frac{3}{8}\ln\left(\frac{x}{4}\right)-\frac{3}{8}\gamma\right)x^2\right] \to \frac{8}{3\pi^2\theta^2}, \tag{6.36}$$

when $x \to 0$. As regards Gravity's Rainbow [9], we can begin by defining a *"rainbow metric"*

$$ds^2 = -\frac{N^2(r)\,dt^2}{g_1^2(E/E_p)} + \frac{dr^2}{\left(1-\frac{d(r)}{r}\right)g_2^2(E/E_p)} + \frac{r^2}{g_2^2(E/E_p)}\left(d\theta + \sin^2\theta d\phi^2\right). \tag{6.37}$$

$g_1(E/E_P)$ and $g_2(E/E_P)$ are two arbitrary functions which have the following property

$$\lim_{E/Ep \to 0} g_1(E/Ep) = 1 \quad and \quad \lim_{E/Ep \to 0} g_2(E/Ep) = 1. \tag{6.38}$$

We expect the functions $g_1(E/E_P)$ and $g2(E/E_P)$ modify the UV behavior in the same way as GUP and Noncommutative geometry do, respectively. Following Ref. [10], in presence of Gravity's Rainbow, we find that Eq. (6.8) changes into

$$\frac{g_2^3(E/Ep)}{\tilde{V}}\frac{\left\langle \Psi \left| \int_\Sigma d^3x \tilde{\Lambda}_\Sigma \right| \Psi \right\rangle}{\langle \Psi | \Psi \rangle} = -\frac{\Lambda}{k}, \tag{6.39}$$

where

$$\tilde{\Lambda}_\Sigma = (2k)\frac{g_1^2(E/Ep)}{g_2^3(E/Ep)}\tilde{G}_{ijkl}\tilde{\pi}^{ij}\tilde{\pi}^{kl} - \frac{\sqrt{\tilde{g}}\tilde{R}}{(2k)\,g_2(E/Ep)}. \tag{6.40}$$

Of course, Eq. (6.39) and Eq. (6.40) reduce to the ordinary Eqs. (6.2; 6.8) and (6.10) when $E/E_P \to 0$ By repeating the procedure leading to Eq. (6.20), we find that the TT tensor contribution of Eq. (6.39) to the total one loop energy density becomes

$$\frac{\Lambda}{8\pi G} = -\frac{1}{3\pi^2}\sum_{i=1}^{2}\int_{E^*}^{+\infty} E_{igl}(E/Ep)\,g_2(E/Ep)\frac{d}{dE_i}\sqrt{\left(\frac{E_i^2}{g_2^2(E/Ep)} - m_i^2(r)\right)^3}\,dE_i. \tag{6.41}$$

where E^* is the value which annihilates the argument of the root. In the previous equation we have assumed that the effective mass does not depend on the energy E. To further proceed, we choose a form of $g_1(E/E_P)$ and $g_2(E/E_P)$ which allows a comparison with the results obtained with a Noncommutative geometry computation expressed by Eq. (6.29). We are thus led to choose

$$g_1(E/Ep) = \left(1+\beta\frac{E}{Ep}\right)\exp\left(-\alpha\frac{E^2}{E_p^2}\right) \quad \text{and} \quad g_2(E/Ep) = 1. \tag{6.42}$$

with $\alpha > 0$ and $\beta \in \mathbb{R}$, because the pure "*Gaussian*" choice with $\beta = 0$ can not give a positive induced cosmological constant.[3] However this is true when the effective masses satisfy relation (6.30). In case relation (6.34) holds the pure "*Gaussian*" choice works for large and small x, where $x = \sqrt{m_0^2(r)/E_P^2}$. The final result is a vanishing induced cosmological constant in both asymptotic régimes. It is interesting to note that Gravity's Rainbow has potential effects on the photon propagation [11]. Indeed, let us consider two photons emitted at the same time $t = -t_0$ at $x_{dS} = 0$: The first photon be a low energy photon ($E \ll E_P$) and the second one be a Planckian photon ($E \sim E_P$): Both photons are assumed to be detected at a later time in $\bar{x}_{ds}$. We expect to detect the two photons with a time delay Dt given by the solution of the equation

$$\bar{x}_{dS}^{E \ll Ep}(0) = \bar{x}_{dS}^{E \sim Ep}(\Delta t), \tag{6.43}$$

that implies

$$\Delta t \simeq g_1(E) \frac{e\sqrt{\Lambda_{eff}/3t0} - e^{\frac{\sqrt{\Lambda_{eff}/3t0}}{g_1(E)}}}{\sqrt{\Lambda_{eff}/3}} \simeq \beta \frac{E}{Ep} t0 \left(1 + \sqrt{\Lambda_{eff}/3t0}\right), \tag{6.44}$$

where we have used (6.42) for the rainbow functions.

References

1. DeWitt BS (1967) Phys Rev 160:1113
2. Vilenkin A (1988) Phys Rev D 37:888
3. R.Garattini (2006) J Phys A 39, 6393 ; gr-qc/0510061
4. Garattini R (2006) Phys J Conf Ser 33, 215 ; gr-qc/0510062
5. Capozziello S, Garattini R (2007) Class Quant Grav 24, 1627; gr-qc/0702075
6. Garattini R (2004) TSPU Vestnik 44 N7, 72 ; gr-qc/0409016
7. Garattini R (2011) The cosmological constant and the Wheeler-DeWitt Equation, PoS CLAQG08 012; arXiv:0910.1735 [gr-qc]
8. Garattini R, Nicolini P (2011) Phys Rev D 83, 064021 (2011); arXiv:1006.5418 [gr-qc]
9. Magueijo J, Smolin L (2004) Class Quant Grav 21, 1725 [arXiv:gr-qc/0305055]
10. Garattini R, Mandanici G (2011) Phys Rev D 83, 084021, arXiv:1102.3803 [gr-qc]
11. Garattini R, Mandanici G, Particle propagation and effective space-time in Gravity's Rainbow. To appear in Phys. Rev.D; e-Print: arXiv:1109.6563 [gr-qc]

[3] See Ref. [10] for technical details.

Chapter 7
Testing the Nature of Astrophysical Black Hole Candidates

Cosimo Bambi

Abstract It is thought that the final product of the gravitational collapse is a Kerr black hole and astronomers have discovered several good astrophysical candidates. For the time being, all the black hole candidates are objects so compact and heavy that cannot be explained otherwise without introducing new physics, but there is no evidence that the geometry of the space-time around them is really described by the Kerr metric. Future space-based gravitational wave detectors will be able to test the nature of the black hole candidates with high precision. In this talk, I will show that the measurement of the radiative efficiency of individual AGN may test the Kerr black hole hypothesis at a similar level of accuracy, but before the advent of gravitational wave astronomy.

7.1 Introduction

General Relativity (GR) has been tested and verified to high precision for distances in the rage ∼1 mm to ∼1 pc and for weak gravitational fields [36]. The research is now moving to check the validity of the theory at cosmological scales, sub-millimeter distances, and for strong gravitational fields. One of the most intriguing predictions of GR is that the collapsing matter inevitably produces singularities in the space-time. According to the weak cosmic censorship conjecture, singularities of gravitational collapse must be hidden within black holes (BHs) [29]. In 4-dimensional GR, BHs are described by the Kerr solution, which is completely specified by two parameters, the mass, M, and the spin angular momentum, J [16, 30]. The condition for the existence of the event horizon is that the spin parameter $a = |J/M^2|$ cannot exceed 1. When $a > 1$, there is no horizon and the central singularity is naked, violating the weak cosmic censorship conjecture.

C. Bambi (✉)
Ludwig-Maximilians-Universität München, Munich, Germany
e-mail: Cosimo.Bambi@physik.uni-muenchen.de

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_7,

Astronomers have discovered at least two classes of BH candidates (for a review, see e.g. [25]): stellar-mass objects in X-ray binary systems ($M \sim 5$–20 Solar masses) and super-massive objects in galactic nuclei ($M \sim 10^5$–10^9 Solar masses). The estimates of the masses of these objects are robust, because obtained via dynamical measurements and without any assumption about the nature of the massive body. The stellar-mass objects in X-ray binary systems are too heavy to be neutron or quark stars for any reasonable matter equation of state, while the super-massive objects at the centers of galaxies are too heavy, compact, and old to be clusters of non-luminous bodies. All these objects are therefore thought to be the BHs predicted by GR, as they cannot be explained otherwise without introducing new physics. However, there is no indication that the geometry around these objects is described by the Kerr metric.

Testing the Kerr BH hypothesis is thus the next step to progress in this research field and several authors have indeed suggested possible ways to do it using present and future data (for a review, see e.g. [3]). A very promising approach is the detection of extreme mass ratio inspirals (EMRIs, i.e. systems consisting of a stellar-mass compact object orbiting a super-massive BH candidate) with future space-based gravitational wave antennas. Missions like LISA will be able to follow the stellar-mass compact object for millions of orbits around the central super-massive BH candidate, and therefore deviations from the Kerr geometry will lead to a phase difference in the gravitational waveforms that grows with the number of observed cycles ([2, 15, 19, 32] etc.). However, these data will not be available shortly, as the first mission will be launched, at best, in the early 2020s. This fact has motivated the study of alternative approaches to test the nature of BH candidates, such as the X-ray continuum fitting method [3], observations of quasi-periodic oscillations [20] and measurements of the cosmic X-ray background [3, 4]. These methods can in principle be applied even with present data, provided that the systematic errors are properly understood. Future observations of the shadow of nearby super-massive BH candidates are another exciting possibility to test the Kerr BH paradigm [9, 10, 14].

Previous studies have clearly pointed out that "rapidly-rotating" objects are the best candidates to test the Kerr BH hypothesis: if the object rotates fast, even a small deviation from the Kerr background can cause significant differences in the properties of the electromagnetic radiation emitted by the gas of the accretion disk and peculiar features, otherwise absent in the Kerr geometry, may show up [12–14]. In this talk, I will consider the most massive BH candidates in AGN. I will show that the measurement of their radiative efficiency may soon provide stringent constraints on possible deviations from the Kerr geometry, comparable to the ones that may be obtained with future space-based gravitational wave detectors. The talk is based on a work in preparation [8].

7.2 Standard Accretion Disk Model

The Novikov-Thorne (NT) model is the standard model for accretion disks [27]. It describes geometrically thin and optically thick disks and it is the relativistic generalization of the Shakura-Sunyaev model [32]. The disk is thin in the sense that the disk opening angle is $h = H/r << 1$, where H is the thickness of the disk at the radius r. Magnetic fields are ignored. In the Kerr background, there are four parameters (BH mass, BH spin parameter, mass accretion rate, and viscosity parameter), but the model can be easily extended to any (quasi-) stationary, axisymmetric, and asymptotically flat space-time. Accretion is possible because viscous magnetic/turbulent stresses and radiation transport energy and angular momentum outwards. The model assumes that the disk is on the equatorial plane and that the disk's gas moves on nearly geodesic circular orbits. Heat advection is ignored (it scales as h^2) and energy is radiated from the disk surface.

The key-ingredient of the NT model is that the inner edge of the disk is at the innermost stable circular orbit (ISCO), where viscous stresses are assumed to vanish. When the gas's particles reach the ISCO, they quickly plunge into the BH, without emitting additional radiation. Neglecting the radiation emitted by the disk and captured by the BH, the maximum value for the radiative efficiency is

$$\eta = 1 - E_{ISCO}, \tag{7.1}$$

where E_{ISCO} is the specific energy of the gas at the ISCO radius and depends uniquely on the background geometry. Eq. 7.1 provides the maximum value for η because a fraction of the gas's gravitational energy may be converted to kinetic energy of jet/wind outflows. In what follows, I will assume the conservative hypothesis that all the gravitational energy of the gas is converted to radiation and η is given by Eq. 7.1.

As a consequence of the accretion process, the BH spin parameter evolves. Since the gas particles arriving at the ISCO plunge quickly into the central object, without emission of additional radiation, the BH changes its mass by $\delta M = E_{ISCO}\delta m$ and its spin angular momentum by $\delta J = L_{ISCO}\delta m$, where L_{ISCO} is the specific angular momentum of the gas at the ISCO, while δm is the gas rest-mass. The evolution of the spin parameter of the BH, a, turns out to be governed by the following equation:

$$da/dlnM = (1/M)L_{ISCO}/E_{ISCO} - 2a. \tag{7.2}$$

If the right hand side of Eq. 7.2 is positive, the accretion process spins the BH up. If it is negative, the BH is spun down. The equilibrium spin parameter a_{eq} is reached when the right hand side of Eq. 7.2 vanishes and its value depends on the geometry of the space-time.

For non-magnetized and weakly-magnetized disks, there is a common consensus that the NT model describes correctly thin disks, $h << 1$, when the viscosity parameter is small [1]. In the case of magnetized disks, the issue is more controversial, as it is not yet possible to perform GRMHD simulations of thin disks (see e.g. [28] and [26]). Here, I will assume that the NT model works, as it is commonly supposed in most studies discussed in the literature. At the observational level, a common criterion to select sources with thin disks is that the bolometric luminosity of the source does not exceed 30 % of its Eddington luminosity [24].

7.3 Radiative Efficiency in Kerr and Non-Kerr Backgrounds

In the NT model, the maximum value of the radiative efficiency can be immediately inferred from Eq. 7.1 and depends only on the background geometry. In the Kerr space-time, there is a one-to-one correspondence between a and η. For corotating disks, the radiative efficiency increases monotonically with the spin parameter, from about 0.057 (Schwarzschild BH, $a = 0$) to about 0.42 (extreme Kerr BH, $a = 1$). If a compact object is a Kerr BH, a measurement of its radiative efficiency can potentially be used to estimate its spin parameter.

To test the Kerr BH hypothesis, we have to consider a more general background that includes the Kerr solution as special case. Roughly speaking, the object will be specified by its mass, its spin angular momentum, and one or more "deformation parameters", measuring deviations from the Kerr geometry. The Kerr metric will be recovered when all the deformation parameters vanish. The idea is to study the properties of the accretion process in this more general space-time and then compare the theoretical predictions with the observational data. If the latter demand that the deformation parameters must vanish, then the Kerr BH hypothesis is verified and our astrophysical BH candidates are really the BHs predicted by GR.

In Fig. 7.1, I show the radiative efficiency η for a subclass of Manko-Novikov space-times [21–23] characterized by one deformation parameter, the anomalous quadrupole moment q (for more details, see [8]). q is related to the mass-quadrupole moment of the compact object by the relation:

$$Q = (1 + q)Q_{KERR}, \tag{7.3}$$

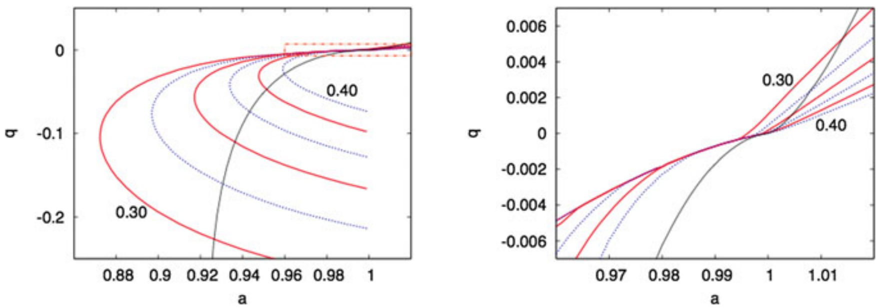


Fig. 7.1 Manko-Novikov space-time with deformation parameter q. Contour plots of the radiative efficiency $\eta = 1 - E_{ISCO}$: $\eta = 0.30$ (*red solid curve*), 0.32 (*blue dotted curve*), 0.34 (*red solid curve*), 0.36 (*blue dotted curve*), 0.38 (*red solid curve*), 0.40 (*blue dotted curve*). The *black solid curve* is the equilibrium spin parameter a_{eq} as inferred from Eq. 7.2. If $a < a_{eq}$, the accretion process spins the compact objects up; if $a > a_{eq}$, the accretion process spins the compact object down. The *right* panel is simply the enlargement of the area inside the orange box in the *left* panel. From [8]

where $Q_{KERR} = -a^2 M^3$ is the mass-quadrupole moment of a Kerr BH with mass M and spin parameter a. The compact object is thus more oblate than a Kerr BH with the same spin when $q > 0$, and it is more prolate when $q < 0$. Because of the introduction of a new parameter, there is now a degeneracy in η and, in general, a certain value of the radiative efficiency cannot be associated with a unique value of the spin parameter and of the anomaous quadrupole moment.

7.4 Constraining Deviations from the Kerr Geometry

At the observational level, the radiative efficiency is defined as the ratio between the bolometric luminosity of the source and the mass accretion rate of the BH candidate. The latter, however, is usually difficult to measure. The Soltan's argument provides an elegant way to determine the mean radiative efficiency of AGN from the mean BH mass density in the contemporary Universe and the AGN luminosity per unit volume integrated over time [33]. There are several sources of uncertainty in the final result, but current estimates suggest a high radiative efficiency (see e.g. [18, 35]).

Recently, Davis and Laor [17] proposed a way to estimate the radiative efficiency of individual AGN. The mass accretion rate can indeed be determined from the low frequency region of the thermal spectrum of the accretion disk of these objects, if the mass of the BH candidate is known. Such a radiation is mainly emitted at large radii, where gravity is Newtonian and the details of all the complicated astrophysical processes, like the viscosity mechanism, are not important. The authors found a strong correlation of η with M, raising from $\eta \sim 0.03$, when $M \sim 10^7$ Solar masses and the bolometric to Eddington luminosity ratio is $L_{bol}/L_{Edd} \sim 1$, to $\eta \sim 0.4$, when $M \sim 10^9$ Solar masses and $L_{bol}/L_{Edd} \sim 0.3$. For our discussion, the crucial point is that the most massive BH candidates in AGN seem to have a high radiative efficiency, $\eta \sim 0.4$, and a moderate mass accretion rate, $L_{bol}/L_{Edd} \sim 0.3$. The high radiative efficiency suggests that they are very rapidly-rotating objects. The moderate luminosity could be interpreted as the indication that their accretion disk is geometrically thin: if this is the case, the measurement of radiative efficiency could provide the specific energy of the gas at the ISCO via Eq. 7.1.

Interestingly, in the case of super-massive BH candidates, we can deduce an upper bound on their spin parameter. These objects must have $a < a_{eq}$, where a_{eq} the equilibrium spin parameter in case of accretion from a thin disk. The point is that the super-massive BH candidates have increased their mass by several orders of magnitude from its original value, the value of the spin parameter at the time of the formation of the object is irrelevant, while other processes (chaotic accretion, minor and major mergers) more likely spin the BH down [5–8]. On the other hand, the process of accretion from a thin disk is the most efficient mechanism to spin these objects up (and it spins them down if $a > a_{eq}$).

If observations can provide a robust lower bound on the radiative efficiency of a source, and we then combine such a measurement with the requirement $a < a_{eq}$, we can obtain interesting constraints on possible deviations from the Kerr background.

For instance, for $\eta > 0.30$, we can deduce the bound $-0.20 < q < 0.005$, see Fig. 7.1. The observation of gravitational waves emitted by an EMRI with future space-based gravitational wave detectors like LISA will be able to constrain the quadrupole moment of the super-massive BH candidates with a precision of order 0.01–0.001 [15]. For $q > 0$, the measurement of the radiative efficiency can likely provide similar results of LISA. Let us also notice that a self-gravitating fluid with reasonable equations of state has q positive and significantly larger than 0.01. For a neutron star, $q > 1$. The constraint in the region $q < 0$ is much weaker with our approach. However, these objects might be excluded from theoretical arguments: their ISCO is marginally unstable along the vertical direction, which means that there are two "centers of attraction", one above and one below the equatorial plane. It is not clear if a similar object can exist and be stable, as it may be necessary a repulsive force between the two centers of attraction in order to balance their gravitational force and maintain them at a fixed distance. On the other hand, if we discovered a BH candidate with $\eta > 0.32$ accreting from a thin disk, the Kerr BH hypothesis may be rejected and the existence of objects with $q < 0$ may be necessary to explain the observation. Astrophysical Kerr BHs can unlikely have a radiative efficiency higher than 0.32 [34].

7.5 Conclusions

There is some evidence that the most massive BH candidates in AGN have a high radiative efficiency and a thin accretion disk. In this case, they could be excellent candidates to test GR in the strong field regime and, in particular, the Kerr BH paradigm. For instance, the confirmation of the existence of BH candidates with $\eta > 0.30$ could test the Kerr BH hypothesis at the level of 0.5 % for objects more oblate than a Kerr BH, and at the level of 20 % for more prolate bodies. These bounds can be compared with the capabilities of future space-based gravitational wave detectors like LISA, which may be able to perform the same test with an accuracy of 0.1–1 %. For the time being, there are a few issues to address before using the measurement of the radiative efficiency of individual AGN to test GR (the validity of the NT model for magnetized disk and the confirmation of the method proposed in [17]), but the approach seems to be promising and capable of providing interesting constraints on the nature of the super-massive BH candidates well before the advent of gravitational wave astronomy.

Acknowledgments This work was supported by Humboldt Foundation.

References

1. Afshordi N, Paczynski B (2003) Astrophys J 592:354–367
2. Apostolatos TA, Lukes-Gerakopoulos G, Contopoulos G (2009) Phys Rev Lett 103:111101
3. Bambi C (2011a) Mod Phys Lett A 26:2453–2468
4. Bambi C (2011b) Phys Rev D 83:103003
5. Bambi C (2011c) Phys Lett B 705:5–8
6. Bambi C (2011d) Europhys Lett 94:50002
7. Bambi C (2011e) JCAP 1105:009
8. Bambi C (2011f) in preparation.
9. Bambi C, Freese K (2009) Phys Rev D 79:043002
10. Bambi C, Yoshida N (2010) Class Quantum Gravity 27:205006
11. Bambi C, Barausse E (2011a) Astrophys J 731:121
12. Bambi C, Barausse E (2011b) Phys Rev D 84:084034
13. Bambi C, Modesto L (2011c) Phys Lett B 706:13–18
14. Bambi C, Caravelli F, Modesto L (2011d) arXiv:1110.2768.
15. Barack L, Cutler C (2007) Phys Rev D 75:042003
16. Carter B (1971) Phys Rev Lett 26:331–333
17. Davis SW, Laor A (2011) Astrophys J 728:98
18. Elvis M, Risaliti G, Zamorani G (2002) Astrophys J Lett 565:L75–L77
19. Glampedakis K, Babak S (2006) Class Quantum Gravity 23:4167–4188
20. Johannsen T, Psaltis D (2011) Astrophys J 726:11
21. Manko VS, Mielke EW, Sanabria-Gomez JD (2000a) Phys Rev D 61:081501
22. Manko VS, Sanabria-Gomez JD, Manko OV (2000b) Phys Rev D 62:044048
23. Manko VS, Novikov ID (1992) Class Quantum Gravity 9:2477–2487
24. McClintock JE et al (2006) Astrophys J 652:518–539
25. Narayan R (2005) New J Phys 7:199
26. Noble SC, Krolik JH, Hawley JF (2010) Astrophys J 711:959–973
27. Novikov ID, Thorne KS (1973) Black Holes. Les Astres Occlus, In, pp 343–450.
28. Penna RF et al (2010) Mon Not R Astron Soc 408:752–782
29. Penrose R (1969) Rivista del Nuovo Cimento. Numero Speciale 1:252
30. Robinson DC (1975) Phys Rev Lett 34:905–906
31. Ryan FD (1995) Phys Rev D 52:5707–5718
32. Shakura NI, Sunyaev RA (1973) Astron Astrophys 24:337–355
33. Soltan A (1982) Mon Not R Astron Soc 200:115–122
34. Thorne KS (1974) Astrophys J 191:507–520
35. Wang J-M, Chen Y-M, Ho LC, McLure RJ (2006) Astrophys J Lett 642:L111–L114
36. Will CM (2006) Living Rev Relativ 9:3

Chapter 8
Are Anomalous Cosmic Flows A Challenge for LCDM?

Vincent Bouillot, Jean-Michel Alimi, Yann Rasera and André Füzfa

Abstract The dipolar moment of the peculiar velocity field, named bulk flow, is a sensitive cosmological probe: in parallel with the estimation issued from density fluctuations, it can give an indication on the value of cosmological parameters. Recent observations, based independently on composite velocity survey or velocity reconstruction from redshift survey, showed the existence of an anomalously high bulk flow in apparent contradiction with the linear prediction in ΛCDM cosmology. Using numerical simulations, we interpret this observation of consistently large cosmic flows on large scales as a signature of a rare event. Supposing we live in such a configuration and building samples with bulk flow profiles in agreement with the observations, we show that the asymmetric distribution of matter of large scales is responsible for the observed high bulk flow. To confirm the possible origin in ΛCDM cosmology of such a bulk flow profile and the agreement with linear theory description, we carefully study the time-dependence of the bulk flow and the distribution of matter.

8.1 Introduction

Recent observations such as velocity surveys [8] and redshift surveys [3] showed the existence of an excess in the dipolar moment of the velocity fields (i.e. bulk flow) at scales up to 50 h^{-1} Mpc. New observations, done through kinetic Sunyaev-Zeldovich effect, show this deviation can even be higher at larger scale [7]. Such discordance can be interpreted only in two ways: either the cosmological model we assume to predict the bulk flow is wrong, either we live in an environment with a very unlikely

V. Bouillot (✉) · J.-M. Alimi · Y. Rasera
LUTH, Observatoire de Paris, CNRS UMR 8102, Université Paris Diderot, 5 Place Jules Janssen, 92195 Meudon Cedex, France

A. Füzfa
Namur Center for Complex Systems, University of Namur (FUNDP), Namur, Belgium

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_8,

bulk flow profile. Several authors (e.g. Feldman [4]) favorize the cosmology as the cause of such a discrepancy.

On the other hand, the question of the realization of rare events is crucial in many fields in physics. Due to our unique position of observers, it is of utmost importance in cosmology to understand the behaviour of a given observable.

In this proceeding, we interpret the observation of the anomalously high bulk flow as a rare event realization. In particular, using high-resolution N-body simulations, we show that the bulk flow can be largely out of the linear prediction even though all the dynamics remain linear. To highlight the linearity of the dynamics, we draw a link between the velocity fields and the asymmetric distribution of matter and study the independence of this link on time evolution.

8.2 Realization of Rare Events

The velocity of matter in the Universe $\mathbf{u}(\mathbf{r})$ can be decomposed into the sum of the mean Hubble expansion $\mathbf{v}_{\mathrm{H}} = \mathrm{H}_0\mathbf{r}$ and a field of velocity fluctuations $\mathbf{v}(\mathbf{r}) = \mathbf{u}(\mathbf{r}) - \mathbf{v}_{\mathrm{H}}$. This field is called the peculiar velocity field.

The bulk (i.e. volume average) flow is defined as the mean of $\mathbf{v}(\mathbf{r})$ in a sphere of growing radius surrounding the observer. Assuming that each component of the peculiar velocity field is isotropic and follow a Gaussian distribution, the Fourier components are uncorrelated. Under those assumptions, the velocity power spectrum (convolved by a top-hat window function in Fourier space) describes entirely the statistical behaviour of the field [5]:

$$v_{bulk}(R) = \sqrt{\frac{1}{2\pi^2}\int_0^{\infty} k^2 P_v(k)\hat{W}(kR)^2 dk}. \tag{8.1}$$

Physically, we understand easily that the matter distribution source velocity fields. Especially, in linear theory, the velocity power spectrum and the density power spectrum are linked in a simple way:

$$P_v(k, z) = \frac{H^2(z) f^2(z)}{k^2} P_\delta(k, z), \tag{8.2}$$

where H is the Hubble factor and $f = \frac{d \ln \delta}{d \ln a}$ is the linear growth rate of density fluctuations. This relation is crucial to describe the time-dependence of the bulk flow since it relates the velocity fluctuations to the density fluctuations, which are evolving according to the linear growth rate D+ of a given cosmological model. Therefore, the time-dependence of the bulk flow can be written:

$$v_{bulk}(R, z) = \left\{ \frac{H(z) f(z) D_+(z)}{H(z=0) f(z=0) D_+(z=0)} \right\} v_{bulk}(R, z=0). \tag{8.3}$$

We are deeply interested in predicting the probability of obtaining the observed bulk flow profile in a given cosmology. For simplicity, we approximate the observational profile [8] by only two data points, namely a depletion point at radius R = 16 h^{-1} Mpc and a bump at radius R = 53 h^{-1} Mpc (see Fig. 8.1).

The probability of having a given value for the norm of the bulk flow is obviously not gaussian. In complete analogy with a gas of particles, the probability distribution of the norm of the velocity vector is a maxwellian. Extending this formalism to the case of the joint probability of having fixed velocities at radius R1 and R2 and following [6], we obtain a two-dimensional maxwellian distribution:

$$P(v_{R1}, v_{R2}) \propto \frac{2}{\pi} \frac{v_{R1}^2 v_{R2}^2}{|\det M|^{3/2}} \times \exp\left(-\frac{1}{2} {}^t\vec{V} M^{-1} \vec{V}\right) \text{ with } \vec{V} = (\|\vec{v}_{R1}\|, \|\vec{v}_{R2}\|) \tag{8.4}$$

Since the bulk flow is computed in a spherical volume, the correlation matrix M cannot be diagonal. The velocities at scale R2 are strongly dependent on the velocities at scale R1. As a result the correlation terms can be expressed as:

$$\gamma_{ij} = \sqrt{\frac{1}{2\pi^2} \int_0^\infty k^2 P_v(k) \hat{W}(kR_i) \hat{W}(kR_j) dk.} \tag{8.5}$$

Computing those quantities, we find that the probabilities of having such a bulk flow, given by the long tail of the maxwellian, is very scarce: only 1.4 % in ΛCDM.

8.3 Bulk Flow and Asymmetric Distribution of Matter

8.3.1 Numerical Set-up

We have performed a large set of high performance N-body simulations of large-scale structures with various dark energy components.

For this proceeding, we consider a simulation done within a cubic region of 648 h^{-1} Mpc with $1{,}024^3$ particles. Thanks to an adaptive mesh refinement method, the maximum resolution in the ΛCDM cosmology reaches 10 h^{-1} kpc.

The bulk flow is computed in a similar way to the definition given in section (2): from the tridimensional velocity fields of N_h objects in a sphere of radius R, we have:

$$v_{bulk}(R) = \left\| \frac{1}{N_{h,r<R}} \sum_{i}^{N_{h,r<R}} \vec{v}_i \right\| \tag{8.6}$$

We throw randomly 20.000 centres in the simulation volume, and around each center, we compute the bulk flow profile. From the whole environment, we could extract two interesting subsets (shown in Fig. 8.1):

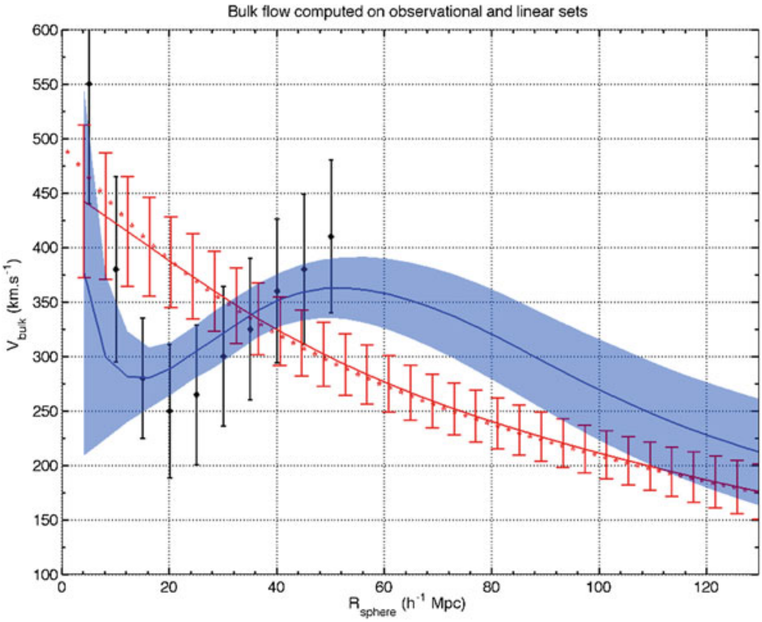


Fig. 8.1 Mean bulk flow profile vs radius of sphere. The observations [8] are indicated with black points; the linear prediction (Eq. (8.1)) with red stars; the linear catalogue is in red and the realistic catalogue is in blue

- Set of centres with a bulk flow profile close to the mean statistical prediction given in equation (1) at 95 % confidence level: this is the *linear* catalogue.
- Set of centres with a numerical bulk flow close to the bulk flow profile measured by [8] at 95 % confidence level. Since the mean bulk flow of this sample is on the mean in agreement with the observations, we call this subset *realistic*.

The probability to find a bulk flow profile in ΛCDM cosmology is given by the ratio of the number of elements in the *realistic* catalogue by the overall number of centres. The analysis of the Grand Challenge simulation exhibits 255 centres over 20.000 in the *realistic* catalogue. The numerical probability is then 1.27 %.

Considering the fact that this experimental probability is computed on all datapoints whereas the theoretical probability is computed from two points only, we have a good agreement between both the numerical and the statistical views. The Watkins-like profile can then be interpreted like a rare event realization.

We see from Fig. 8.1 that there is a strong disagreement for the bulk flow between the *realistic* sample and the linear prediction. Such a bulk flow profile traces possibly an asymmetric matter distribution. Indeed, the bulk flow is a vectorial quantity, which keeps track of local overdensities. Therefore, to understand bulk flow profile, we then have to quantify the local overdensities in a given direction with respect to the other direction.

8.3.2 Dynamical Origin of the Bulk Flow

How to characterize the asymmetry in a sphere of a growing radius R? Mathematically, the *asymmetry index* is obtained by maximizing the difference of the density fields ρ of the hemispheres at radius R:

$$A_R = \max_{\phi_0 \in [0,2\pi], \theta_0 \in [0,2\pi]} \left\{ \frac{1}{\rho_{mean}} \iint_{S^2/2} \rho_{<R}(\theta + \theta_0, \phi + \phi_0) \right.$$
$$\left. - \rho_{<R}\left(\pi - (\theta + \theta_0)\,,\, \pi + (\phi + \phi_0)\right) d\Omega \right\} \tag{8.7}$$

The *direction of the asymmetry index* is given by the direction of the north pole of the densest hemisphere. The more the asymmetry index is close to one, the more asymmetric the density field is.

Figure (8.2) shows the dependence of the asymmetry index on the radius for linear and realistic catalogues. The linear catalogue exhibits a constant decrease with radius whereas two main characteristics are clear on the realistic asymmetric index: from 40 to 76 h^{-1} Mpc, the realistic sample is more symmetric than the linear one; from 76 to 128 h^{-1} Mpc, the realistic catalogue is less symmetric than the linear. Due to the relation between gravity and velocity fields, one can wonder if the excess of the asymmetry around 80 h^{-1} Mpc is the cause of the bump of the bulk flow at 53 h^{-1} Mpc.

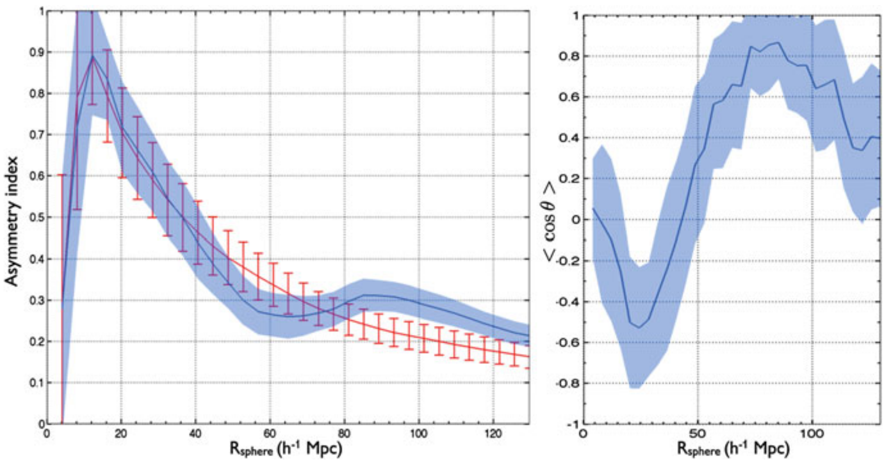


Fig. 8.2 Left panel: Asymmetry index vs radius of the sphere. Linear catalogue is in red line and the realistic sample is in blue. Right panel: Normalized scalar product of the differential direction of the asymmetry index and the bulk flow at 53 h^{-1} Mpc Vs radius

To answer this question, we have to compute the scale of alignment f the bulk flow at the bump position (53 h^{-1} Mpc) with the differential (i.e. in shells) asymmetry index. The right panel of Fig. (8.2) shows this normalized scalar product peaking around a particular scale. Therefore, the sourcing scale of the bulk flow is found to be 85 h^{-1} Mpc. More details as well as a full equivalence with a centre of mass approach can be found in [2].

8.4 Bulk Flow as a Linear Quantity

Even if the amplitude of the bulk flow is largely out of the mean linear prediction, this means only that we've to deal with a rare event. The linear nature of the bulk flow is confirmed by its linear time-evolution.

As a matter of fact, we compute the bulk flow and the asymmetry index through time for all objects of the realistic catalogue. Using Eq.(8.3), we renormalize bulk flow profiles at various redshifts with the bulk flow profile at z = 0. The same renormalization procedure is followed for the asymmetry index:

$$A_R(z) = \frac{D_+(z)}{D_+(z=0)} A_R(z=0). \tag{8.8}$$

Figure (8.3) shows clearly that the bulk flow (as well as the asymmetry index) renormalized by the linear evolution, remains the same through time. Since the asymmetry index is build from density contrasts, which is a scalar quantity, it is not surprising that the dynamical evolution is linear.

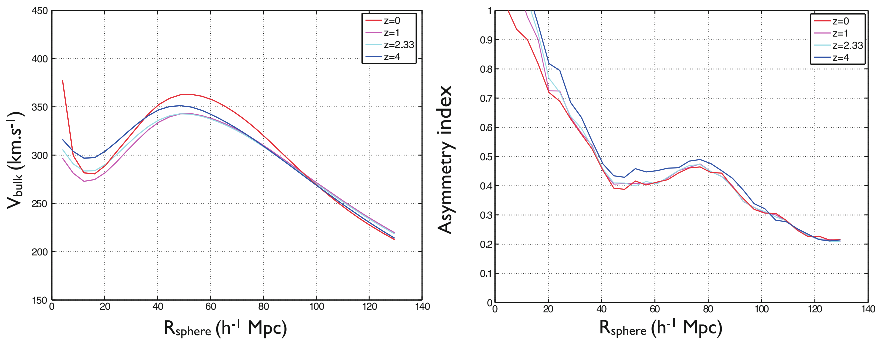


Fig. 8.3 Left panel: Bulk flow at different redshifts normalized to the bulk flow at z = 0. Right panel: Asymmetry index at various redshifts normalized to the asymmetry index at z = 0

8.5 Conclusion

Although the anomalously high amplitude of the bulk flow profile is often interpreted as a challenge for ΛCDM, we show that such a profile can be seen as the imprint of a rare event.

Using numerical simulations, we build samples with bulk flow profiles in agreement with the observations. The probability of having such events is allowed in a ΛCDM model with initial Gaussian condition and is 1.4 %, in agreement with the probability issued from our numerical catalogue.

The origin of such a bulk flow profile results from the asymmetric tridimensional distribution of matter. In particular, the bump of the bulk flow at 53 h^{-1} Mpc is explained by the asymmetric distribution of matter at 85 h^{-1} Mpc. Finally, we show that the time-evolution of the bulk flow for such samples is really predicted by the linear theory.

At larger scales (Fig. 8.1), the amplitude converges towards the expected amplitude of the mean linear prediction in ΛCDM. The distance between the bump and this point of reconvergence is a feature of cosmological models [1].

References

1. Alimi J.-M., Bouillot V, Rasera Y (2012) Constraining cosmology from large scale bulk flow MNRAS (in prep).
2. Bouillot V, Alimi J.-M., Füzfa A, Rasera Y (2012) Retrieving cosmological signal using cosmic flows MNRAS (in prep).
3. Erdogdu P, Lahav O, Huchra JP (2006) Reconstructed density and velocity fields from the 2MASS. Redshift Surv MNRAS 373:45
4. Feldman H-A, Watkins R, Hudson MJ (2010) Cosmic flows on 100 h^{-1} Mpc. MNRAS 407:23
5. Gramann M (1998) Power spectrum of velocity fluctuations in the universe. ApJ 493:28
6. Lavaux G, Tully RB, Mohayaee R, Colombi S (2010) Cosmic flow from two micron all-sky redshift survey. ApJ 709:483
7. Kashlinsky A, Atrio-Barandela F, Ebeling H, Edge A, Kocevski D (2009) A new measurement of the bulk flow of X-ray luminous clusters of galaxies. ApJ 712:81
8. Watkins R, Feldman HA, Hudson M-J (2009) Consistently large cosmic flows on scales of 100h^{-1} Mpc. MNRAS 392:743

Chapter 9
New Strength to Planck's Length Choice

Giuseppe Fazio, Mauro Giaconi and Davide Quatrini

Abstract The calculation of the total mass of a spacetime in which an observer-independent scale of length exist is possible through the use of Heisenberg Uncertainity Principle. In the following paragraphs we show that such a spacetime has a total mass equal to the mass of our Universe observed by the WMAP Nasa spacecraft if the cited observer-independent scale of length is determined by the Planck's length lp. This result gives new strength to fundamental theories that make use of lp as a length with a "special role", like the various string theories.

9.1 Introduction

The theoretical description of a general class of spacetimes in which observer-independent scales of both velocity and length exist can be found in recent-years published papers [1]. Besides, from another point of view, it is possible to affirm that a "stringy nature" (i.e. a Universe in which a string theory can be successfully applied) needs only two constants: a maximum velocity for satisfying the relativistic invariance principle and a minimum length for quantization [2]. For these reasons the two cited values are of great importance in the search for a theory that binds together the small and the large scales of the Universe, i.e. a theory that could lead to sinergies between General Relativity and Quantum Mechanics.

The debate related to the value of the maximum velocity needed for the relativistic invariance was solved long time ago with the selection of c, i.e. the speed of light in vacuum. This value is universally accepted because of its role in General Relativity.

Regarding the minimum length the debate is still open, but the most "appreciated" candidate is Planck's length lp, the only length that can be obtained combining General Relativity constants (c and G) together with the Quantum Mechanics constant $\hbar$:

G. Fazio (✉) · M. Giaconi · D. Quatrini
Electronic Engineering Department, Faculty of Engineering, University of Rome
"Tor Vergata", Via del Politecnico, 1–00133 Darmstadt, Rome

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_9,

$$lp = \sqrt{(\hbar G/c^3)} \sim 10^{-35}\,\text{m} \tag{9.1}$$

The aim of the present work is giving an evidence based on real experimental data to support the choice of Planck's length as the correct minimum length in our Universe. We give a simple way for calculating the mass of the Universe through the use of Planck's length and then we compare the obtained result with the data provided by the Nasa WMAP spacecraft. The selection of this process was obvious, for us, because we believe that the properties of such a "unifying" constant like the Planck's length can emerge only applying the rules and laws of Quantum Mechanics to large systems mainly ruled by General Relativity (like our Universe at large scale is).

9.2 Universe Mass' Calculation Through Heisenberg Uncertainity Principle

The calculation of the Universe mass is possible using the Heisenberg Uncertainity Principle (HUP). It is sufficient obtaining an expression for the mass from the best-case HUP for the length-momentum couple (i.e. selecting the = operator instead of the $\geq$) and then maximizing the obtained expression:

$$\Delta \times \Delta p = \hbar/2 \tag{9.2}$$

$$m\Delta \times \Delta v = \hbar/2 \tag{9.3}$$

$$m = (\hbar/2)/\Delta \times \Delta v \tag{9.4}$$

The maximization of this expression requires the selection of minimum values for $\Delta\times$ and Δv. According to [1] it is possible to select the Planck's length for the minimum value of $\Delta\times$, and then proceed from there for the selection of the minimum value of Δv. Using the simplest definition for velocity, i.e. v = s/t, we can minimizing it through the selection of the biggest amount of time and the smallest length. For this last one our selection is coherent with the above considerations, so we choose lp. For the biggest amount of time we can use the age of the Universe, or its good approximation given by the 1/H value, where H is the Hubble constant. In this way we obtain the smallest observable velocity at any given time in our spacetime, because calculated through the ratio between the smallest selected length (lp) and the largest possible observation period (the age of the Universe). Our research group is at work for finding deeper implications on the existence of the cited "minimum value of velocity", however, such as possible impacts on the quantization of fundamental quantities. With the chosen values the mass obtained is:

$$m = (\hbar/2)/(lp^*(lp^*H)) \tag{9.5}$$

$$m = (c^3)/(2GH) \sim 10^{52}\,\text{Kg} \tag{9.6}$$

This is the largest calculable mass in a spacetime in which HUP is valid and the Planck's length is the minimum length.

9.3 Universe Mass' Calculation Through 2008 WMAP Nasa Spacecraft Data

2008 WMAP NASA spacecraft data [3] confirmed that the the density of the Universe (ρ) is equal to its critical density calculated by the Friedmann equations [4], i.e. $3(H^2)/8\pi G$. This fact confirmed also the Euclidean geometrical structure for our Universe, leading to an estimation of Universe mass based on the assumption that its volume could be calculated as the volume of a sphere, i.e. $V = (4/3)\pi(r^3)$. In other words:

$$m = \rho V \tag{9.7}$$

$$m = ((H^2)(r^3))/(2G) \tag{9.8}$$

The selection of the value for the radius r can be done considering the space travelled by a ray of light for the entire duration of the Universe (Ua), i.e. r = c*Ua. Using the above mentioned approximation for which Ua = 1/H we obtain the following value for the mass of the Universe:

$$m = (c^3)/(2GH) \sim 10^{52}\ \text{Kg} \tag{9.9}$$

Please note that this value, and the expressions used for its calculation, are widely accepted by the scientific community, and are also used by Nasa itself in their official publications [5].

9.4 Conclusions

A spacetime in which Relativity applies, in which observer-independent scales of both velocity and length exist and are respectively determined by *c* and lp, and in which HUP is valid has a a total mass equal to:

$$m = (c^3)/(2GH) \sim 10^{52}\ \text{Kg} \tag{9.10}$$

In our Universe Relativity applies, an observer-independent scale of velocity exist and is determined by *c*, HUP is valid and the total mass is equal to:

$$m = (c^3)/(2GH) \sim 10^{52}\ \text{Kg} \tag{9.11}$$

For these reasons it is very probable that in our Universe an observer-independent scale of length exists and it is determined by Planck's length lp.

References

1. Amelino-Camelia G (2001) Testable scenario for relativity with minimum length. Phys Lett B 510:255–263
2. Veneziano G (1986) A Stringy Nature Needs Just Two Constants. Europhys. Lett. 2:199
3. NASA / WMAP team (2008) WMAP Press Release: WMAP reveals neutrinos, end of dark ages, first second of universe. http://map.gsfc.nasa.gov/news/5yr_release.html
4. Friedmann A (1922) Über die Krümmung des Raumes. Z. Phys. 10(10):377–386
5. NASA Glenn Research Centre: On the expansion of the Universe. http://www.grc.nasa.gov/WWW/K-12/Numbers/Math/documents/ON_the_EXPANSION_of_the_UNIVERSE.PDF

Part III
Particle Physics and High Energy Physics

Chapter 10
Results from the Atlas Experiment at the LHC

Antonio Sidoti

Abstract ATLAS is a general purpose detector located at one of the four interaction points of the Large Hadron Collider at the CERN laboratory near Geneva, Switzerland. In 2010 and 2011, LHC collided proton beams at the unprecedented center of mass energy of 7 TeV and the ATLAS detector has collected more than 5.25 fb^{-1} of data. The detector performance and results for Standard Model and Beyond the Standard Model physics are presented here.

10.1 Introduction

The LHC collider at CERN started its operation with first proton–proton collisions in 2009. After a commissioning phase, thanks to the excellent performance of the CERN Accelerator Division crew, the ATLAS detector collected 5.25 fb^{-1} of proton–proton collisions at $\sqrt{s} = 7$ TeV respectively in 2011. In these proceedings, we will report on the latest physics results obtained by the ATLAS experiment. We will show the Standard Model (SM) physics measurements performed with already competitive precision with respect to former collider experiment. Comparing those against predictions from SM will lead to evidences of Physics Beyond Standard Model if discrepancies are found. Direct searches of SM Higgs boson, supersymmetric or exotic physics searches will be shown.

On behalf of the ATLAS collaboration.

A. Sidoti (✉)
Istituto Nazionale Fisica Nucleare—Sezione Roma, Rome, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_10,

10.2 Standard Model Physics Measurements

A detailed description of the ATLAS detector can be found in Aad (2008). The more important SM processes have been tested at the unprecedented center of mass energy of 7 TeV. We will show various SM measurements performed by ATLAS starting from processes with larger cross sections to the ones with smaller ones.

10.2.1 Hard QCD Results

The inclusive jet and the dijet mass cross section are measured in different rapidity regions. The inclusive jet measurement is performed in the p_T range 20 GeV–1.5 TeV,[1] and it extends to a wide rapidity range ($|y| < 4.4$) (Fig. 10.1: Left). Jets are reconstructed using the *anti*-k_T algorithm with different R distance parameters ($R = 0.4$ and 0.6). These different R parameters make the jets differently sensitive to the hadronization, underlying event and pile up. Jet energies are corrected for detector effects taking into account non linearity due to the different calorimetric response to electromagnetic and hadronic showers, the energy loss in non instrumented material, and the bending of charge particles in the magnetic field. The measured cross sections are unfolded to the particle level and have been compared with pure next to leading order (NLO) predictions, corrected for the nonperturbative effects, and with POWHEG Monte Carlo simulations, which perform a NLO prediction coherently interfaced with parton shower, hadronization and underlying event simulations [2].

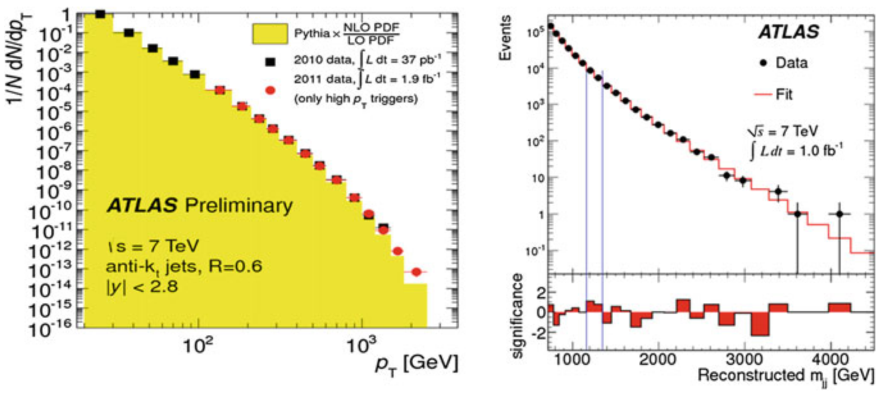


Fig. 10.1 *Left* Inclusive Jet P_T differential cross section $d\sigma/dP_T$ measured with 2010 and 2011 data [2]. *Right* Dijet mass distribution. The *blue lines* indicates the mass region where the deviation from the fit with a smooth functional form describing the QCD background theoretical expectations is more important. However, the significance of this region is well below the two standard deviations [3]

[1] In the following we will use the natural units where $c = \hbar = 1$.

In the measurement of the dijet mass cross section (Fig. 10.1: Right), the first two jets, ordered in decreasing p_T, are used to measure the dijet invariant mass M_{jj}, and the angular variable $|y|_{max} = max(|y|_1; |y|_2)$. The leading jet p_T is required to be above 30 GeV, and the second above 20 GeV. They are both selected in the rapidity range $|y| < 2.8$. The dijet mass cross section covers a range from 70 GeV to 4 TeV and shows no evidence of resonance production over background. Limits are set at 95 % CL for several new physics hypotheses: excited quarks are excluded for masses below 2.99 TeV, axigluons are excluded for masses below 3.32 TeV and colour octet scalar resonances are excluded for masses below 1.92 TeV [3].

10.2.2 W and Z Boson Cross Section Measurements

At hadron colliders W and Z bosons can be identified only in their leptonic decay final states. For the W boson, the presence of the undetected neutrino in the final state gives an event signature with one isolated high P_T lepton and large missing transverse energy (ME_T). The kinematics of the event cannot be completely reconstructed since the longitudinal information of the missing energy is missing. The transverse mass m_T is then defined as ${m_T}^2 = 2E_T ME_T(2 - \cos\varphi)$ where φ is the angle between the lepton p_T and the missing transverse momentum. The transverse mass is shown (Fig. 10.2: Left) in the W→eν channel with the expected backgrounds. The dominant one is given by "QCD background", (dijet final states) where one jet mimics the electron and the other one is missed or mismeasured giving a large ME_T.

The Z boson final state is characterized by two high p_T, opposite sign and isolated leptons. The W and Z production cross section have been measured taking into account the background estimations, the trigger and reconstruction efficiencies, the geometric and kinematic acceptance and the integrated luminosity. The results are

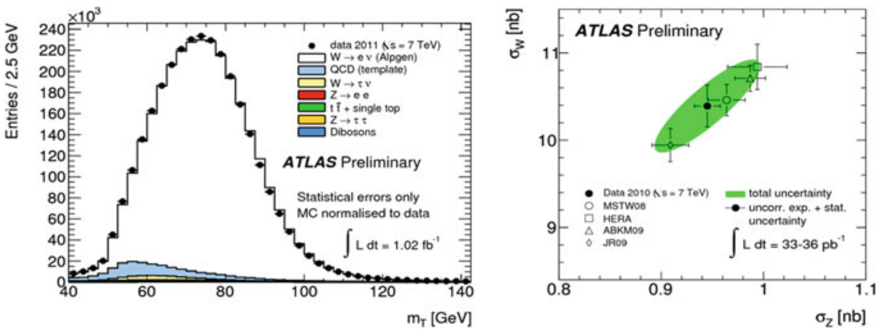


Fig. 10.2 *Left* Transverse mass distribution for W→eν candidates. QCD background has been determined from data while the other background contributions are derived from Monte Carlo simulations [4]. *Right* W boson measured cross section as a function of the Z boson measured cross section together with the total uncertainty. The theoretical expectations are reported using different Parton Distribution Functions (PDF) [4]

shown in (Fig. 10.2 Right) together with expectations using different Parton Distribution Functions (PDF) sets.

10.2.3 Top Quark Measurements

The top quark is the heaviest elementary particle of the SM. At the LHC it can be produced in top–antitop pairs *via* the strong interactions or as *single top* with electroweak processes. In the following we will focus on the top quark pair production process. Once produced, the top quark decays with an electroweak process in $t \to Wb$.[2] The different decay modes of the W boson determine the final state of the top–antitop pair. We can identify three main final states:

- the *dilepton* one, where both W decay leptonically in electron or muon. This final state is characterized by an excellent S/B ratio and a small branching fraction (BF)
- the $lepton + jet$ one, where one W decays in electron/muon and the other hadronically. This final state is characterized by a larger BF and a slightly larger background contamination
- the *all-hadronic* one, where both W bosons decay hadronically. This final state is characterized by the largest BF, but the signal is overwhelmed by the QCD multijet production.

ATLAS has measured the top–antitop production cross section in all these decay channels, also including final states with tau lepton. The measured cross section are shown (Fig. 10.4: Left) together with the theoretical expectations (NNLO) for $M_{top} = 172.5$ GeV for the dilepton and lepton+jet channel.

The top quark mass is one of the key parameters in the electroweak fit. The three-jet reconstructed top mass is reported in the muon+jet channel with the various background contributions (Fig. 10.4: Right). After combination with the electron channel, the top mass has been measured to be:

$$M_{top} = 175.9 \pm 0.9\,(stat.) \pm 2.7\,(syst.)\,GeV$$

So far, in all these measurements measured already with small uncertainties, no deviations from SM expections have been found.

10.3 Direct Searches

In this section we will show direct searches of the Standard Model Higgs boson and Supersymmetric processes.

[2] Assuming $|V_{tb}|=1$.

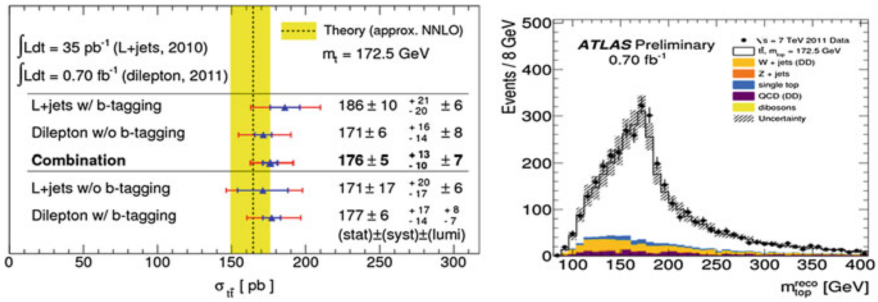


Fig. 10.3 *Left* Measured top–antitop cross section for the lepton+jet channel @ $M_{top} = 172.5$ GeV with the NNLO theoretical prediction [6]. *Right* Three-jet reconstructed mass in the muon+jet channel. The top quark mass measurement is obtained comparing the data distribution with distributions obtained with different M_{top} hypothesis [7]

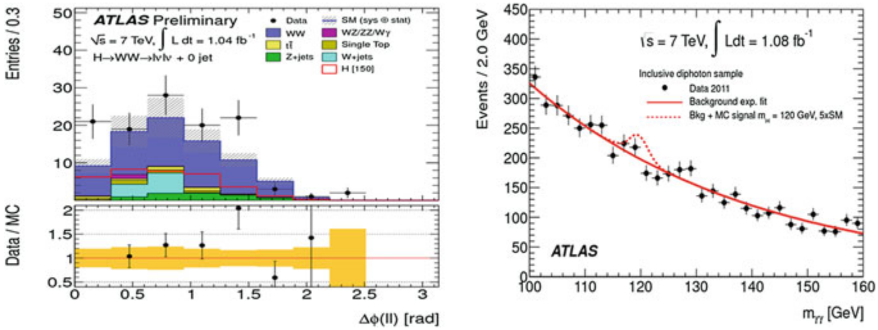


Fig. 10.4 *Left* Distribution of the diphoton reconstructed invariant mass. The expected signal for a Higgs boson of 120 GeV mass with five times the SM production cross section is shown [10]. *Right* $H \rightarrow WW^{(*)} \rightarrow$lνlν selection, 0-jet bin: the azimuthal opening angle $\Delta\varphi$ll of the two selected leptons after preselection cuts [11]

10.3.1 SM Higgs Boson Searches

Higgs boson discovery or exclusion is one of the main goals of the ATLAS experiment. A review of theoretical predictions can be found in Ref. [8]. Looking at the $\sigma \times$ BF[3] as a function of the Higgs boson mass we can identify two regions: a "high mass region" ($M_H > 140$ GeV) where the most sensitive channels are the diboson decays H→ $WW^{(*)}$ or H→ $ZZ^{(*)}$ and a "low mass region" ($M_H < 140$ GeV) where the most sensitive channels are H→ $\gamma\gamma$, H→ $\tau\tau$ and H→bb for the Vector Boson Fusion production process with still some contributions from diboson decays.

The most sensitive channel in the "low mass" region is the decay into a pair of photons. This decay channel has a low branching ratio (2×10^{-3} for M_H = 120 GeV), a clear signature and thus a good sensitivity for $M_H < 140$ GeV. The

3 Production cross section times branching fraction of the decay.

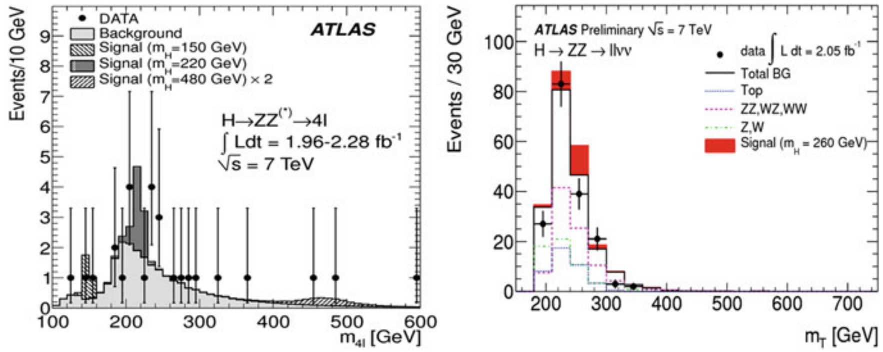


Fig. 10.5 *Left* Distribution of the reconstructed invariant mass in the four lepton channel. Invariant mass distributions for different Higgs boson mass hypothesis (M_H =150, 180 and 480 GeV) are shown [12]. *Right* Transverse mass distribution of the H→ZZ(*)→llνν candidates shown together with the expectations from a Higgs boson mass of $M_H = 260$ GeV [13]

search is performed in the $110 < M_H < 150$ GeV mass range. The candidate events are separated in different categories according to the presence of converted photons and impact points in the calorimeter improving the overall sensitivity. The main backgrounds in this channel are: diphoton production (irreducible), photon-jet production with one fake photon, di-jet production with two fake photons and Drell–Yan events where both electrons are misidentified as photons. The diphoton invariant mass for the 1.08 fb^{-1} data sample is reported (Fig. 10.5: Left) together with five times the expected contribution from a Higgs boson with $M_H = 120$ GeV. In the "high mass region" one of the most promising channels is the H→ $WW^{(*)}$ decay channel. In particular for the $120 < M_H < 200$ GeV mass interval the full leptonic channel H→ $WW^{(*)}$ →lνlν is the most sensitive. Given the fact that the Higgs boson is scalar, the angle between the two leptons φ(ll) can be used as a discriminating variable between the SM Higgs boson production process and its main backgrounds (dilepton WW/WZ production,W + jets, etc.) (Fig. 10.5: Right).

H→ $ZZ^{(*)}$ candidates are searched in the $ZZ^{(*)}$ →llνν and $ZZ^{(*)}$ →llqq final states and the "four lepton golden mode" channel $ZZ^{(*)}$ →llll. Figure 10.6 (Left) shows the four lepton invariant mass with three different Higgs boson mass hypothesis. The former channels contribute with larger BF but with a larger background contamination. Figure 10.6 (Right) shows the transverse mass distribution for the H→ $ZZ^{(*)}$ →llνν candidates.

No significant excess of events is found with respect to the expectations from SM processes, therefore exclusion limits on the production cross section are set at 95 % Confidence Level (CL) with a statistical analysis based on the CL method [9]. The contributions from the different channel are shown individually together with their statistical combination (Fig. 10.7: Left). These results have been combined with CMS ones to maximize the overall sensitivity. The observed data are compatible with the background only hypothesis and the SM Higgs boson is excluded at 95 % CL or

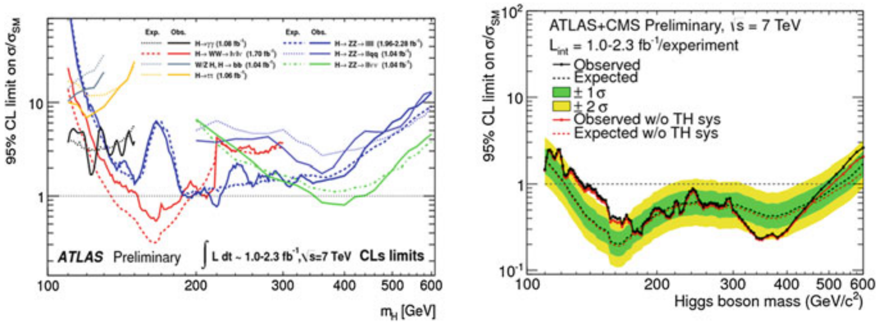


Fig. 10.6 *Left* The expected (*dashed*) and observed (*solid*) cross section 95 % exclusions limits for the individual search channels, normalised to the Standard Model Higgs boson cross section, as functions of the Higgs boson mass [14]. *Right* The combined ATLAS-CMS 95 % CL upper limits on the σ / σ_{SM} ratio obtained with the CLs method, as a function of the SM Higgs boson mass in the range 110–600 GeV. The observed limits are shown by solid symbols. The *dashed line* indicates the median expected value for the background-only hypothesis, while the *green* (*yellow*) bands indicate the ranges expected to contain 68 % (95 %) of all observed limit excursions from the median. The limits obtained without the theoretical systematic uncertainties are also shown for comparison [15]

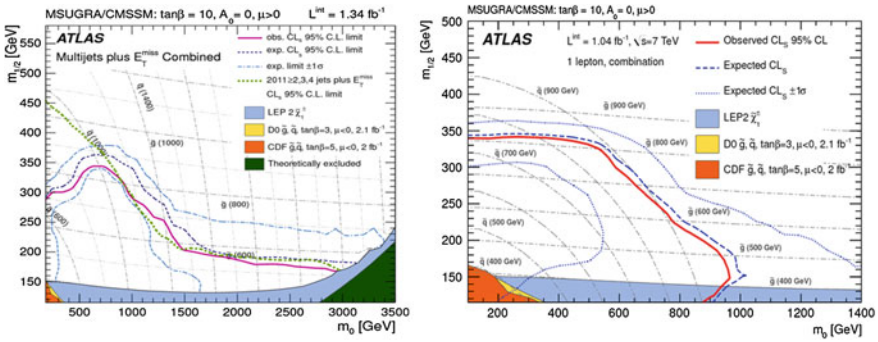


Fig. 10.7 *Left* Combined exclusion limits from the 0-lepton analysis in the $(m_0, m_{1/2})$ plane for mSUGRA/CMSSM models with $\tan\beta = 10$, $A_0 = 0$ and $\mu > 0$. The *dashed-blue* and the *red lines* correspond to the expected and observed 95 % CL limits, respectively. The *dotted blue lines* correspond to the $\pm 1\,\sigma$ variation in the expected limits. The observed ATLAS limit from 2010 analysis ($\tan\beta = 3$) is shown by the *solid black line*. The *star* indicates the position of the mSUGRA reference point with $m_0 = 660$ GeV, $m_{1/2} = 240$ GeV, $A_0 = 0$, $\tan\beta = 10$ and $\mu > 0$ [16]. *Right* observed and expected 95 % CL exclusion limits from the 1-lepton analysis, as well as the $\pm 1\,\sigma$ variation on the expected limit, in the combined electron and muon channels [17]

higher in the mass range 141–476 GeV. The region from 146 to 443 GeV is excluded at the 99 % CL (Fig. 10.7: Right).

10.3.2 SUSY Searches

SUSY is one of the most popular extensions of the Standard Model of particle physics. Each fundamental Standard Model fermion has a supersymmetric boson partner and vice versa. SUSY solves the mass hierarchy problem. A quantum number R is defined to be $R = (-1)^{2s+3B+L}$.[4] In the following we will focus on the SUSY class of theories where R-parity is conserved. In those cases, the Lightest Supersymmetric Particle (LSP) is produced at the end of all *s*-particle cascade-decay and is stable. LSP are natural candidates for dark matter. LSP are also weakly interacting and escape detection giving large transverse missing energy. The production of colored *s*-particles could be possible at the LHC giving final states with high p_T jets. Also additional high P_T leptons can be produced from model dependent cascades. Final states are then characterized by several high momentum jets, large missing transverse energy and 0, 1 or 2 high momentum leptons. Here we will show the results in the 0-lepton and 1-lepton bin. No excess of events over the background expectation is seen in any signal region, and exclusion limits have been set in the $(m_0, m_{1/2})$ mSUGRA/CMSSM plane (Fig. 10.7). Squarks and gluinos of equal mass have been excluded at 95 % CL for masses below 980 GeV.

10.4 Conclusions

The ATLAS detector is working well since the start of LHC collisions in 2009. The integrated luminosity of proton-proton collisions delivered by LHC is awesome. We have shown only a minor part of the physics measurement performed by ATLAS. The differential jet cross section as a function of the momentum has been shown together with the dijet mass distribution that excludes several new physics hypotheses. The W and Z cross section has been measured in the electron, muon and tau channel. We have reported the top-antitop cross section measurement in most of the final states and the top mass measurement in the letpon+jet channel. Direct searches of SM Higgs boson in the most promising final states have been presented. Finally, direct SUSY searches have been shown. The analysis presented there are based at most on 2.28 fb^{-1} and analysis are progressing fast on the remaining data collected in 2011. So far no direct or indirect indications of Higgs boson or Physics Beyond SM have been observed. For 2012 we expect at least to double the statistics collected so far.

10.5 Post Scriptum

At the time of writing of this paper, the ATLAS Collaboration has reported a substantial update on the search for the SM Higgs boson in the gamma-gamma and

[4] With spin S, baryon number B, and lepton number L. All Standard Model particles have R-parity of 1 while supersymmetric particles have R-parity -1.

four-lepton final states, based on the analysis of the full 2011 data, and on their combination with the results already available during the summer. The results of this updated study are shown in Ref. [18].

References

1. JINST (2008) The ATLAS Experiment at the CERN Large Hadron Collider. 3:S08003.
2. ATLAS (2011a) ATL-COM-PHYS-2011-1082. http://cdsweb.cern.ch/record/1375181
3. ATLAS (2011b) CERN-PH-EP-2011-127, arXiv: 1108.6311 (hep-ex).
4. ATLAS (2011c) CERN-PH-EP-2011-143, arXiv: 1109.5141 (hep-ex).
5. ATLAS (2011d) CERN-PH-EP-2011-122 arXiv: 1108.4101 (hep-ex).
6. ATLAS (2011e) ATLAS-CONF-2011-108. http://cdsweb.cern.ch/record/1373410
7. ATLAS (2011f) ATLAS-CONF-2011-120. http://cdsweb.cern.ch/record/1376412
8. Djouadi A (2011) Higgs: theory. In: Proceedings of Lepton-Photon 2011, Mumbai.
9. Read AL (2011) NJ Phys G 28:2963
10. ATLAS (2011g) arXiv: 1108.5895 (hep-ex).
11. ATLAS (2011h) ATLAS-CONF-2011-134. http://cdsweb.cern.ch/record/1383837
12. ATLAS (2011i) arXiv: 1109.5945 (hep-ex).
13. ATLAS (2011j) arXiv: 1109.3357 (hep-ex).
14. ATLAS (2011k) ATLAS-CONF-2011-135. http://cdsweb.cern.ch/record/1383838
15. ATLAS (2011l) ATLAS-CONF-2011-157. http://cdsweb.cern.ch/record/1399599
16. ATLAS (2011m) CERN-PH-EP-2011-155, arXiv: 1110.2299 (hep-ex).
17. ATLAS (2011n) CERN-PH-EP-2011-153, arXiv: 1109.6606 (hep-ex).
18. ATLAS (2011o) ATLAS-COM-CONF-2011-163. http://cdsweb.cern.ch/record/1406358

Chapter 11
Covariant Perturbations Theory in General Multi-Fluids Cosmology

Vincent Bouillot, Jean-Michel Alimi and Cristiano Germani

Abstract We develop a variational approach, inspired by the work of Maldacena (Maldacena 2003, JHEP 0305), to study covariant cosmological perturbations in general multi-fluids extended gravity. A special attention is paid to the minimization of the propagating degrees of freedom to obtain simple equations of motion. In particular, by parametrizing perfect fluids with scalar fluids and by working within the ADM formalism, we manage to introduce new gauge invariant quantities, generalizing Mukhanov-Sasaki variables. Those developments are crucial to deeply understand the influence of general extended gravity either in the inflationary epoch or in the late-time inflation. In this proceeding, we present only the quadratic Lagrangian at first order for scalar perturbations deduced from the gravitational constraint equations.

11.1 Introduction

The predictions of General Relativity are in good agreement with the cosmological and astronomical observations only if we suppose the evolution of our Universe is driven by unknown forms of energy: Dark matter and Dark energy. Dark matter dominates the gravitational attraction at small scales and seems well described by non-relativistic particles, weakly coupled to baryons. At larger scales, the evolution of the Universe can be associated with a form of energy behaving like a cosmological constant.

V. Bouillot (✉) · J.-M. Alimi
LUTH, Observatoire de Paris, CNRS UMR 8102, Université Paris Diderot,
5 Place Jules Janssen, 92195 Meudon Cedex, France

C. Germani
Arnold Sommerfeld Center, Ludwig-Maximilians-University, Theresienstr. 37,
80333 Muenchen, Germany

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_11,

The interpretation of dark energy purely in term of a new constant of Nature suffers however several problems. To go beyond this simple proposal, we can make this cosmological constant vary, thus introducing a scalar field.

The main goal of this proceeding is to describe **in full generality** the equations of motion for wide classes of dark energy models i.e. dark energy as a scalar field (non-)minimally coupled to multiple matter components.

In the observable frame, named Dicke-Jordan frame, the action we consider writes down:

$$S_{BD} = \int d^4x\sqrt{-\tilde{g}}\left[\frac{\phi}{2\kappa^2}\tilde{R} - \frac{\omega(\phi)}{2\kappa^2\phi}\partial_\mu\phi\partial_\upsilon\phi + \tilde{U}(\phi) + \sum_i \tilde{P}(\tilde{X}_{\psi_i})\right] \quad (11.1)$$

with $\kappa^2 = 8\pi G$ the gravitational coupling and $\omega(\phi)$ the Dicke-Jordan coupling function.

11.2 Perfect Fluid as a Scalar Field

To minimize the degrees of freedom, we have to parametrize the matter part of action Eq. (11.1).

A barotropic and irrotational perfect fluid propagates only one degree of freedom. Therefore, at Lagrangian level, the question of the treatment of perfect fluids in terms of scalar fields can be asked. The answer is positive as shown in [1]. In the following, we develop this formulation in the Dicke-Jordan frame (i.e. observational frame).

The energy-momentum tensor of a perfect fluid reads:

$$\tilde{T}_{\mu\upsilon} = (\tilde{\rho} + \tilde{p})\tilde{u}_\mu\tilde{u}_\upsilon + \tilde{p}\tilde{g}_{\mu\upsilon} \quad (11.2)$$

where $\tilde{\rho}$ and $\tilde{p}$ are respectively the energy density and pressure of the fluid in the Dicke-Jordan, while $\tilde{u}_\mu$ is the fluid four-velocity.

Introducing the Lagrangian $L = \tilde{P}(\tilde{X})$ with $\tilde{X} \equiv -\tilde{g}^{\mu\upsilon}\partial_\mu\psi\partial_\upsilon\psi$, associated to the pressure of a scalar field, and varying it with respect to the metric, we find the stress-energy tensor:

$$\tilde{T}_{\mu\upsilon} = 2\tilde{P}(\tilde{X})_{,X}\partial_\mu\psi\partial_\upsilon\psi + \tilde{P}(\tilde{X})\tilde{g}_{\mu\upsilon} \quad (11.3)$$

This tensor looks quite similar to Eq. 11.2. An identification can be done in the following way:

$$\tilde{\rho} = 2\tilde{X}\tilde{P}_{,X} - \tilde{P}, \quad \tilde{p} = \tilde{P} \text{ and } \tilde{u}_\mu = \frac{\partial_\mu\psi}{\sqrt{\tilde{X}}}. \quad (11.4)$$

Specializing to a barotropic fluid i.e. $\tilde{p} = \omega\tilde{\rho}$, we can explicitly write the pressure and the energy density in terms of $\tilde{X}$ alone (At homogeneous order: $p_i = A_i^{2(2-\alpha)}\dot{\psi}_i^{2\alpha}$):

$$\tilde{P} = \tilde{X}^{\alpha} \text{ with } \alpha = \frac{1+\omega}{2\omega} \tag{11.5}$$

for $w \neq 0$. However, as proven in [1], the limit $w \to 0$ can be taken at the level of the equations of motion and corresponds to the case of dust (baryons or cold dark matter). This case is of special interest since it describes the energy density component of baryons and cold dark matter in the late time Universe.

In this proceeding, we are interested in the case in which a scalar field ϕ is, in full generality, non-minimally coupled to gravity. In this case, the action in Einstein-Hilbert frame can be rewritten via a conformal transformation $\tilde{g}_{\alpha\beta} = A^2_{\psi}(\varphi) g_{\alpha\beta}$ with $A_i(\varphi)$ the conformal coupling:

$$S_{EH} = \int d^4x\sqrt{-g}\left[\frac{1}{2\kappa^2}R + \frac{1}{2}(\partial\varphi)^2 - V(\varphi) + \sum_i P(X_i, A_i)\right] \tag{11.6}$$

with $\kappa^2 = 8\pi G$ the gravitational coupling and $\omega(\phi)$ the Dicke-Jordan coupling function.

11.3 Constraints at Homogeneous Order

The consistency of our approach can be checked through the derivation of the Friedmann and conservation equations. Since the Friedmann equations are resulting of the variation of the action with respect to the metric $g_{\mu\nu}$, they remained unchanged under our ersatz.

We assume a flat Friedmann-Lemaître-Robertson-Walker (FLRW) spacetime with metric:

$$ds^2 = -dt^2 + a(t)^2\delta_{ij}dx^i dx^j \tag{11.7}$$

Using this metric, action (11.6) and the perfect fluid ersatz, we find the Klein-Gordon equation for an interacting scalar fluid and a set of equations describing the motion of various 'matter' scalar fields.

$$\ddot{\varphi} + 3H\dot{\varphi} + V_{,\varphi} + \sum_i \frac{d\ln A_i^{2(2-\alpha)}}{d\varphi}(\rho_i - p_i) = 0 \tag{11.8}$$

$$(2\alpha - 1)\ddot{\psi}_i + 3H\dot{\psi}_i + \frac{d\ln A_i^{2(2-\alpha)}}{dt}\dot{\psi}_i = 0$$

For simplicity, in this proceeding,[1] we suppose all the 'matter' fields have the same physical origin: we consider only one α parameter that has the same equation

1 The full case is described in [2]

of state for all matter components. However, every component has its own coupling function A_i to gravity.

11.4 Toward Observables at Linear Order

11.4.1 Linear Perturbations

To express observables at linear order, we split our perturbed metric using the ADM formalism:

$$ds^2 = -N(t)^2dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt) \tag{11.9}$$

Thanks to diffeomorphism invariance, we can choose a gauge such that: $N = 1 + \delta N$, $N^i = a^{-2}\partial_i\beta$, $h_{ij} = a^2 e^{2\zeta}\delta_{ij}$ and $\delta\varphi = 0$. As soon as we're dealing with perturbations in cosmology, the question of the gauge is crucial. In fact, two gauge choices are preferred in our formalism: the unperturbed scalar field gauge ($\delta\varphi = 0$) and the spatially flat gauge ($\zeta = 0$). However, since the scalar field φ is non-minimally coupled to all other 'matter' scalar fields, it plays a very specific role. Therefore, the gauge comoving with the scalar field is the most natural gauge choice. In this gauge, the action writes down:

$$S = \int d^4x\sqrt{\det(h_{ij})}\left\{\frac{1}{2\kappa^2}\left[N^{(3)}R + N^{-1}(E_{ij}E^{ij} - E^2)\right] + N^{-1}\frac{\dot{\varphi}}{2} - NV(\varphi) + N\sum_i P(X_i, A_i)\right\} \tag{11.10}$$

The main feature of the ADM formalism is the non-propagation of the metric components N and N^i. Instead of having two propagating equations, we obtain the Hamiltonian and momentum constraints as follows:

$$\begin{aligned}
\delta N &= \frac{\dot{\zeta}}{H} + \frac{\kappa^2}{H}\sum_i \alpha P_i \frac{\delta\psi}{\dot{\psi}} \\
H\partial^2\beta + \partial^2\zeta - 3H\dot{\zeta} &= \sum_i \alpha(1-2\alpha)\kappa^2 P_i \frac{\delta\dot{\psi}}{\dot{\psi}} \\
&\quad + \sum_i (1-\alpha)(1-2\alpha)\kappa^2 P_i \delta N - \kappa^2 \delta N V
\end{aligned} \tag{11.11}$$

Substituting these two constraints into the action (11.10), we find the quadratic action on linear perturbations.

Please note that a dot is a derivative with respect to time whereas, in the following, a prime denotes a derivative with respect to conformal time.

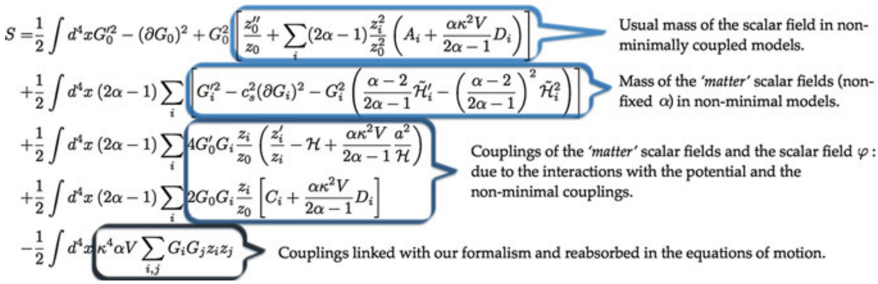


Fig. 11.1 Quadratic action at first order

11.4.2 Quadratic Action at First Order

During the computation of the quadratic action at first order, a generalized form of the Mukhanov-Sasaki variables, widely used in inflation, arises. Those quantities, formally equivalent to the potential of the velocity fields, are gauge invariants:

$$G_0 = Hz_0\left(\frac{\zeta}{H} - \frac{\delta\varphi}{\dot{\varphi}}\right) \text{ and } G_i = Hz_i\left(\frac{\zeta}{H} - \frac{\delta\psi_i}{\dot{\psi}_i}\right) \text{ with } z_\mu = \kappa a\frac{\sqrt{\rho_\mu + p_\mu}}{H} \tag{11.12}$$

Therefore, the gauge invariant quadratic action at first order is given in Fig. 11.1. The quantities $\tilde{H}_i$, A_i, C_i and D_i determining the evolution of the perturbations are only functions of the background variables and the couplings.

References

1. Boubekeur L, Creminelli P, Norena J, Vernizzi F (2008) Action approach to cosmological perturbations: the second-order metric in matter dominance JCAP 0808:028
2. Alimi J-M, Bouillot V, Germani C (2012) Linear perturbations in coupled models JCAP (in prep).
3. Maldacena J (2003) Non-gaussian features of primordial fluctuations. JHEP 0305:013
4. Mukhanov VF, Feldman HA, Brandenberger RH (1992) Theory of cosmological perturbations. Phys Rept 215:203

Chapter 12
Susy Results at the LHC with the Atlas Detector

Simone Brazzale

Abstract The data collected during 2011 with the ATLAS detector has been used to perform searches for signals due to R-parity conserving supersymmetry. The results of different analyses targeting jets, isolated leptons and missing transverse energy in the final state, a promising venue for the discovery of supersymmetry, are presented in some details in the following.

12.1 Introduction

Supersymmetry (SUSY) is one of the most popular extensions of the Standard Model (SM). SUSY relates each elementary SM particle of one spin to another particle named superpartner, from which it differs by half a unit of spin (Fig. 12.1).

In case of R-parity[1] conservation, charginos and neutralinos[2] can be produced directly in pairs or in the decay chains of squarks and gluinos. Moreover, the Lightest Supersymmetric Particle (LSP) is stable and therefore escapes any detector, while charginos decay into LSPs can yield high-pT leptons. A common signature for SUSY searches at high energy colliders is thus a high Transverse Missing Energy (E^T_{miss}) due to the LSPs, multiple high energetic jets from quark hadronization and eventually additional leptons coming from $X^{\pm}_1$ decay to the LSP (Fig. 12.2).

The ATLAS Collaboration has searched these final states during the first year of proton-proton collisions at a centre-of-mass energy of 7 TeV, which have been delivered by the LHC in 2011.

[1] $R = (-1)^{3(B-L)+2s}$, where B, L and S are respectively the barionic number, the leptonic number and the spin.

[2] Charginos ($X^{\pm}_i$) and neutralinos (X^0_i) are mass eigenstates of the superpartners of the bosons.

S. Brazzale (✉)
Università di Udine e INFN Gruppo Collegato di Udine, Udine, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_12,

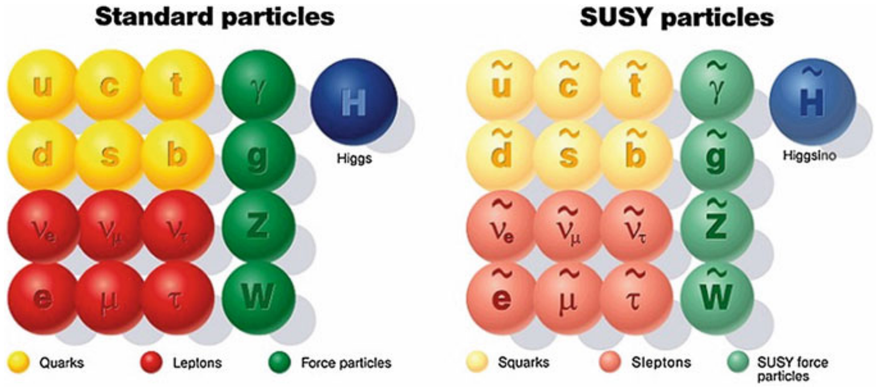


Fig. 12.1 SM particles and their supersymmetric partners in the SUSY scenario

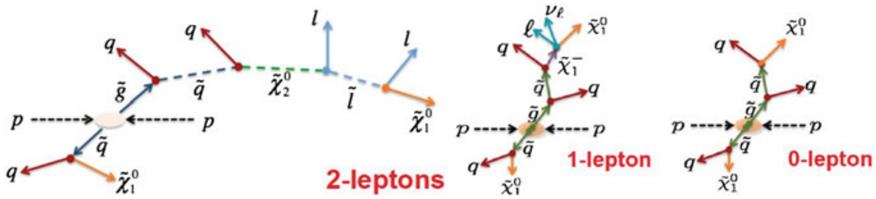


Fig. 12.2 Example of proton-proton interactions where supersymmetric partners are produced in pairs: the final state includes E^T_{miss}, some jets and eventually one or more leptons

During this year, the ATLAS detector has worked remarkably well, recording around $5\,\mathrm{fb}^{-1}$ of collisions with an overall data taking efficiency of 94 %. The proton run ended in October, when the LHC switched to heavy ions collisions.[3]

12.2 0-Lepton Final State

When squarks and gluinos decay directly to quarks and LSPs (Fig. 12.2), leptons are not produced in the decay chain and therefore no electrons nor muons appear in the final state. This case is favoured when the squark and gluino masses are not heavy enough to decay into heavier neutralinos or charginos. With $1\,\mathrm{fb}^{-1}$ of ATLAS data, the number of expected SM events has been compared with the number of observed events in the 0-lepton final state, and no excess has been found over the expectation. The SM background processes, such as W/Z+jets, top pair production and QCD multijets, have been estimated by defining different Control Regions (CR) and by

[3] All the analysis presented in this paper apply to $1\,\mathrm{fb}^{-1}$ of proton–proton collisions.

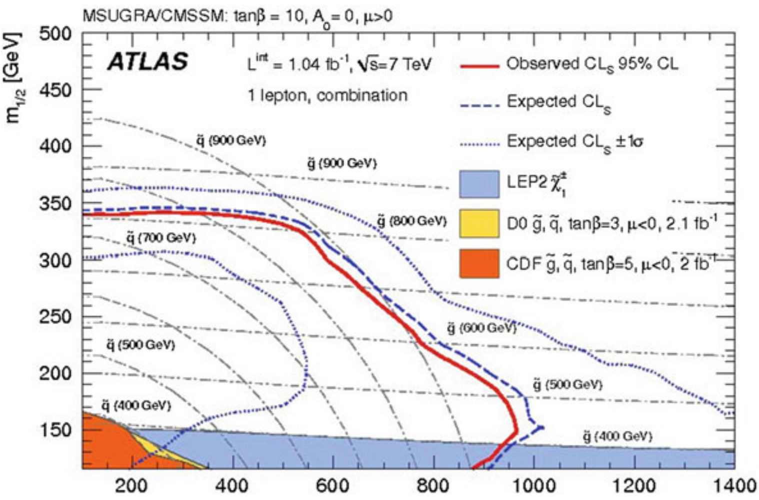


Fig. 12.3 Observed and expected CL_s limits at 95 % C.L., as well as the ±1 sigma variation on the expected limits, in the combined 1-lepton channel for an integrated luminosity of about 1 fb^{-1}

extracting the amount in the Signal Regions (SR) by means of transfer factors.[4] With no excess, it has been possible to set limits on the SUSY particle masses and subsequently to interprete the results within different theoretical models. Gluino and squark masses below 1000 GeV have been excluded at the 95 % Confidence Level (C.L.) within the main models. The most important uncertainties taken into account in this analysis are Jet Energy Scale (JES), Jet Energy Resolution (JER), pileup effect, luminosity, Monte Carlo (MC) statistics and the uncertainty on cross sections.

12.3 1-Lepton Final State

In case a squark decays via chargino and a quark (Fig. 12.2), then the chargino may produce a high-p_T lepton in the final state. Scenarios with large E^T_{miss}, jets and one electron or muon have been studied by ATLAS, but no excess has been found with respect to the SM predictions.

The main background processes for this final state are W+jets, top pair production with a semi-leptonic decay and QCD processes with a jet misidentified as a lepton. To measure them, the number of expected background events in the defined SR has been extracted with data-driven techniques from different CRs. Good agreement was also observed between predicted and observed events in every CR. Systematic uncertainties for this analysis are mainly due to theoretical uncertainties such as the

[4] A Control (Signal) Region is defined with loose (tight) cuts to enhance the presence of background (signal).

prediction of cross section, scale factors and PDFs (20–30 %), MC statistics (15 %), JES and JER (1–10 %), pileup (1–10 %) and luminosity (3.7 %).
Final results led to exclude masses of gluinos and squarks up to 875 GeV at the 95 % C.L. (Fig. 12.3).

12.4 2-Leptons Final State

The final state with two high-p_T, well-isolated leptons is a very promising venue to discover SUSY at the LHC. Two leptons can emerge either from the chargino decay into lepton-neutrino-neutralino in both legs or from a single chargino decay into two leptons plus a lighter chargino (Fig. 12.2).

The SRs in this analysis have been optimized with studies on specific models and include final states with opposite sign leptons (OS), same sign (SS) and a flavour subtraction of events with same flavour and different flavour leptons. The major source of background for the OS SR comes from top pair production with a di-leptonic decay. SS leptons events have smaller SM background, but they suffer from low statistics. The estimate of all these processes has been handled with data- driven or semi data-driven techniques. In particular, the QCD misidentified leptons have been measured with a data-driven technique called Matrix Method. The total systematic uncertainty on the number of expected events for this channel ranges from 25 to 70 %, depending on the cuts in the SR.

As for the other channels, no excess over SM background has been observed and limits on the masses and cross sections have been set. Charginos masses up to 200 GeV are excluded at 95 % C. L.

12.5 Conclusions

A wide range of SUSY signatures has been investigated by ATLAS with 1 fb^{-1} of proton-proton collisions data. No hints for supersymmetry was observed in all channels and limits on the supersymmetric particles masses were set depending on the theoretical model.

At 95 % C.L. squarks and gluinos masses up to 1000 GeV are excluded in most principal models. For the year 2012, the ATLAS Collaboration is moving to study all the remaining range of SUSY signatures as well as to improve the existing analyses using the full 2011 statistics.

The parameter space where SUSY particles are hiding is being continuously reduced.

References

1. The ATLAS Coll (2011) Search for squarks and gluinos using final states with jets and missing transverse momentum with the ATLAS detector in sqrt(s) = 7 TeV proton-proton collisions, ArXiV:1109.572.
2. The ATLAS Coll (2011) Search for supersymmetry in pp collisions at sqrt(s) = 7 TeV in final states with missing transverse momentum, b-jets and no leptons with the ATLAS detector, ATLAS-CONF-2011-098.
3. The ATLAS Coll (2011) Search for supersymmetry in final states with jets, missing transverse momentum and one isolated lepton in sqrt(s) = 7 TeV pp collisions using 1 fb-1 of ATLAS data, ArXiV:1109.6606.
4. The ATLAS Coll (2011) Searches for supersymmetry with the ATLAS detector using final states with two leptons and missing transverse momentum in sqrt(s) = 7 TeV proton-proton collisions, ArXiV:1110.6189.
5. The ATLAS Experiment. http://www.atlas.ch. Accessed Oct 2011

Part IV
Gravitation and Cosmology

Chapter 13
Dark Energy from Curvature and Ordinary Matter Fitting Ehlers-Pirani-Schild: Foundational Hypothesis

M. De Laurentis, Lorenzo Fatibene and Mauro Francaviglia

Abstract We discuss in a critical way the physical foundations of geometric structure of relativistic theories of gravity by the so-called Ehlers-Pirani-Schild formalism. This approach provides a natural interpretation of the observables showing how relate them to General Relativity and to a large class of Extended Theories of Gravity. In particular we show that, in such a formalism, geodesic and causal structures of space-time can be safely disentangled allowing a correct analysis in view of observations and experiment. As specific case, we take into account the case of $f(R)$ gravity.

13.1 Introduction

Einstein General Relativity (GR) is a self-consistent theory that dynamically describes Space, Time and Matter under the same standard. The result is a deep and beautiful scheme that, starting from some first principles, is capable of explaining a huge number of gravitational phenomena, ranging from laboratory up to cosmological scales. Its predictions are well tested at Solar System scales and give rise to a comprehensive cosmological model that agrees with the Standard Model of particles, with the

---

PACS numbers: Keywords: g.

M. De Laurentis (✉)
Dipartimento di Scienze Fisiche, Universit'a di Napoli "Federico II", Compl. Univ. di Monte S. Angelo, Edificio G, Via Cinthia, I-80126 Napoli, Italy

M. De Laurentis
INFN Sez. di Napoli, Compl. Univ. di Monte S. Angelo, Edificio G, Via Cinthia, 80126 Naples, Italy

L. Fatibene · M. Francaviglia
Dipartimento di Matematica, Universit di Torino, Turin, Italy

L. Fatibene · M. Francaviglia
INFN Sez. di Torino, Via Carlo Alberto 10, 10123 Turin, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_13,

recession of galaxies, with the cosmic nucleosynthesis and so on. Despite these good results, the recent advent of the so-called Precision Cosmology tests from astrophysics (rotation curves of galaxies) and possible some tests coming from the Solar System outskirts (e.g. the Pioneer anomaly) entail that the self-consistent scheme of GR seems to disagree with an increasingly high number of observational data, as e.g. those coming from IA-type Supernovae, used as standard candles, large scale structure ranging from galaxies up to superclusters. Furthermore, being not renormalizable, GR seems to fail to be quantized in any classical way (see [4]). In other words, it seems, from ultraviolet up to infrared scales, that GR is not and cannot be the definitive theory of Gravitation heven if it successfully addresses a wide range of phenomena.

Many attempts have been therefore made both to recover the validity of GR at all scales, on one hand, and to produce theories that suitably generalize Einsteins one, on the other hand. In order to interpret a large number of recent observational data inside the paradigm of GR, the introduction of DarkMatter (DM) and Dark Energy (DE) seemed to be necessary: the price of preserving the *simplicity* of the Hilbert Lagrangian has been, however, the introduction of rather odd-behaving physical entities which, up to now, have not been revealed by any experiment at fundamental scales. In other words, we are observing the large scale effects of missing matter (DM) and the accelerating behaviour of the Hubble flow (DE) but no final evidence of these ingredients exists, if we want to deal with them as standard quantum particles or fields. However, from an observational point of view, considering GR + cosmological constant + DM gives an extremely good snapshot of the currently observed Universe. The problem is that dynamics of previous epochs cannot be reconstructed and addressed in a self-consistent way starting from the present status of observations. Furthermore, it seems that the type of DM to be considered strictly depends on the size of selfgravitating structures (e.g. the dynamical behavior of DM in small galaxies, in giant galaxies and in galaxy clusters is completely different). So, besides the issue to find out DM and DE at fundamental scales, it seems hard to find out a general dynamics involving such components working at all cosmic epochs and at any astrophysical size. With these considerations in mind, one can wonder if extending gravity sector could be a more economic and useful approach which does not involve too much exotic ingredients but retains all the good results achieved by GR (for a review, see e.g.[5–8]). In this paper we address some of the recent issues concerning the geometrical structure of "physically reasonable" gravitational theories, starting from the fundamental work of Elehers-Pirani-Shild [9–11] about the geometric and physical foundations of relativistic theories of gravitation and revisiting them, *á la Palatini*, in view of applications to the new challenges discussed above [12, 13].The outline of the paper is as follows. In Sect. 13.2 we introduce the EPS framework. Section 13.3 is devoted to the EPS formalism in GR while Sect. 13.4 is a critical discussion of such an approach. In Sect. 13.5, we discuss EPS from the point of view of Extended Theories of Gravity (ETG). In particular, the straightforward extension of GR, $f(R)$-gravity, is taken into account. Sections 13.6 and 13.7 are devoted to discussion and conclusions. A new paradigm for gravitational theories is proposed assuming the EPS paradigm.

13.2 Ehlers-Pirani-Schild-Theory

We first summarize Ehlers-Pirani-Schild (EPS) analysis of the mathematical structures that lie at the basis of all "reasonable" relativistic gravitational theories [9–11]. In early 70s EPS started from a set of well motivated physical property of light rays matter in a relativistic framework to derive the geometrical structure of space-time from potentially observable objects. This is particularly suitable to discuss which geometric structure is observable and which are conventionals. In this way it provides stronger physical motivation and understanding not only of space-time geometry as such, but also in comparison with more general geometries (as candidates for mathematically modeling physical space-time). EPS specifically highlighted the potential role of space-time models based on Weyl geometry. Supplying this new axiomatic characterization of the otherwise mathematically familiar space-time geometry structure, EPS also brings relevant new insight even from a strictly matical (geometrical) standpoint. Einstein's GR usesses advanced mathematical ideas. Things like 4-dimensional curved spacetime are not easy to grasp. Even if one masters the math behind it, the essential physical meaning and content is not obvious. EPS is one of a series of attempts to clarify the physics behind the math. Unfortunately and unavoidably, getting there requires even more abstract math. At first sight, this seems selfdefeating; however, some of these mathematical ideas are chosen so as to be closer to 'operational' physical interpretation, representing more elementary physical observation, measurement and construction. In the upshot, EPS ends up with Lorentzian metric ($L4$), rather then of accepting it as starting point: the idea is to rebuild $L4$ from scratch, using only bricks with intuitively clear physical meaning to the extent possible, and at the cost of some extra math. For example as far as metric structure $L4$ is concerned EPS clearly showed that what is physically well defined is a conformal structure (the class of all a conformally equivalent Lorentzian metrics such a representative g a specific Lorentzian metric) can be singled out only by convention of an observer. As is typical in axiomatic reconstructions like EPS, one exploits the benefit of hindsight, as the intended result (in this case: $L4$ spacetime of General Relativity) is already known. So this in no way detracts from Einstein's original feat, on the contrary. The scope of EPS is limited to the kinematics of space-time itself; the problem of any possible axiomatic derivation or reconstruction of Einstein field equations (that is dynamics) governing matter and gravity within such a space-time model, is left open.

> "The approach shows how quantitative measures of time, angle and distance, and a procedure of parallel displacement... can be obtained constructively from 'geomtry free' assumptions about light-rays and freely falling particles; pseudo-Riemannian (or Weylian) geometry is recognized even more clearly than before as the appropriate language for a generalized kinematics which allows for the unavoidable and ever-present 'distortions' called gravitational fields." (Ehlers)

13.2.1 A. Outline of Ehelers-Pirani-Shild Construction

With the above considerations in mind, let us outline the main points of EPS conceptual construction. The construction of EPS space-time proceeds in steps as sketched below, each one enriching the axiomatic content of the underlying set of events. Roughly, the underlying idea is the following. From differential geometry, one knows that the geodesics determine 'their affine connection' (assuming torsion to be zero, for instance) and hence a corresponding metric. Now, in contrast to the metric itself, these geodesics do possess an immediate physical interpretation (as light ray worldlines for null geodesics or particle world lines for timelike ones). So in very general terms, one tries to reconstruct the sought after metric from known geodesics that fulfill certain qualitative criteria (postulates), which are themselves physically meaningful and plausible.

- **Particles and light rays in event space.**
  EPS adopts a set M of events (to become the space-time manifold) as its backdrop. On this, a set of particles p and a set of light rays l are assumed given. Each particle and each light ray are identified with their 'world line' of events.
- **Smooth radar coordinates for events**
  As subsets of the space of events, particle and light ray world lines are taken to be smooth one dimensional manifolds. A permissible local coordinate represents time as measured by a (possibly irregular) local clock. Light ray messages between particles p and q smoothly relate their private time parameters, the timing of echoes received back by p also relate smoothly to that of the message flashes it sent out to q to begin with. Using 'radar soundings' in this way, pairs of 'observer' particles set out to map surrounding events by assigning 2 time values each, or a total of 4 coordinates each. Postulating that this process may cover the entire event set, the events form a smooth 4-dimensional space (manifold).
- **Light propagation ensures local validity of pointwise causality**
  At each point of space-time (event), the propagation of light determines an infinitesimal null cone, amounting to a conformal structure C of Lorentzian signature. This assertion is stated operationally, This assertion is stated operationally, demanding that one may (topologically) distinguish between C-time-like, space-like and null vectors, directions and curves at an event. Null curves lying on a null hypersurface are singled out as null geodesics.
- **Free falling particles encode influence of gravity on particle motion** Among the timelike curves, the free-falling particles form a preferred family of wordlines. Imposing a generalized law of inertia provides a projective structure, with free-fall world lines as its (C-time-like) geodesics.
- **Free fall implicitly define a projective structure** P.
  They in turns determine, by a canonical gauge fixing, a preferred connection space-time.
- **Light and particle motion agree** Then one can define two compatible conformal and projective structures on space-time. The choice of representatives is a conventional gauge fixings. The conventional nature of metrics and connections

is important to be noticed in view of which quantities are to be considered physically sound. In particular, one can choose canonically a standard representative of projective structure imposing

$$\nabla^{(\Gamma)}_{\mu} g_{\alpha\beta} = 2V_{\mu} g_{\alpha\beta} \tag{13.1}$$

for some covector *V*. Then there is a canonical connection $\Gamma^{\alpha}_{\beta\eta}$ which, of course, depends on extra degrees of freedom depending on *A*. The triple (M; C;Γ) is called a *Weyl-geometry*. It is called metric if there exists a representative $\Gamma \in \mathcal{C}$ such that $\Gamma = \{g\}$ coincides with the Christoffel symbols of the metric *g*. In this case, the metric *g* describes light rays and particles free fall, as it is assumed in standard GR. However, in general one needs two different (still compatible) structures to describe light rays and matter free fall. Let us stress once again that there is no reason at this stage to assume that the Weyl-geometry obtained on space-time is metric. A Weyl space possesses a unique affine structure *A*: *A* geodesics are *P* and *A* parallel displacement preserves *C* nullity. In a Weyl space, one may construct a "proper time" arc length (up to linear transformation) along non-null curves by purely geometrical means (i.e. using light rays reflected from particles only, so without any need for atomic clocks). In technical terms, one employs affine parallel displacement, and congruence in the tangent space, as defined by *C*. This 'geodesic' clock is known as the Marke-Wheeler clocks [14].

13.2.2 B. Hypotesis

In summary EPS analysis is based on a number of assumptions: It physically distinguishes the Principle of Equivalence from the Principle of Causality and investigates the need of measuring and describing Space-time structure through light rays. The need of measuring in Space-time and using light rays requires that Space-time carries a (Lorentzian) metric while the Principle of Equivalence and interaction with matter ("Free Fall" under gravitational pull) requires that Space-time carries also a (Linear or Affine) Connection. The Connection, an object that can be reduced to be zero at each single point, is the potential of the gravitational field. The Metric determines causality and photon propagation. According to EPS analysis, in order for a Gravitational theory being physically reasonable, compatibility conditions should exist between the Metric and the Connection. The Connection defines a family of autoparallel lines (also called improperly geodesics). They establish the free fall of pointlike (in principle massive) "test particles". The Metric defines light cones and a family of geodesics. Null geodesics of the Metric are paths of light rays (photons). The family of autoparallel lines of the Connection determine an equivalence class of "Projectively Equivalent Connections". Along them free fall is the same, only proper time changes. The light cones of the Metric define an equivalence class of Conformally Equivalent Metrics. Along them units and measuring devices change point by point, but light rays and photon trajectories are the same. The required compatibility condition amounts to pretend that the two families of autoparallel lines of the (projective equivalence class of) Connections and the family of null geodesics of the conformal equivalence class of Metrics are in a precise relation: each null geodesic of the Metric has to be one of the autoparallel lines of the Connection At this point of their fresh analysis all Foundational Axioms have been satisfied. In order to recover General Relativity as the Unique Relativistic Theory of Gravitation Elhers, Pirani and Schild make some further axiomatic hypotheses:

- **Speed of time does not depend on path**
  A final physical assumption (expressed mathematically as an axiom) ensures the existence of a Lorentzian metric, which determines both light cones and free fall. "Equally spaced clock ticks" along one particle world line are transported to a nearby particle by Einstein simultaneity. Imposing that this must generate (approximately) equidistant ticks also for the second particle and applying the equation of geodesic deviation for the curvature tensor given by A implies (through the vanishing of the Weyl 'track curvature') the existence of a single Lorentzian metric compatible to both C with A.
  This finally 'reduces' Weyl space to $L4$. Requiring in this way that 'time runs equally fast along all paths' amounts to denying the existence of a 'second clock effect'. Indeed, in (Lorentzian) GR, only the 'time interval' between 2 events is path dependent (i.e. the 'first clock effect'); not the 'speed' of time.
  "Metricity Axiom": a single Metric is chosen in the conformal class and the Connection is chosen while be the Levi-Civita Connection of this Metric.
  With this above hypothesis the compatibility conditions are met. Notice, however, that this just amounts to say that the gravitational theory is of "purely metric" nature. To recover GR as the unique Relativistic Theory of Gravitation one has in fact to make a further assumption.

13.3 Elehers-Pirani-Schild and General Relativity

In order to recover General Relativity as the Unique Relativistic Theory of Gravitation one has in fact to make the following further axiomatic hypotheses:

"Lagrangian Axiom": the Lagrangian that governs gravitational field equations (in absence of Matter) is the Scalar Curvature.

The "Metricity Axiom" has in fact no real physical grounds. According to EPS (and to physical needs) a Metric has to exist to define rods and clocks, but there is no need to pretend from the very beginning that it defines also the gravitational potential, i.e. the Connection. Assuming that the Metricity Axiom holds is just a "matter of taste" and in a sense it corresponds to have a great mathematical simplification. From the viewpoint of Lagrangian Mechanics it is a purely kinematical restriction imposed a priori on Dynamics. Physically spacing, it is much better not to impose a priori purely kinematical restriction on Dynamics. Physics requires that possible restrictions should be obtained from dynamics rather than imposed a priori as a constrains. This point was perfectly clear to Albert Einstein when, in 1923, he tried to establish a more general setting for Gravity (and Electromagnetism) by assuming a priori that both a Metric and a Connection must be chosen, from the beginning, as dynamical variables. So-called "Palatini formalism" was born. Also the "Lagrangian Axiom" had in fact no real physical grounds. Once it is clear which are the variables that have to enter dynamics, the choice of a Lagrangian for them is again a "matter of taste" or it should be at least determined on the basis of Phe- nomenology, in order to fit observational data. When Hilbert, in 1916, in the purely metric framework (the only one that was available before 1919 and Levi-Civitas work on Linear Connections) assumed the Lagrangian to be the Scalar Curvature of the Metric this was, in a sense, an obliged choice. Dictated by "simplicity". The choice of the Hilbert-Einstein Lagrangian *R(g)*, made in 1916, was not only the "simplest one". It satisfied the will to obtain second-order field equations suitably generalizing Newtons law, fitting all astronomical predictions, satisfying conservation of matter and being compatible with Maxwell.

$$G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta} = \kappa T \alpha\beta, \tag{13.2}$$

where $G_{\alpha\beta}$ is the Einstein tensor, a combination of curvature invariants derived from Bianchi's identities, $T\alpha\beta$ is the stress-energy momentum tensor and $\kappa = 16\pi G$ is the gravitational coupling constant.

In the new framework introduced by assuming both metric and connection among the variables, Einstein decided to take into account again, in 1923, the Lagrangian to be the scalar curvature (of metric and connection), again for the sake of simplicity. At that time there were very few observations fine enough to be used as tests and all of them agree with purely metric predictions. Thus the first test for an extended theory was to reproduce standard GR in purely metric formalism. When Einstein, in 1923, in the new framework he introduced by assuming both a Metric and a Connection among the variables he decided to assume again the Lagrangian to be the Scalar Curvature of the Metric and the Connection, again for the sake of simplicity. In this new framework and with the Linear Lagrangian $R(g, \Gamma)$ he proved that no realy new Physics comes on stage. Field equations impose in fact, a posteriori, hat the Connection is nothing but the Levi-Civita Connection of the Metric, so that GR is eventually recovered. Einstein did not investigate, however, what happens when Matter is coupled to the Linear Lagrangian $R(g, \Gamma)$. In this case just a few slight changes are necessary if Matter couples with the Metric *g* but great difficulties arise if Matter couples with the Connection Γ (as it should). Around the sixties a number of mathematical papers were written about possible generalizations of Einsteins Theory by reverting to Non-Linear Lagrangians, more complicated than R(g). These Higher Order Theories remained just as a mathematical game for long time. Renewed interest towards Non-Linear Lagrangians more complicated than R(g) (Higher Order Theories) was lately determined by new phenomenology, such as: Inflation, Acceleration in the Expansion, Dark Matter, Quantum Gravity, Low Energy Limit of String Models [15, 16].

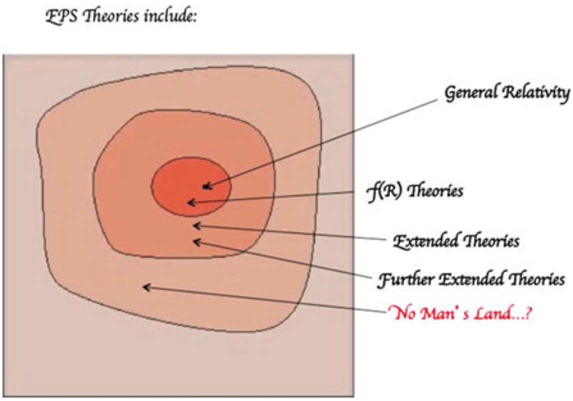


Fig. 13.1 The eps theories

13.4 Elehers-pirani-schild revised formalism

The analysis of EPS—concerning the mathematical and physical foundations of relativistic theories of gravitation and the compatibility between (conformal classes of) metrics and (projective classes of) connections—is worth of being revisited. EPS have shown that the family of gravitational theories that satisfy all of their Axioms (with the exception of the "Metricity Axiom" and the "Lagrangian Axiom") includes many (but not all) of the currently investigated frameworks for (relativistic) gravitation. First of all, it suggest that the correct and most general framework for dealing with gravity is the Palatini formalism, since it is based on the physical and mathematical distinction between the Principle of Equivalence and the Principle of Causality that for obvious reasons are mathematically and physically distinct. They imply the necessity of introducing a priori distinct and separate structures to full fill them, even if compatibility is required a posteriori on the mathematical and physical structures they induces on space-time. Within this formalism, the most general class of theories that care be considered without renauncing to the physical requirements point out by EPS analysis, is the family of so-called "Further Exteneded Theories of Gravity" that has been explicitly introduced in [17, 18]. This class includes all gravitational theories in which the gravitational Lagrangian depend on g and the (Ricci) curvature of the connection, the matter Lagrangian interacts allows in principle interaction of matter with both g and Γ and, a posteriori or a priori, field equations imply EPS compatibility. Of course one is free to work in more general frameworks for gravitation, but in such a case one has to remind that at list one EPS requirements will fail.

Our choice will be more restrictive and will be therefore based on three assumptions:

1. Assume Palatini—EPS framework and accept the view that in Palatini formalism the gravitational field is encoded in to the dynamical connection (i.e. free fall) while the dynamical metric has more to do with measures, rods, clocks and causality. We accept moreover that dynamics, and in particular the interaction

with matter, will determine a posteriori the relations between the metric and the connection, so to satisfy EPS compatibility requirements.
2. Assume that the Lagrangian is a possibly non-linear function of curvature of g and ;
3. Assume that the Lagrangian is "simple"—the simplest choice is of course $f(R(g;\Gamma)$ but it is not the only simple case.

13.5 The Ehelers-Pirani-Schild approach for f(r)-Gravity

With this choice one can show that if no Matter is present, purely gravitational equations entail that the Connection Γ entering dynamics is still the Levi-Civita Connection of the Metric g while a (quantized) Cosmological Constant enters the game and somehow determines the asymptotic freedom for Gravity. One should remark that if Matter is present and couples only to the Metric g things change if and only if the trace t of the Stress Tensor is different from zero. In particular, thence, nothing changes when Electromagnetism couples, so that the light cone structure and photons are not affected when passing to Palatini framework. It is however known that both in purely metric formalism (higher order gravity) and in the Palatini approach (first order gravity) coupling with matter generates relativistic effects that are not present in vacuum. This is particularly evident when one relies on non-linear Lagrangians of the type $f(R)$. In $f(R)$ gravity in the Palatini approach in presence of Matter coupled only with g field equations still imply that the Connection is metric, but now it is the Levi-Civita Connection of a new Metric h, conformally related with the Metric g given in the Lagrangian.

Being this the core point of our discussion, we want to derive in details the field equations of $f(R)$ gravity in Palatini formalism and then perform the EPS analysis in this framework.

Let us first consider on M metric field g, a torsionless connection Γ and a generic tensor density A of rank 1 and weight -1. The covariant derivative of then defined as

$$\overset{\Gamma}{\nabla}_{\mu}A_{\nu} = d_{\mu}A_{\nu} - \Gamma^{\lambda}_{\nu\mu}A_{\lambda} + \Gamma^{\lambda}_{\lambda\mu}A_{\nu}, \tag{13.3}$$

Accordingly, we have

$$\begin{aligned}\overset{\Gamma}{\nabla}_{\mu A_{\nu}} &= d_{(\mu A_{\nu})} - (\Gamma^{\epsilon}_{\nu\mu} - \delta_{(\nu\Gamma^{\lambda}_{\mu})\lambda})A_{\epsilon} \\ &= d_{(\mu A_{\nu})} - u^{\epsilon}_{\mu\nu}A_{\epsilon},\end{aligned} \tag{13.4}$$

where we set $u^{\epsilon}_{\mu\nu} := \Gamma^{\epsilon}_{\mu\nu} - \delta^{\epsilon}_{(\mu}\Gamma^{\lambda}_{\nu)\lambda}$.

Let us consider the following Lagrangian (density)

$$\mathcal{L} = \frac{1}{\kappa}\sqrt{g} f(R) + g g^{\mu\nu} \overset{\Gamma}{\nabla}_{\mu} A_{\nu}, \tag{13.5}$$

where, $g = |det(g_{\nu\omega})$. $R = g^{\mu\nu} R_{\nu\omega}(\Gamma)$ is the scalar curvature of (g; Γ), and f(R) is a generic (analytic) function. By variation of this Lagrangian and usual covariant integration by parts one obtains

$$\begin{aligned}
\delta\mathcal{L} = & \frac{\sqrt{g}}{\kappa}\left(f'(R)R_{\alpha\beta} - \frac{1}{2}f(R)g_{\alpha\beta} - \kappa T_{\alpha\beta}\right)\delta g^{\alpha\beta} \\
& - g g^{\alpha\beta} A_{\lambda}\delta u^{\lambda}_{\alpha\beta} + \frac{\sqrt{g}}{\kappa} g^{\alpha\beta} f\prime(R) \overset{\Gamma}{\nabla}_{\lambda}\delta u^{\lambda}_{\alpha\beta} \\
& + g g^{\mu\nu} \overset{\Gamma}{\nabla}_{\mu}\delta A_{\nu} \\
= & \frac{\sqrt{g}}{\kappa}\left(f'(R)R_{(\alpha\beta)} - \frac{1}{2}f(R)g_{\alpha\beta} - \kappa T\alpha\beta\right)\delta g^{\alpha\beta} \\
& - \frac{1}{\kappa}\left(\overset{\Gamma}{\nabla}_{\lambda}(\sqrt{g}g^{\alpha\beta}f'(R)) + \kappa g g^{\alpha\beta}A_{\lambda}\right)\delta u^{\lambda}_{\alpha\beta} \\
& - \overset{\Gamma}{\nabla}_{\mu}(g g^{\mu\nu})\delta A_{\nu} \\
& + \overset{\Gamma}{\nabla}_{\lambda}\left(\frac{\sqrt{g}}{\kappa}g^{\alpha\beta}f'(R)\delta u^{\lambda}_{\alpha\beta} + g g^{\lambda\nu}\delta A_{\nu}\right),
\end{aligned} \tag{13.6}$$

where we used the well-known identity $^{\delta R}(\alpha\beta) = \overset{\Gamma}{\nabla}\lambda\delta u^{\lambda}_{\alpha\beta}$ and we set for the energy-momentum tensor $T_{\alpha\beta} = \sqrt{g}\left(g\alpha\beta g^{\mu\nu}\overset{\Gamma}{\nabla}_{\mu}A_{\nu} - \overset{\Gamma}{\nabla}_{(\alpha A_{\beta})}\right)$. Field equations are

$$\begin{cases} f'(R)R_{\alpha\beta} - \frac{1}{2}f(R)g_{\alpha\beta} = \kappa T_{\alpha\beta}, \\ \overset{\Gamma}{\nabla}_{\lambda}(\sqrt{g}g^{\alpha\beta}f'(R)) = \alpha\lambda\sqrt{g}g^{\alpha\beta}f'(R), \\ \overset{\Gamma}{\nabla}_{\mu}(g g^{\mu\nu}) = 0, \end{cases} \tag{13.7}$$

where we set $\alpha_{\lambda} := -\kappa\frac{\sqrt{g}}{f'R}A_{\lambda}$. Notice that the third equation (that is the matter field equation) is not enough to fix the connection due to the contraction. Notice also that these are more general than field equations of standard $f(R)$ theories due to the rhs of the second equation (that is originated by the coupling between the matter field A and the connection Γ). Nevertheless one can analyze these field equations along the same lines used in $f(R)$ theories. Let us thence define a metric $h\mu\nu = f'(R)g_{\mu\nu}$ and rewrite the second equation as

$$\overset{\Gamma}{\nabla}_{\lambda}\left(\sqrt{h}h^{\alpha\beta}\right) = \alpha_{\lambda}\sqrt{h}h^{\alpha\beta}. \tag{13.8}$$

According to the analysis of EPS-compatibility done in [17, 18] this fixes the connection as

$$\Gamma^{\alpha}_{\beta\mu} := \{h\}^{\alpha}_{\beta\mu} - \frac{\kappa}{2f'(R)}\left(h^{\alpha\epsilon}h_{\beta\mu} - 2\delta^{\alpha}_{\beta\delta^{\epsilon}_{\mu}}\right)\alpha_{\epsilon}, \tag{13.9}$$

where for notational convenience we introduced the 1-form $\alpha_{\epsilon} := \sqrt{g}A_{\epsilon}$. For later convenience let us notice that we have

$$K^{\alpha}_{\beta\mu} \equiv \Gamma^{\alpha}_{\beta\mu} - \{h\}^{\alpha}_{\beta\mu} = -\frac{\kappa}{2f'(R)}\left(h^{\alpha\epsilon}h_{\beta\mu} - 2\delta^{\alpha}_{(\beta\delta^{\epsilon}_{\mu})}\right)(d, 0) \tag{13.10}$$

Now we can define the tensor $H^{\alpha}_{\beta\mu} := \Gamma^{\alpha}_{\beta\mu} - \{g\}^{\alpha}_{\beta\mu}$ and obtain

$$\begin{aligned} H^{\alpha}_{\beta\mu} &= K^{\alpha}_{\beta\mu} - \frac{1}{2}\left[g^{\alpha\lambda}g\beta\mu - 2\delta^{\alpha}_{(\beta\delta^{\alpha}_{\mu})}\right]\delta_{\lambda}\ \mathrm{In} f'(R) \\ &= -\frac{1}{2f'(R)}\left[g^{\alpha\epsilon}f\beta\mu - 2\delta^{\epsilon}_{(\beta\delta^{\alpha}_{\mu})}\ [[\kappa a_{\epsilon} + \delta_{\epsilon}f'(R)], \right. \end{aligned} \tag{13.11}$$

By substituting into the third field equation we obtain

$$\begin{aligned} &\overset{g}{\nabla}_{\mu}(gg^{\mu\nu}) + g(H^{\mu}_{\lambda\mu}g^{\lambda\nu} + H^{\nu}_{\lambda\mu} - 2H^{\lambda}_{\lambda\mu}g^{\mu\nu}) = 0, \\ &\Rightarrow H^{\nu}_{\lambda\mu}h^{\mu\lambda} - H^{\lambda}_{\lambda\mu}h^{\mu\nu} = 0, \\ &\Rightarrow -\frac{1}{2f'(R)}\left[\left(h^{\nu\epsilon h_{\lambda\mu}} - 2\delta^{\epsilon}_{(\lambda\delta^{\nu}_{\mu})}\right)h^{\mu\epsilon}\right. \\ &\quad \left. + \left(h^{\lambda\epsilon}h_{\lambda\mu} - 2\delta^{\epsilon}_{(\lambda\delta^{\lambda}_{\mu})}\right)h^{\mu\nu}\right](\kappa a_{\epsilon} + \delta_{\epsilon}f'(R)) = 0, \\ &\Rightarrow -\frac{3}{f'(R)}h^{\nu\epsilon}(\kappa a_{\epsilon} + \delta_{\epsilon}f'(R)) = 0, \\ &\Rightarrow a_{\epsilon} = -\frac{1}{\kappa}\delta_{\epsilon}f'(R), \end{aligned} \tag{13.12}$$

where $\overset{g}{\nabla}_{\mu}$ is now the covariant derivative with respect to the metric g. Hence the matter field $A_{\epsilon} = \sqrt{g}a_{\epsilon} = -\frac{\sqrt{g}}{K}\delta_{\epsilon} = f'(R)$ has no dynamics and it is completely deter- mined in terms of the other fields. We can also express the connection as a function of g alone (or, equivalently, of h alone)

$$\Gamma^{\alpha}_{\beta\mu} := \{h\}^{\alpha}_{\beta\mu} + \frac{1}{2}\left(h^{\alpha\epsilon}h_{\beta\mu} - 2\delta^{\alpha}_{(\beta\delta^{\epsilon}_{\mu})}\right)\delta_{\epsilon} \in f'(R) \equiv \{g\}\Gamma^{\alpha}_{\beta\mu} \tag{13.13}$$

This behaviour, which has been introduced by the matter coupling, is quite peculiar; the model resembles in the action an *f (R)* theory but in solution space the connection is directly determined by the original metric rather than by the conformal metric h as in *f (R)* theories. Still the metric g obeys modified Einstein equations. In fact, we have the first field equation which is now depending on g alone, since the matter and the connection have been determined as functions of g. The master equation is obtained as usual by tracing (using $g^{\alpha\beta}$)

$$f'(R)R - 2f(R) = \kappa T \quad \Rightarrow f(R) = \frac{1}{2}(f'(R)R - \kappa T), \tag{13.14}$$

where we set $T := T_{\alpha\beta}g^{\alpha\beta}$ the trace of the stress-energy tensor. By substituting back into the first field equation, $T = -\frac{3}{\kappa}\Box f'(R)$ being , we obtain

$$\begin{aligned} & f'(R)\left[R_{\alpha\beta} - \frac{1}{4}Rg_{\alpha\beta}\right] - \frac{3}{4}\Box f'(R)g_{\alpha\beta} \\ & = \nabla_\alpha\nabla_\beta f'(R) - \Box f'(R)g_{\alpha\beta}, \\ & \Rightarrow R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} = \\ & \frac{1}{f'(R)}\left[\nabla_\alpha\nabla_\beta f'(R) - \frac{1}{4}(\Box f'(R) + f'R)g_{\alpha\beta}\right], \end{aligned} \tag{13.15}$$

where now the curvature and covariant derivatives refer to g. These are exactly the field equations obtained in the corresponding purely-metric $f(R)$ theory [5–8].

Hence we have that, regardless of the function *f (R)*, when there is no matter field other than the field A all these models behave exactly as metric *f (R)* theories. The conformal factor is *(R)* and *R* can be calculated in terms of *T*, i.e. $\phi(T) = {}^!(R(T))$. Field equations then imply that Einstein equations hold for the new metric $\hat{g}$ (corresponding to the above h), with a suitably modified stress-energy tensor that takes into account extra effects due to the conformal factor. The previous Einstein equations are recovered

$$\hat{R}_{\alpha\beta} - \frac{1}{2}\hat{R}\hat{g}_{\alpha\beta} = \kappa\hat{T}_{\alpha\beta} \tag{13.16}$$

with

$$\hat{T}_{\alpha\beta} = \frac{1}{f'}\left[T_{\alpha\beta} + T^{(grav)}_{\alpha\beta}\right], \tag{13.17}$$

where the first term on the *rhs* is due to a standard matter term- Clearly Eq. 13.17 means that the extra degrees of freedom coming from $f(R)$ gravity can be managed as a further contribution to the stress-energy tensor and the above observational shortcomings, related to GR (e.g. DM and DE), can be, in principle, solved in a geometrical way.

Which are the physical implications from the EPS formalism point of view?

1. being g and $\hat{g}$ conformally related photon propagation does not change;
2. Einstein equations hold for the new metric $\hat{g}$ with extra stress-energy tensor directly generated by "ordinary" matter *T*;
3. rods and clocks change pointwise, by a factor depending on *T*.

In summary, EPS formalism works also for ETG and further information can be always enclosed in a suitable definition of stress-energy tensor.

13.6 A New Paradigm for Gravity

The coupling of a non-linear gravitational Lagrangian $f(R)$ with matter Lagrangians, depending on the metric g in an arbitrary way or even on the connection Γ in a peculiar way (dictated by EPS compatibility), generate a set of modified Einstein equations in which the following effects are easily recognizable:

A new metric $\hat{g}$ conformally related to the original metric g arises. The conformal factor is a computable function of curvature and, through functional inversion, of the trace of the stress tensor that corresponds to the "ordinary" matter distribution (including possible DM and DE effects). In the Palatini approach, the new metric generates the connection Γ as its Levi-Civita connection, so that it describes the free fall of ordinary matter. This new metric induces, in fact, a change of rulers and clocks that affects measurements and conservation laws, while the original g is directly related to light propagation. Due to conformal equivalence, light propagates on the same null geodesics of both g and $\hat{g}$, although clock rates are different in presence of matter.

The net effect of non-linearity and of (non trivial) interaction with matter resides in a change of the stress tensor that couples to the Einstein tensor of $\hat{g}$; a change that induces additions to the previously existing one (directly generated from the matter Lagrangian as discussed above).

This new stress-energy tensor defines conservation laws that are fully covariant with respect to the Einstein frame of $\hat{g}$. Furthermore, it contains an additional term, that can be interpreted under the form of a "space-time varying cosmological constant" $\Lambda(\mathrm{x})$ in turn determined by distribution of ordinary (and Dark) Matter so that the residual amount could be interpreted as a net curvature effect (DE) due to the change of rules and clocks induced by EPS compatibility [5–8]. In other words, the observational effects of such a dynamics are the clustering of astrophysical structures (DM) and the revealed cosmic speed up (DE).

13.7 Conclusions and Remarks

To conclude, we can say that very likely Einstein today, after the new phenomenological evidences would much probably come back onto his own steps and accept, as he always did, that models are not eternal and should be dictated by phenomenology rather than by preestablished rules and prejudices. Why should we insist on pre-judicial rules that impose metricity *a priori* (and metricity with respect to a given metric!) and insist on the choice of the "simplest" Hilbert-Lagrangian, when cosmology, quantum Issues and strings suggest instead to us to strictly follow the beautiful analysis of EPS, and work at least *a priori*, in the extended framework of Palatini-EPS formalism and in a much larger class of Lagrangians?

Moreover, let us remark that working in the extended setting suggested by the Palatini-EPS framework requires to reconsider all the machinery and settings of the observational paradigms and protocols have to be carefully analyzed to disentangle purely metrical effects from effects that measure the interaction with free-fall (and therefore with the connection) that in purely metric formalism GR are necessarily mixed up and entangled by the a priori requirement that free-fall is also driven by the metric.

References

1. Riess AG et al (1998) Astron J 116:1009
2. Riess AG et al (2004) ApJ 607:665
3. Cole S et al (2005) MNRAS 362:505
4. Utiyama R, DeWitt BS (1962) Renormalization of a classical gravitational field interacting with quantized matter fields. J Math Phys 3:608
5. Capozziello S, Francaviglia M (2008) Gen Rel Grav 40:357
6. Capozziello S, De Laurentis M, Faraoni V (2009) Open Astr J 2:1874
7. Nojiri S, Odintsov SD (2011) Phys Rep 505:59–144
8. Sotiriou TP, Faraoni V (2010) Rev Mod Phys 82:451
9. Elehers J, Pirani FAE, Schild A (1972) The geometry of free fall and light propagation.
10. Ehlers J, Kohler E (1977) J Math Phys 18:2014
11. Ehlers J (1973) The nature and structure of spacetime. In: Mehra J (ed) The physicist's conception of nature. D. Rei- del Pub, Dordrecht.
12. Palatini A (1919) Rend Circ Mat Palermo 43:203
13. Ferraris M, Francaviglia M, Reina C (1982) Gen Rel Grav 14:243
14. Misner CW, Thorne KS, Wheeler JA (1973) Gravitation. W.H Feeman & Co, New York
15. Capozziello S, Faraoni V (2010) Beyond Einstein gravity: A survey Of gravitational theories for cosmology and astrophysics. Springer, New York
16. Capozziello S, De Laurentis M (2011) Invariance principles and extended gravity: Theory and probes. Nova Science Publishers, New York
17. Di Mauro M, Fatibene L, Ferraris M, Francaviglia Int M (2010) J Geom Meth Mod Phys 7:887.
18. Fati-bene L, Ferraris M, Francaviglia M, Mercadante S (2010) Int J Geom Meth Mod Phys 7:899

Chapter 14
The Palatini Approach Beyond Einstein's Gravity

Gonzalo J. Olmo

Abstract I review recent results obtained for extensions of general relativity formulated within the Palatini formalism, an approach in which metric and connection are treated as independent geometrical entities. The peculiar dynamics of these theories, governed by second-order equations and having no new degrees of freedom, makes them specially suitable to address certain aspects of quantum gravity phenomenology, construct nonsingular bouncing cosmologies, and explore black hole interiors, which in the Reissner-Nordström case develop a compact core of finite density instead of a point-like singularity.

14.1 Introduction

General relativity (GR) has been confronted with experiments in scales that range from millimeters to astronomical distances, scales in which weak and strong field phenomena can be observed [1]. The theory is so successful in those regimes and scales that it is generally accepted that it should also work at larger and shorter distances, and at weaker and stronger regimes. However, for standard sources of matter and radiation, the theory predicts that the Universe emerged from a singularity and that the fate of sufficiently massive stars is the formation of black holes, which posses a singularity behind their event horizon.

The general perception is that in such extreme scenarios GR should be replaced by some improved description able to avoid the singularities. This, in particular, has motivated the study of different approaches to the quantization of gravity and also numerous phenomenological extensions of GR. Among the former we find the very famous *string theory* [2, 3] and *loop quantum gravity* [4–6]. The phenomenological approaches include theories characterized by higher-order curvature terms

G. J. Olmo
Departamento de Física Teórica and IFIC, Centro Mixto Universidad de Valencia and CSIC. Facultad de Física, Universidad de Valencia, Burjassot 46100, Valencia, Spain

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_14,

and/or higher-order derivatives [7–11], models inspired by higher dimensions and the brane-world scenario [12, 13], scalar-tensor and scalar-vector-tensor theories [14], and many others. From the existing literature, most of these phenomenological approaches are formulated within the so-called *metric approach*, in which the affine connection is defined using the Christoffel symbols of the metric.

A different and, in general, inequivalent approach consists on formulating those theories à la Palatini, i.e., assuming no a priori relation between the metric and the connection [15]. This possibility is supported by the geometrical nature of gravitation, which follows from the Einstein equivalence principle, and by the fact that metric and connection are independent and fundamental geometrical entities. Therefore, in the construction of extended theories of gravity, Ockham's razor suggests that we should give higher priority to metric-affine theories, in which metric and connection are independent, than to purely metric theories, in which compatibility between metric and connection is implicitly imposed somehow arbitrarily by sociological or educational tradition. In Palatini theories, on the contrary, the connection is determined by solving its corresponding field equation, which is obtained from the action according to standard variational methods.

In this talk I present recent results obtained within a particular extension of GR formulated à la Palatini. This model allows to explore the potential effects that a minimum length (such as the Planck length) could have on relativistic field theories [16], produces consistent cosmological models that avoid the big bang singularity by means of a cosmic bounce [17, 18], and modify the internal structure of black holes in such a way that their central singularity is replaced by a compact nucleus that may be nonsingular [19, 20].

14.2 Quantum Gravity Phenomenology. Introduction of a Minimum Length

The combination of special relativity, quantum theory, and gravity suggests that relativistic quantum gravitational effects could arise at length scales of order $l_P = \sqrt{\hbar G/c^3} \sim 10^{-35}$ m. Though this scale is well beyond our current experimental capabilities, its mere existence raises doubts as to how a length, which is not a relativistic invariant, could be consistently introduced in our current field theories to explore the potential phenomenology associated to quantum gravity.

To address this problem, we note that special relativity was built by requiring that the speed of light were an invariant and universal magnitude. To combine the speed of light and the Planck length l_P in a way that preserves the invariant and universal nature of both quantities, we first note that Minkowski space-time allows to interpret the relativity principle in geometrical terms. In this way, though c^2 has the dimensions of a squared velocity it needs not be seen as a privileged 3-velocity. Rather, it can be regarded as a geometrical invariant in a 4D space-time. Analogously, we may see l_P^2 as an invariant with dimensions of length squared in a 4D space-time. Dimensional compatibility with a curvature suggests that $l_P^2 \equiv 1/R_P$ could be introduced in the theory via the gravitational sector. For this reason, we consider the following action

$$S[g_{\mu\nu}, \Gamma^{\alpha}_{\beta\gamma}, \psi] = \frac{\hbar}{16\pi l_P^2} \int d^4x \sqrt{-g} \left[R + l_P^2 (aR^2 + R_{\mu\nu}R^{\mu\nu}) \right] + S_m[g_{\mu\nu}, \psi], \tag{14.1}$$

where $R \equiv g^{\mu\nu}R_{\mu\nu}$, $R_{\mu\nu} \equiv R^{\rho}{}_{\mu\rho\nu}$ is assumed symmetric $R_{\mu\nu} = R_{\nu\mu}$ (for the implications of a non-symmetric piece in the Ricci tensor see, for instance, [21–23]), and $R^{\alpha}{}_{\beta\mu\nu} = \partial_\mu \Gamma^{\alpha}_{\nu\beta} - \partial_\nu \Gamma^{\alpha}_{\mu\beta} + \Gamma^{\alpha}_{\mu\lambda}\Gamma^{\lambda}_{\nu\beta} - \Gamma^{\alpha}_{\nu\lambda}\Gamma^{\lambda}_{\mu\beta}$ represents the components of the Riemann tensor, the field strength of the connection $\Gamma^{\alpha}_{\mu\beta}$. The field equations for metric and connection that follow from the above action are [15]

$$f_R R_{\mu\nu} - \frac{1}{2} f g_{\mu\nu} + 2 f_Q R_{\mu\alpha} R^{\alpha}{}_{\nu} = \kappa^2 T_{\mu\nu} \tag{14.2}$$

$$\nabla_\alpha \left[\sqrt{-g} \left(f_R g^{\beta\gamma} + 2 f_Q R^{\beta\gamma} \right) \right] = 0. \tag{14.3}$$

where $\kappa^2 = 8\pi l_P^2/\hbar$, $f = R + l_P^2(aR^2 + R_{\mu\nu}R^{\mu\nu})$, $f_R \equiv \partial_R f = 1 + 2l_P^2 aR$, and $f_Q \equiv \partial_Q f = l_P^2$. Defining the tensor $P_\mu{}^\nu = R_{\mu\alpha}g^{\alpha\nu}$, (14.2) can be seen as a matrix equation,

$$2 f_Q P_\mu{}^\alpha P_\alpha{}^\nu + f_R P_\mu{}^\nu - \frac{1}{2} f \delta_\mu{}^\nu = \kappa^2 T_\mu{}^\nu, \tag{14.4}$$

which establishes an algebraic relation between the components of $P_\mu{}^\nu$ and those of $T_\mu{}^\nu = T_{\mu\alpha}g^{\alpha\nu}$, i.e., $P_\mu{}^\nu = P_\mu{}^\nu(T_\alpha{}^\beta)$. Once the solution of (14.4) is known, the equation for the independent connection can be solved by means of algebraic manipulations. One then finds that this connection can be written as the Levi-Civita connection of a new auxiliary metric $h_{\mu\nu}$ (see [24] for details) which is related to $g_{\mu\nu}$ through the following non-conformal relation

$$h^{\mu\nu} = \frac{g^{\mu\alpha}\Sigma_\alpha{}^\nu}{\sqrt{\det \Sigma}}, \tag{14.5}$$

where $\Sigma_\alpha{}^\nu = f_R \delta_\alpha^\nu + 2 f_Q P_\alpha{}^\nu$ is a function of $T_\mu{}^\nu$ and, therefore, depends on the local densities of energy and momentum. For instance, if we take the $T_\mu{}^\nu$ of a scalar field with kinetic energy $\chi \equiv g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$ and Lagrangian $\mathcal{L} = \chi + 2V(\phi)$, $h_{\mu\nu}$ and $g_{\mu\nu}$ turn out to be related by

$$g_{\mu\nu} = \frac{1}{\Omega} h_{\mu\nu} + \frac{\Lambda_2}{\Lambda_1 + \chi\Lambda_2} \partial_\mu\phi\partial_\nu\phi \tag{14.6}$$

where $\Omega = [\Lambda_1(\Lambda_1 + \chi\Lambda_2)]^{1/2}$, $\Lambda_1 = \sqrt{2f_Q}\lambda + \frac{f_R}{2}$, $\Lambda_2 = \sqrt{2f_Q}(-\lambda \pm \sqrt{\lambda^2 + \kappa^2\chi})/\chi$, and $\lambda^2 = f/2 + f_R^2/8f_Q - \kappa^2\mathcal{L}/2$.

To better understand the dynamics of our theory, we can use the relation (14.5) to write the field Equation (14.2) in the following compact form

$$R_\mu{}^\nu(h) = \frac{1}{\sqrt{\det \hat{\Sigma}}}(\frac{f}{2}\delta_\mu{}^\nu + \kappa^2 T_\mu{}^\nu), \tag{14.7}$$

where $R_\mu{}^\nu(h) \equiv R_{\mu\alpha}(h)h^{\alpha\nu}$ and $T_\mu{}^\nu \equiv T_{\mu\alpha}g^{\alpha\nu}$. In vacuum ($T_\mu{}^\nu = 0$) this equation boils down exactly to GR with (possibly) an effective cosmological constant (depending on the form of the Lagrangian). This can be seen by rewriting (14.4) in vacuum as

$$2f_Q\left(\hat{P} + \frac{f_R}{4f_Q}\hat{I}\right)^2 = \left(\frac{f_R^2}{8f_Q} + \frac{f}{2}\right)\hat{I}, \tag{14.8}$$

where $\hat{P}$ and $\hat{I}$ denote the matrices $P_\mu{}^\nu$ and $\delta_\mu{}^\nu$, respectively. The physical solution to this equation, which recovers the $f(R)$ theory in the limit $f_Q \to 0$, is of the form

$$P_\mu{}^\nu = -\frac{f_R}{4f_Q}\left(1 - \sqrt{1 + \frac{4f_Q f}{f_R^2}}\right)\delta_\mu{}^\nu \equiv \Lambda(R, Q)\delta_\mu{}^\nu. \tag{14.9}$$

This equation can be used to compute $R_0 \equiv P_\mu{}^\mu|_{vac} = 4\Lambda(R_0, Q_0)$ and $Q_0 = P_\mu{}^\alpha P_\alpha{}^\mu|_{vac} = 4\Lambda(R_0, Q_0)^2$, which lead to the characteristic relation $Q_0 = R_0^2/4$ of de Sitter spacetime. For the quadratic models $f(R, Q) = R + aR^2/R_P + Q/R_P$, for instance, one can also use the trace of (14.2) with $g^{\mu\nu}$ to find that $R_0 = 0$, from which $Q_0 = R_0^2/4 = 0$ follows. For a generic $f(R, Q)$ model, in vacuum one finds that $\Sigma_\mu{}^\nu = a(R_0)\delta_\mu{}^\nu$ and $h_{\mu\nu} = a(R_0)g_{\mu\nu}$, with $a(R_0) = f_R\left(1 + \sqrt{1 + \frac{4f_Q f}{f_R^2}}\right)/2$ evaluated at R_0. Therefore, in vacuum (14.7) can be written as $R_\mu{}^\nu(h) = R_\mu{}^\nu(g) = \Lambda_{eff}\delta_\mu{}^\nu$, with $\Lambda_{eff} = f(R_0, Q_0)/2a(R_0)^2$, which shows that the field equations coincide with those of GR with an effective cosmological constant.

For the particular model (14.1) with $a = -1/2$ coupled to a scalar field, the low energy-density limit $|\mathcal{L}/\rho_P| \ll 1$ (where $\rho_P \equiv c^5/8\pi\hbar G^2 \sim 10^{94}$ g/cm^3 is the Planck matter density) leads to

$$R_{\mu\nu}(h) \approx \kappa^2\left(\partial_\mu\phi\partial_\nu\phi + \frac{V}{2}h_{\mu\nu}\right) + \frac{1}{\rho_P}\left[(V - \chi)\partial_\mu\phi\partial_\nu\phi + \left(\frac{2\kappa^2V^2 + \kappa^2\chi^2}{4}\right)h_{\mu\nu}\right] \tag{14.10}$$

which is in agreement with GR up to corrections of order $O(1/\rho_P)$. This indicates that $h_{\mu\nu}$ is mainly determined by integrating over the sources (cumulative effects of gravity), whereas Ω and the last term of (14.6) represent local energy-density contributions to the metric. By neglecting the cumulative effects of gravity, which corresponds to the limit $h_{\mu\nu} \approx \eta_{\mu\nu}$, we obtain a kind of special relativistic limit of the theory (or a DSR-like theory [25–28]). In this limit, the metric becomes

$$g_{\mu\nu} \approx \eta_{\mu\nu} + \frac{2}{\rho_P}\left(V\eta_{\mu\nu} + \partial_\mu\phi\partial_\nu\phi\right) + O\left(\frac{1}{\rho_P^2}\right). \tag{14.11}$$

From (14.11) we see that the leading order corrections to the Minkowski metric are strongly suppressed by inverse powers of the Planck density, which indicates that a perturbative study of such contributions in field theories is feasible at low energy densities. This should provide an idea of the kind of corrections induced by the Planck-scale modified Palatini dynamics on Minkowskian field theories. In fact, one can use the metric (14.11) to estimate the first-order modifications of the scalar field equation $\Box\phi - V_\phi = 0$ due to the local energy-density dependence of the metric. After some lengthy algebra, one finds

$$\partial^2\phi - V_\phi \approx 0 + \frac{1}{\rho_P}\left[V_\phi\left(2V - 3\partial^\alpha\phi\partial_\alpha\phi\right) + 2(\partial^\mu\phi\partial^\nu\phi)\partial_\mu\partial_\nu\phi\right], \tag{14.12}$$

where $\partial^2 \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu$. For a massive scalar with $V(\phi) = m^2\phi^2/2$, the term $V_\phi V$ on the right hand side produces the same effect as a $\lambda\phi^4$ interaction in the Lagrangian with $\lambda \equiv m^4/4\rho_P$. The terms involving derivatives of the field are expected to modify the dispersion relation $E^2 = m^2 + k^2$ when the scalar amplitude is sufficiently high. This contrasts with other approaches to quantum gravity phenomenology where the proposed modifications of the dispersion relations introduce higher powers of k^2 but are independent of the field amplitude. The nonlinear dependence on the field amplitude found here is a distinctive characteristic of Palatini theories, which signals the energy-density dependence of its modified dynamics.

14.3 Nonsingular Bouncing Palatini Cosmologies

In the Sect. 14.2 we have studied some perturbative properties of Palatini theories. The full dynamics can be explored in simplified scenarios such as cosmological models. In this sense, it is remarkable that a simple quadratic Lagrangian of the form $f(R) = R + R^2/R_P$ (where $R_P = 1/l_P^2$) does exhibit non-singular solutions [14] for certain equations of state depending on the sign of R_P. To be precise, if $R_P > 0$ the bounce occurs for sources with $w = P/\rho > 1/3$. If $R_P < 0$, then the bouncing condition is satisfied by $w < 1/3$ (see Fig. 14.1). This can be easily understood by having a look at the expression for the Hubble function in a universe filled with radiation plus a fluid with generic equation of state w and density ρ

$$H^2 = \frac{1}{6f_R}\frac{\left[f + (1+3w)\kappa^2\rho + 2\kappa^2\rho_{rad} - \frac{6Kf_R}{a^2}\right]}{\left[1 + \frac{3}{2}\Delta_1\right]^2} \tag{14.13}$$

where $\Delta_1 = -(1+w)\rho\partial_\rho f_R/f_R = (1+w)(1-3w)\kappa^2\rho f_{RR}/(f_R(Rf_{RR} - f_R))$. Due to the structure of Δ_1, one can check that H^2 vanishes when $f_R \to 0$. A more careful analysis shows that $f_R \to 0$ is the only possible way to obtain a bounce with a Palatini $f(R)$ theory that recovers GR at low curvatures if w is constant. In the case of $f(R) =$

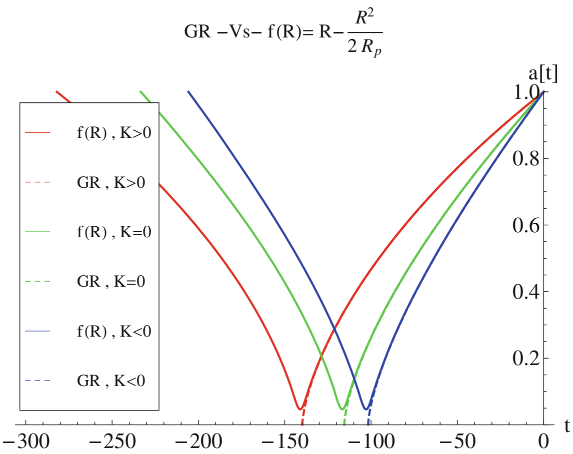


Fig. 14.1 Time evolution of the expansion factor for the model $f(R) = R - R^2/2R_P$ and $w = 0$ for $K > 0$, $K = 0$, and $K < 0$ (*solid curves* from *left* to *right*). From *left* to *right*, we see that the universe is initially contracting, reaches a minimum, and then bounces into an expanding phase. The *dashed lines*, which are only discernible near the bounces, represent the expanding solutions of GR, which begin with a big bang singularity ($a(t) = 0$) and quickly tend to the nonsingular solutions

$R + R^2/R_P$, it is easy to see that $f_R = 0$ has a solution if $1 + 2R_{Bounce}/R_P = 0$ is satisfied for $\rho_{Bounce} > 0$, where $R_{Bounce} = (1 - 3w)\kappa^2\rho_{Bounce}$, which leads to the cases mentioned above. It is worth noting, see Fig. 14.1, that the expanding branch of the non-singular solution rapidly evolves into the solution corresponding to GR. The departure from the GR solution is only apparent very near the bounce, which is a manifestation of the non-perturbative nature of the solution. Note also that in GR there is a solution that represents a contracting branch that ends at the singularity where the expanding branch begins (this solution is just the time reversal of the expanding branch). The Palatini model $f(R) = R - R^2/2R_P$ represented here allows for a smooth transition from the initially contracting branch to the expanding one.

The robustness of the bounce under perturbations can be tested by studying the solutions of these theories in anisotropic spacetimes of Bianchi-I type

$$ds^2 = -dt^2 + \sum_{i=1}^{3} a_i^2(t)(dx^i)^2. \tag{14.14}$$

Despite the complexity of this new scenario, one can derive a number of useful analytical expressions for arbitrary Lagrangian of the type $f(R)$. In particular, one finds that the expansion $\theta = \sum_i H_i$ and the shear $\sigma^2 = \sum_i \left(H_i - \frac{\theta}{3}\right)^2$ (a measure of the degree of anisotropy) are given by

$$\frac{\theta^2}{3}\left(1+\frac{3}{2}\Delta_1\right)^2=\frac{f+\kappa^2(\rho+3P)}{2f_R}+\frac{\sigma^2}{2} \tag{14.15}$$

$$\sigma^2=\frac{\rho^{\frac{2}{1+w}}}{f_R^2}\frac{(C_{12}^2+C_{23}^2+C_{31}^2)}{3}, \tag{14.16}$$

where the constants $C_{ij}=-C_{ji}$ set the amount and distribution of anisotropy and satisfy the constraint $C_{12}+C_{23}+C_{31}=0$. In the isotropic case, $C_{ij}=0$, one has $\sigma^2=0$ and $\theta^2=9H^2$, with H^2 given by Eq. (14.13). Now, since homogeneous and isotropic bouncing universes require the condition $f_R=0$ at the bounce, a glance at (14.16) indicates that the shear diverges as $\sim 1/f_R^2$. This shows that, regardless of how small the anisotropies are initially, isotropic $f(R)$ bouncing models with a single fluid characterized by a constant equation of state will develop divergences when anisotropies are present. This negative result, however, does not arise in extended theories of the form (14.1). For that model one finds that $R=\kappa^2(\rho-3P)$, like in GR, and $Q=Q(\rho,P)$ is given by

$$\frac{Q}{2R_P}=-\left(\kappa^2P+\frac{\tilde{f}}{2}+\frac{R_P}{8}\tilde{f}_R^2\right)+\frac{R_P}{32}\left[3\left(\frac{R}{R_P}+\tilde{f}_R\right)-\sqrt{\left(\frac{R}{R_P}+\tilde{f}_R\right)^2-\frac{4\kappa^2(\rho+P)}{R_P}}\right]^2, \tag{14.17}$$

where $\tilde{f}=R+aR^2/R_P$, and the minus sign in front of the square root has been chosen to recover the correct limit at low curvatures. In a universe filled with radiation, for which $R=0$, the function Q boils down to

$$Q=\frac{3R_P^2}{8}\left[1-\frac{8\kappa^2\rho}{3R_P}-\sqrt{1-\frac{16\kappa^2\rho}{3R_P}}\right]. \tag{14.18}$$

This expression recovers the GR value at low curvatures, $Q\approx 4(\kappa^2\rho)^2/3+32(\kappa^2\rho)^3/9R_P+\ldots$ but reaches a maximum $Q_{max}=3R_P^2/16$ at $\kappa^2\rho_{max}=3R_P/16$, where the squared root of (14.18) vanishes. At ρ_{max} the shear also takes its maximum allowed value, namely, $\sigma_{max}^2=\sqrt{3/16}R_P^{3/2}(C_{12}^2+C_{23}^2+C_{31}^2)$, which is always finite, and the expansion vanishes producing a cosmic bounce regardless of the amount of anisotropy (see Fig. 14.2). Our model, therefore, avoids the well-known problems of anisotropic universes in GR, where anisotropies grow faster than the energy density during the contraction phase leading to a singularity that can only be avoided by sources with $w>1$.

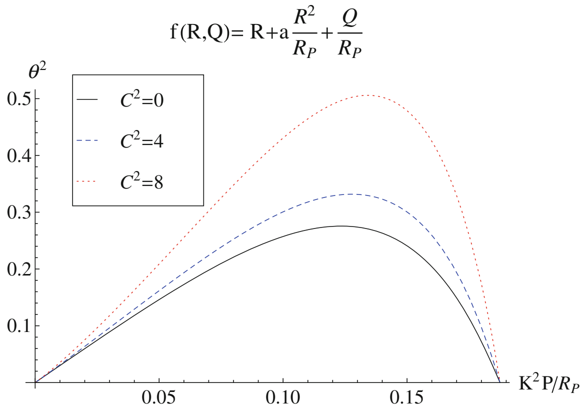


Fig. 14.2 Evolution of the expansion as a function of $\kappa^2\rho/R_P$ in radiation universes with low anisotropy, which is controlled by the combination $C^2 = C_{12}^2 + C_{23}^2 + C_{31}^2$. The case with $C^2 = 0$ corresponds to the isotropic flat case, $\theta^2 = 9H^2$

14.4 Black Holes

Besides early time cosmology, black hole spacetimes represent another scenario where Palatini theories can be tested. Since the modified dynamics of these theories is induced by the existence of matter sources, vacuum configurations are not suitable for our purposes, because they yield exactly the same solutions as in GR. In particular, though the Schwarzschild black hole is the most general spherically symmetric, non-rotating vacuum solution of GR and also of (14.1), that solution assumes that all the matter is concentrated on a point of infinite density, which is not consistent with the dynamics of (14.1). In fact, if one considers the collapsing object as described by a perfect fluid that behaves as radiation during the last stages of the collapse, explicit computation of the scalar $Q = R_{\mu\nu}R^{\mu\nu}$ [see (14.18)] shows that the energy density ρ is bounded from above by $\kappa^2\rho_{max} = 3R_P/16$, as we saw in the Sect.14.3. Therefore, one should study the complicated dynamical process of collapse of a spherical non-rotating object to determine how the Schwarzschild metric is modified in our theory. For this reason it is easier to study instead vacuum space-times with an electric field, which possess a non-zero stress-energy tensor able to excite the Palatini dynamics even in static settings. The resulting solutions should thus be seen as Planck-scale modifications of the usual Reissner-Nordström solution of GR.

In the context of $f(R)$ theories, electrically charged black holes do not produce any new structures unless one considers that the electromagnetic field is described by some non-linear extension of Maxwell's electrodynamics. This is so because the modified dynamics of $f(R)$ theories is sensitive only to the trace of the energy-momentum tensor of the matter-energy sources, which in the case of Maxwell's theory is zero. For non-linear theories of electrodynamics, like the Born-Infeld model, the corresponding energy-momentum tensor is not traceless and, therefore, is able to produce departures from GR. These black holes have been recently studied in detail in [19], where it has been shown that exact analytical solutions can be found. In that work one finds that the combination of a quadratic Palatini Lagrangian with

Born-Infeld theory can dramatically reduce the intensity of the divergence associated to the singularity. For instance, in GR with Maxwell's electrodynamics one finds that the Kretschmann scalar takes the form

$$R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta} = \frac{48M_0^2}{r^6} - \frac{48M_0 r_q^2}{r^7} + \frac{14r_q^4}{r^8}, \tag{14.19}$$

which implies a strong divergence, $\sim 1/r^8$, as $r \to 0$. In GR coupled to Born-Infeld theory, one finds that the divergence is dominated by $R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta} \sim 1/r^4$. If the gravity Lagrangian is taken as $f(R) = R - l_P^2 R^2$, then the divergence is further reduced to $R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta} \sim 1/(r - r_+)^2$, where $r_+ > 0$ defines the surface of a sphere that contains all the matter and charge of the black hole. Though this solution does not avoid the singularity, it does introduce an important qualitative change with respect to GR, namely, that the matter and charge distribution of the collapsed object are no longer concentrated on a point, but on a compact sphere.

The results of [19] that we have just summarized suggest that nonperturbative quantum gravitational effects could halt gravitational collapse and produce regular objects sustained by some kind of quantum degeneracy pressure induced by the gravitational interaction, in much the same way as neutron stars and white dwarfs arise when the quantum degeneracy pressure of matter dominates in the interior of stars. The theory (14.1) does exactly this [20] . If standard electrically charged black holes are considered under the gravity theory (14.1), one finds that completely regular solutions exist. These solutions exhibit a compact core of area $A_{core} = N_q\sqrt{2\alpha_{em}}A_P$, where $A_P = 4\pi l_P^2$ is Planck's area, N_q is the number of charges, and α_{em} is the electromagnetic fine structure constant, which contains all the mass of the collapsed object at a density $\rho^*_{core} = M_0/V_{core} = \rho_P/4\delta_1^*$, independent of q and M_0. The electric charge is distributed on the surface of this core and its density is also a universal constant independent of q and M_0, namely, $\rho_q = q/(4\pi r_{core}^2) = (4\pi\sqrt{2})^{-1}\sqrt{c^7/(\hbar G^2)}$. The impact that these results could have for the theoretical understanding of black holes and the experimental search of compact objects in particle accelerators are currently under investigation.

14.5 Conclusion

We have shown that a simple extension of general relativity at the Planck scale formulated à la Palatini successfully addresses different aspects of quantum gravity phenomenology, such as the consistent introduction of a minimum length compatible with the principle of relativity, the avoidance of the big bang singularity, and also the modification of black hole interiors developing a nonsingular compact core that contains all the mass and charge of the collapsed object. In summary, the model (14.1) does everything it was expected to do and lacks of any known instabilities.

Acknowledgments Work supported by the Spanish grant FIS2008-06078-C03-02, and the Programme CPAN (CSD2007-00042).

References

1. Will CM (2005) Living Rev Rel 9:3 [arXiv:gr-qc/0510072].
2. Ortin T (2004) Gravity and strings. Cambridge University Press, Cambridge
3. Green M, Schwarz J, Witten E (1987) Superstring theory. Cambridge University Press, Cambridge
4. Thiemann T (2007) Modern canonical quantum general relativity. Cambridge Univeristy Press, Cambridge
5. Rovelli C (2004) Quantum gravity. Cambridge Univeristy Press, Cambridge
6. Ashtekar A, Lewandowski J (2004) Class Quant Grav 21:R53
7. Capozziello S, De Laurentis M, arXiv:1108.6266 [gr-qc].
8. De Felice A, Tsujikawa S (2010) Living Rev Rel 13:3 [arXiv:1002.4928 [gr-qc]].
9. Sotiriou TP, Faraoni V (2010) Rev Mod Phys 82:451–497 arXiv:0805.1726 [gr-qc].
10. Capozziello S, Francaviglia M (2008) Gen Rel Grav 40:357
11. Nojiri S, Odintsov SD (2007) Int J Geom Methods Mod Phys 4:115 arXiv:1011.0544 [gr-qc].
12. Lue A (2006) Phys Rept 423:1–48 [astro-ph/0510068]
13. Sahni V [astro-ph/0502032].
14. J. Bruneton, P, Esposito-Farese G (2007), Phys Rev D 76:124012 [arXiv:0705.4043 [gr-qc]].
15. Olmo GJ (2011) Int J Mod Phys D 20:413 [arXiv:1101.3864 [gr-qc]].
16. Olmo GJ (2011) JCAP10 018 [arXiv:1101.2841 [gr-qc]].
17. Barragan C, Olmo GJ (2010) Phys Rev D82:084015
18. Barragan C, Olmo GJ, Sanchis-Alepuz H (2009) Phys Rev D80:024016
19. Olmo GJ, Rubiera-Garcia D arXiv:1110.0850 [gr-qc].
20. Olmo GJ, Rubiera-Garcia D arXiv:1112.0475 [gr-qc].
21. Buchdahl HA (1979) J Phys A:Math Gen 12:1235
22. Tapia V, Ujevic M (1998) Class Quant Grav 15:3719
23. Vitagliano V, Sotiriou TP, Liberati S (2010) Phys Rev D 82:084007
24. Olmo GJ, Sanchis-Alepuz H, Tripathi S (2009) Phys Rev D 80:024013
25. Amelino-Camelia G (2002) Int J Mod Phys D 11:35
26. Magueijo J, Smolin L (2002) Phys Rev Lett 88:190403
27. Magueijo J, Smolin L (2003) Phys Rev D 67:044017
28. Magueijo J, Smolin L (2004) Class Quant Grav 21:1725

Chapter 15
Extended Gravity from Noncommutativity

Paolo Aschieri

Abstract We present a first order theory of gravity (vierbein formulation) on noncommutative spacetime. The first order formalism allows to couple the theory to fermions. This NC action is then reinterpreted (using the Seiberg-Witten map) as a gravity theory on commutative spacetime that contains terms with higher derivatives and higher powers of the curvature and depend on the noncommutativity parameter θ. When the noncommutativity is switched off we recover the usual gravity action coupled to fermions. The first nontrival corrections to the usual gravity action coupled to fermions are explicitly calculated.

15.1 Introduction

In the passage from classical mechanics to quantum mechanics classical observables become noncommutative. Similarly we expect that in the passage from classical gravity to quantum gravity, gravity observables, i.e. spacetime itself, with its coordinates and metric structure, will become noncommutative. Thus by formulating Einstein gravity on noncommutative spacetime we may learn some aspects of quantum gravity.

Planck scale noncommutativity is further supported by Gedanken experiments that aim at probing spacetime structure at very small distances. They show that due to gravitational backreaction one cannot test spacetime at those distances. For example, in relativistic quantum mechanics the position of a particle can be detected with a precision at most of the order of its Compton wave length $\lambda_C = \hbar/mc$. Probing spacetime at infinitesimal distances implies an extremely heavy particle that in turn curves spacetime itself. When λ_C is of the order of the Planck length, the spacetime

*The results presented in these proceedings are based on joint work with Leonardo Castellani.

P. Aschieri (✉)
Dipartimento di Scienze e Innovazione Tecnologica and INFN Gruppo collegato di Alessandria, Università del Piemonte Orientale, Viale T. Michel 11, 15121 Alessandria, Italy
e-mail: aschieri@to.infn.it

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_15,

curvature radius due to the particle has the same order of magnitude and the attempt to measure spacetime structure beyond Planck scale fails.

This Gedanken experiment supports finite reductionism. It shows that the description of space time as a continuum of points (a smooth manifold) is an assumption no more justified at Planck scale. It is then natural to relax this assumption and conceive a noncommutative spacetime, where uncertainty relations and discretization naturally arise. In this way the dynamical feature of spacetime that prevents from testing sub-Plankian scales is explained by incorporating it at a deeper kinematic level. A similar mechanism happens for example in the passage from Galilean to special relativity. Contraction of distances and time dilatation can be explained in Galilean relativity: they are a consequence of the interaction between ether and the body in motion. In special relativity they become a kinematic feature.

The noncommutative gravity theory we present following [1, 2] is an effective theory that may capture some aspects of a quantum gravity theory. Furthermore we reinterpret spacetime noncommutativity as extra interaction terms on commutative spacetime, in this way the theory is equivalent to a higher derivative and curvature extension of Einstein general relativity. We have argued that spacetime noncommutativity should be relevant at Planck scale, however the physical phenomena it induces can also appear at larger scales. For example, due to inflation, noncommutativity of spacetime at inflation scale (that may be as low as Planck scale) can affect cosmological perturbations and possibly the cosmic microwave background spectrum; see for example [3]. We cannot exclude that this noncommutative extension of gravity can be relevant for advancing in our understanding of nowadays open questions in cosmology.

In this contribution, after a short overview of possible noncommutative approaches, we outline the Drinfeld twist approach and review the geometric formulation of theories on noncommutative spacetime [4]. This allows to construct actions invariant under diffeomorphisms. In Sect. 15.4 we first present usual gravity coupled to fermions in an index free formalism suited for its generalization to the noncommutative case. Then we discuss gauge theories on noncommutative space and in particular local Lorentz symmetry ($SO(3,1)$-gauge symmetry), indeed we need a vierbein formulation of noncommutative gravity in order to couple gravity to spinor fields. The noncommutative Lagrangian coupled to spinor fields is then presented. In Sect. 15.12 we reinterpret this NC gravity as an extended gravity theory on commutative spacetime. This is done via the Seiberg-Witten map from noncommutative to commutative gauge fields. The resulting gravity theory then depends on the usual gravitational degrees of freedom plus the noncommutative degrees of freedom, these latter are encoded in a set of mutually commuting vector fields $\{X_I\}$. The leading correction terms to the usual action are explicitly calculated in Sect. 15.13. They couple spinor fields and their covariant derivatives to derivatives of the curvature tensor and of the vierbein. It is interesting to consider a kinetic term for these vector fields, so that the noncommutative structure of spacetime, as well as its metric structure depend on the matter content of spacetime itself. A model of dynamical noncommutativity is presented in [5]. Noncommutative vierbein gravity can also be coupled to scalar fields [5] and to gauge fields [6].

15.2 NC Geometry Approaches

Before entering the details of the theory, we briefly frame it in the context of noncommutative geometry approaches.

The easiest way to describe a noncommutative spacetime is via the noncommutative algebra of its coordinates, i.e., we give a set of generators and relations. For example

$$\left[x^{\mu}, x^{\nu}\right] = \iota\theta^{\mu\nu} \qquad \textit{canonical} \tag{15.1}$$

$$\left[x^{\mu}, x^{\nu}\right] = i f_{\sigma}^{\mu\nu} x^{\sigma} \qquad \textit{Lie algebra} \tag{15.2}$$

$$x^{\mu}x^{\nu} - q^{x^{\mu}x^{\nu}} = 0 \qquad \textit{quantum (hyper)plane} \tag{15.3}$$

where $\theta^{\mu\nu}$ (a real antisymmetric matrix), $f_{\sigma}{}^{\mu\nu}$ (real structure constants) , q (a complex number, e.g. a phase) are the respective concommutativity parameters. Quantum groups and quantum spaces ([7, 8]) are usually described in this way. In this case we do not have a space (i.e. a set of points), rather we have a noncommutative algebra generated by the coordinates x^{μ} and their relations; when the noncommutativity parameters ($\theta^{\mu\nu}$, $f_{\sigma}{}^{\mu\nu}$, q) are turned off this algebra becomes commutative and is the algebra of functions on a usual space.

Of course we can also impose further constraints, for example periodicity of the coordinates describing the canonical noncommutative spacetime (15.1) (that typical of phase-space quantum mechanics) leads to a noncommutative torus rather than to a noncommutative (hyper)plane. Similarly, constraining the coordinates of the quantum (hyper)plane relations (15.3) we obtain a quantum (hyper)sphere.

This algebraic description should then be complemented by a topological approach. One that for example leads to the notions of continuous functions. This is achieved completing the algebra generated by the noncommutative coordinates to a $C^{\bigstar}$-algebra. Typically $C^{\bigstar}$-algebras arise as algebras of operators on Hilbert space. Connes noncommutative geometry [9] starts from these notions and enriches the $C^{\bigstar}$-algebra structure and its representation on Hilbert space so to generalize to the noncommutative case also the notions of smooth functions and metric structure.

Another approach is the $\bigstar$-product one. Here we retain the usual space of functions from commutative space to complex numbers, but we deform the pointwise product operation in a $\bigstar$-product one. A $\bigstar$-product sends two functions (f , g) in a third one ($f\bigstar g$). It is a differential operator on both its arguments (hence it is frequently called a bi-differential operator). It has the associative property $f\bigstar(g\bigstar h) = (f\bigstar g)\bigstar h$. The most known example is the Gronewold- Moyal-Weyl star product on R^{2n},

$$\left(f\bigstar h\right)(x) = e^{\frac{i}{2}\theta^{\mu\nu}\frac{\partial}{\partial x^{\mu}}\otimes\frac{\partial}{\partial y^{\nu}}} f(x)h(y)\Big|_{x=y}. \tag{15.4}$$

Notice that if we set

$$\mathcal{F}^{-1} = e^{\frac{i}{2}\theta^{\mu\nu}\frac{\partial}{\partial x^\mu}\otimes\frac{\partial}{\partial y^\nu}}$$

then

$$(f \star h)(x) = \mu \circ \mathcal{F}^{-1}(f \otimes h)(x)$$

where μ is the usual product of functions $\mu(f \otimes g) = fh$. The element $\mathcal{F} = e^{-\frac{i}{2}\theta^{\mu\nu}\frac{\partial}{\partial x^\mu}\otimes\frac{\partial}{\partial y^\nu}}$ is an example of a Drinfeld twist. It is defined by the exponential series in powers of the noncommutativityparameters $\theta^{\mu\nu}$,

$$\mathcal{F} = e^{-\frac{i}{2}\theta^{\mu\nu}\frac{\partial}{\partial x^\mu}\otimes\frac{\partial}{\partial y^\nu}} = 1 \otimes 1 - \frac{i}{2}\theta^{\mu\nu}\partial_\mu \otimes \partial_\nu - \frac{1}{8}\theta^{\mu_1\nu_1}\theta^{\mu_2\nu_2}\partial_{\mu_1}\partial_{\mu_2} \otimes \partial_{\nu_1}\partial_{\nu_2} + \dots$$

It is easy to see that $x^\mu \star x^\nu - x^\nu \star x^\mu = i\theta^{muv}$ thus also in this approach we recover the noncommutative algebra (15.1).

In this paper noncommutative spacetime will be spacetime equipped with a ★-product. We will not discuss when the exponential series $f \star g = fg - \frac{i}{2}\theta^{\mu\nu}\partial_\mu(f)\partial_\nu(g) + \dots$ defining the function f ★ g is actually convergent. We will therefore work in the well established context of formal deformation quantization [10]. In the latter part of the paper we will consider a series expansion of the noncommutative gravity action in powers of the noncommutativity parameters θ, we will present the first order in θ (a second order study appears in [2]), therefore the convergence aspect won't be relevant.

The method of constructing ★-products using twists is not the most general method, however it is quite powerful, and the class of ★-products obtained is quite wide. For example choosing the appropriate twist we can obtain the noncommutative relations (15.1) (15.2) and also (depending on the structure constant explicit expression) some of the Lie algebra type (15.3).

15.3 Twists and ★-Noncommutative Manifolds

Let M be a smooth manifold, a twist is an invertible element F ∈ UΞ⊗ UΞ where UΞ is the universal enveloping algebra of vector fields, (i.e. it is the algebra generated by vector fields on M and where the element XY - Y X is identified with the vector field [X, Y]). The element F must satisfy some further conditions that we do not write here, but that are satisfied if we consider abelian twists, i.e., twists of the form

$$\begin{aligned}\mathcal{F} &= e^{-\frac{i}{2}\theta^{IJ}X_I\otimes X_J}\\ &= 1 \otimes 1 - \frac{i}{2}\theta^{IJ}X_I \otimes X_j \frac{1}{8}\theta^{I_1J_1}\theta^{I_2J_2}X_{I_1}X_{I_2} \otimes X_{J_1}X_{J_2} + \dots\end{aligned}$$

were the vector fields $X_I(I = 1, ...s$ with s not necessarily equal to m = dim M) are mutually commuting $[X_I, X_J] = 0$ (hence the name abelian twist).

It is convenient to introduce the following notation

$$\begin{aligned}\mathcal{F}^{-1} &= 1 \otimes 1 + \frac{i}{2}\theta^{IJ} X_I \otimes X_J \frac{1}{8}\theta^{I_1 J_1}\theta^{I_2 J_2} X_{I_1} X_{I_2} \otimes X_{J_1} X_{J_2} + \dots \\ &= \bar{\mathrm{f}}^{\alpha} \otimes \bar{\mathrm{f}}_{\alpha}\end{aligned} \tag{15.5}$$

where a sum over the multi-index α is understood.

Let A be the algebra of smooth functions on the manifold M. Then, given a twist F , we deform A in a noncommutative algebra $A_{\bigstar}$ by defining the new product of functions

$$f \bigstar h = \bar{\mathrm{f}}^{\alpha}(f)\bar{\mathrm{f}}_{\alpha}(h)$$

we see that this formula is a generalization of the Gronewold-Moyal-Weyl star product on R^{2n} defined in (15.4). Since the vector fields X_I are mutually commuting then this $\leftrightarrow$-product is associative. Note that only the algebra structure of A is changed to $A_{\bigstar}$ while, as vector spaces, A and $A_{\bigstar}$ are the same. We similarly consider the algebra of exterior forms $\Omega^{\bullet}_{\bigstar}$ with the wedge product $\wedge$, and deform it in the noncommutative exterior algebra $\Omega^{\bullet}_{\bigstar}$ that is characterized by the graded noncommutative exterior product $\wedge_{\bigstar}$ given by

$$\tau \wedge_{\bigstar} \tau' = \bar{\mathrm{f}}^{\alpha}(\tau) \wedge \bar{\mathrm{f}}(\tau'),$$

where τ and τ' are arbitrary exterior forms. Notice that the action of the twist on τ and τ^I is via the Lie derivative: each vector field $X_{I_1}, X_{I_2}, X_{J_1}, X_{J_2} \dots$ in (15.5) acts on forms via the Lie derivative.

It is not difficult to show that the usual exterior derivative is compatible with the new $\wedge_{\bigstar}$-product,

$$d(\tau \wedge_{\star} \tau') = d(\tau) \wedge_{\star} \tau' + (-1)^{deg(\tau)} \tau \wedge_{\star} d\tau' \tag{15.6}$$

this is so because the Lie derivative commutes with the exterior derivative.

We also have compatibility with the usual undeformed integral (graded cyclicity property):

$$\int \tau \wedge_{\star} \tau' = (-1)^{degr(\tau)deg(\tau')} \int \tau' \wedge_{\star} \tau \tag{15.7}$$

Note. We remark that all these properties are due to the special nature of the $\leftrightarrow$-product we consider. As shown in [10] $\leftrightarrow$-products are in 1-1 correspondence with Poisson structures { , } on the manifold M. The Poisson structure the twist F induces is $\{f, g\} = i\theta^{IJ} X_I(f) X_J(g)$. However the twist F encodes more information than the Poisson bracket { , }. The key point is that the twist is associated with the Lie algebra Ξ (and morally with the diffeomorphisms group of M). Given a twist we can deform the Lie algebra Ξ (the diffeomorphisms group of M) and then, using the Lie derivative action (the action of the diffeomorphisms group) we induce noncommutative deformations of the algebra of functions on M, of the exterior algebra and more generally of the differential and Riemannian geometry structures on M [11], leading to noncommutative Einstein equations for the metric tensor [11].

15.4 Noncommutative Vierbein Gravity Coupled to Fermions

15.5 Classical Action

The usual action of first-order gravity coupled to spin $\frac{1}{2}$ fields reads:

$$S = \varepsilon_{abcd} \int R^{ab} \wedge V^c \wedge V^d - i\bar{\psi}\gamma^a V^b \wedge V^c \wedge V^d \wedge D\psi - i(D\psi)\gamma^a \wedge V^b \wedge V^c \wedge V^d \psi \tag{15.8}$$

with $R^{ab} = d\omega^{ab} - \omega^a \wedge \ \ \omega^{cb}$, and the Dirac conjugate defined as usual: $\bar{\psi} = \psi^\dagger \gamma_0$. This action can be recast in an index-free form ([12], Aschieri 2009), convenient for generalization to the non-commutative case:

$$S = \int Tr(iR \wedge V \wedge V\gamma_5) + \bar{\psi} V \wedge V \wedge V\gamma_5 D\psi + D\bar{\psi} \wedge V \wedge V \wedge V\gamma_5\psi \tag{15.9}$$

where

$$R = d\Omega - \Omega \wedge \Omega, \quad D\psi = d\psi - \Omega\psi, \quad D\bar{\psi} = \overline{D\psi} = d\bar{\psi} + \bar{\psi}\Omega \tag{15.10}$$

with

$$\Omega \equiv \frac{1}{4}\omega^{ab}\gamma_{ab}, \quad V \equiv V^a\gamma_a, \quad R \equiv \frac{1}{4}R^{ab}\gamma_{ab}$$

taking value in Dirac gamma matrices. Use of the gamma matrix identities $\gamma_{abc} = i\varepsilon_{abcd}\gamma^d\gamma_5$, $Tr(\gamma_{ab}\gamma_c\gamma_d\gamma_5) = -4i\varepsilon_{abcd}$ in computing the trace leads back to the usual action (15.8).

The action (15.9) is invariant under local diffeomorphisms (because it is the integral of a 4-form on a 4-manifold) and under local Lorentz rotations. In the index-free form they read

$$\delta_\varepsilon V = -[V, \varepsilon], \delta_\varepsilon \Omega = d\varepsilon - [\Omega, \varepsilon], \delta_\varepsilon \psi = \varepsilon\psi, \delta_\varepsilon \bar{\psi} = -\bar{\psi}\varepsilon \tag{15.11}$$

with $\varepsilon = \frac{1}{4}\varepsilon^{ab}\gamma_{ab}$. The local Lorentz invariance of the index free action follows from $\delta_\varepsilon R = -[R, \varepsilon]$ and $\delta_\varepsilon D\psi = \varepsilon D\psi$, the cyclicity of the trace Tr and the fact that the gauge parameter ε commutes with γ_5.

15.6 Noncommutative Gauge Theory and Lorentz Group

Consider an infinitesimal gauge transformation $\lambda = \lambda^A T^A$, where the generators T^A belong to some representation (the fundamental, the adjoint, etc.) of a Lie group G. Since two consecutive gauge transformations are a gauge transformation, the commutator of two infinitesimal ones $[\lambda, \lambda']$, closes in the Lie algebra of G. This in general is no more the case in noncommutative gauge theories. In the noncommutative case the commutator of two infinitesimal gauge transformations is

$$[\lambda \star \lambda'] \equiv \lambda \star \lambda' - \lambda' \star \lambda = \frac{1}{2}\{\lambda^A \star \delta'^B\}[T^A, T^B] + \frac{1}{2}[\lambda^A \star \lambda'^B]\{T^A, T^B\}.$$

We see that also the anticommutator $\{ T^A, T^B \}$ appears. This is fine if our gauge group is for example $U(N)$ or $GL(N)$ in the fundamental or in the adjoint, since in this case $\{ T^A, T^B\}$ is again in the Lie algebra, however for more general Lie algebras (including all simple Lie algebras) we have to enlarge the Lie algebra to include also anticommutators besides commutators, i.e. we have to consider all possible products $T^A T^B \dots T^C$ of generators.

Our specific case is the Lorentz group in the spinor representation given by the Dirac gamma matrices γ_{ab}. The algebra generated by these gamma matrices is that of all even 4×4 gamma matrices. The noncommutative gauge parameter will therefore have components

$$\varepsilon = \frac{1}{4}\varepsilon^{ab}\gamma_{ab} + i\varepsilon 1\!\!1 + \tilde{\varepsilon}\gamma_5.$$

The extra gauge parameters $\varepsilon, \tilde{\varepsilon}$ can be chosen to be real (like ε_{ab}). Indeed the reality of $\varepsilon_{ab}, \varepsilon, \tilde{\varepsilon}$ is equivalent to the hermiticity condition

$$-\gamma_0\varepsilon\gamma_0 = \varepsilon^\dagger \tag{15.12}$$

and if the gauge parameters $\varepsilon, \varepsilon^I$ satisfy this condition then also $[\varepsilon \leftrightarrow \varepsilon^I]$ is easily seen to satisfy this hermiticity condition.

We have centrally extended the Lorentz group to

$$SO(3, 1) \rightarrow SO(3, 1) \times U(1) \times R^+,$$

or more precisely, (since our manifold M has a spin structure and we have a gauge theory of the spin group $SL(2, C)$)

$$SL(2, C) \rightarrow GL(2, C).$$

The Lie algebra generator $i\,\Pi$ is the anti-hermitian generator corresponding to the $U(1)$ extension, while γ_5 is the hermitian generator corresponding to the noncompact R^+ extension.

Since under noncommutative gauge transformations we have

$$\delta_\varepsilon \Omega = d\varepsilon - \Omega \star \varepsilon + \varepsilon \star \Omega \tag{15.13}$$

also the spin connection and the curvature will be valued in the $GL(2, C)$ Lie algebra representation given by all the even gamma matrices,

$$\Omega = \frac{1}{4}\omega^{ab}\gamma_{ab} + i\omega\Pi + \tilde{\omega}\gamma_5, \quad R = \frac{1}{4}R^{ab}\gamma_{ab} + ir\Pi + \tilde{r}\gamma_5. \tag{15.14}$$

Similarly the gauge transformation of the vierbein,

$$\delta_\varepsilon V = -V \star \quad \varepsilon + \varepsilon \quad \star V, \tag{15.15}$$

closes in the vector space of odd gamma matrices (i.e. the vector space linearly generated by γ^a, $\gamma^a\gamma_5$) and not in the subspace of just the γ^a matrices. Hence the noncommutative vierbein are valued in the odd gamma matrices

$$V = V^a\gamma_a + \tilde{V}^a\gamma_a\gamma_5. \tag{15.16}$$

15.7 Noncommutative Gravity Action and its Symmetries

We have all the ingredients in order to generalize to the noncommutative case the gravity action coupled to spinors of Sect. 15.3: an abelian twist giving the star products of functions and forms on the spacetime manifold M (and compatible with usual integration on M); an extension to $GL(2, C)$ of the Lorentz gauge group, so that infinitesimal noncommutative gauge transformations close in this extended Lie algebra. The action reads

$$S = \int Tr(iR \wedge_\star V \wedge_\star V\gamma_5 + \bar{\psi} \star V \wedge_\star V \wedge_\star V \wedge_\star \gamma_5 D\psi + D\bar{\psi} \wedge_\star V \wedge_\star V \wedge_\star V \star \gamma_5\psi \tag{15.17}$$

with

$$R = d\Omega - \Omega \wedge_\star \Omega, \quad D\psi = d\psi - \Omega \star \psi, \quad D\bar{\psi} = d\bar{\psi} + \bar{\psi} \star \Omega \tag{15.18}$$

15.8 *Gauge Invariance*

The invariance of the noncommutative action (15.17) under the $\star$-variations is demonstrated in exactly the same way as for the commutative case: noting that besides (15.13) and (15.15) we have

$$\delta_\varepsilon \psi = \varepsilon \star \psi, \quad \delta_\varepsilon \bar{\psi} = \delta_\varepsilon \bar{\psi} \star \varepsilon, \quad \delta_\varepsilon D\psi = \varepsilon \star D\psi, \quad , \delta_\varepsilon D\bar{\psi} = -D\bar{\psi} \star \varepsilon,$$
$$\delta_\varepsilon R = -R \star \varepsilon + \varepsilon \star R, \tag{15.19}$$

and using that ε commutes with γ_5, and the cyclicity of the trace together with the graded cyclicity of the integral with respect to the $\leftrightarrow$-product.

15.9 *Diffeomorphisms Invariance*

The $\star$-action (15.17) is invariant under usual diffeomorphisms, being the integral of a 4-form. Under these diffeomorphisms the vector fields X_I transform covariantly. We also mention that since the vector fields X_I appear only in the $\star$-product, the action is furthermore invariant under $\star$-diffeomorphisms as defined in [11], see discussion in [4], Sect. 7.2.4. Under these deformed diffeomorphisms the vector fields X_I do not transform.

15.10 *Reality of the Action*

Hermiticity conditions can be imposed on the fields V and Ω as done with the gauge parameter ε in (15.12) :

$$\gamma_0 V \gamma_0 = V^\dagger, \quad -\gamma_0 \Omega \gamma_0 = \Omega^\dagger. \tag{15.20}$$

These hermiticity conditions are consistent with the gauge variations and can be used to check that the action (15.17) is real by comparing it to its complex conjugate (obtained by taking the Hermitian conjugate of the 4-form inside the trace in the integral). As previously observed for the component gauge parameters ε^{ab}, ε, $\tilde{\varepsilon}$, the hermiticity conditions (15.20) imply that the component fields V^a, $\tilde{V}^a$, ω^{ab}, ω, and $\tilde{\omega}$ are real fields.

15.11 *Charge Conjugation Invariance*

Noncommutative charge conjugation is the following transformation (extended linearly and multiplicatively):

$$\psi \to \psi^C \equiv C(\bar{\psi})^T = -\gamma_0 C \psi^*, V \to V^C \equiv C\ V^T C, \quad \Omega \to \Omega^C \equiv C \Omega^T C,$$
$$\star_\theta \to \star_\theta^C \equiv \star_{-\theta}, \tag{15.21}$$

and consequently $\wedge_{\star\theta} \to \wedge_{\star\theta}^C = \wedge_{\star-\theta}$. Then the action (15.17) is invariant under charge conjugation. For example

$$
\begin{aligned}
S^C_{bosonic} &= i\int Tr(R^C \wedge_{-\theta} V^C \wedge_\theta V^C\gamma_5)^T = -i\int Tr(R^T \wedge_{-\theta} V^T C\gamma_5 C^{-1})^T \\
&= -i\int Tr\Big((V^T \wedge_{-\theta} V^T\gamma_5^T)^T \wedge_\star R\Big) = i\int Tr\Big(-(V^T\gamma_5^T)^T \wedge_\star V \wedge_\star R\Big) \\
&= i\int Tr(\gamma_5 V \wedge_\star V \wedge_\star R) = i\int Tr(R \wedge_\star \gamma_5 V \wedge V) = i\int Tr(R \wedge_\star V \wedge_\star V\gamma_5) \\
&= S_{bosonic} \qquad (15.22)
\end{aligned}
$$

Similarly the fermionic part of the action satisfies $S^C{}_{fermionic} = S_{fermionic}$.

The vanishing of the $\tilde{V}^a$ components in the classical limit is achieved by imposing charge conjugation conditions on the fields [1]:

$$
CV_\theta(x)C = V_{-\theta}(x)^T, \quad C\Omega_\theta(x)C = \Omega_{-\theta}(x)^T, \quad C\varepsilon_\theta(x)C = \varepsilon_{-\theta}(x)^T \qquad (15.23)
$$

These conditions involve the θ-dependence of the fields. This latter is due to the $\star$-product θ-dependence (recall that the $\leftrightarrow$-product is defined as an expansion in power series of the noncommutativity parameter θ). Since noncommutative gauge transformations involve the $\leftrightarrow$-product, the gauge transformed fields will be θ-dependent and hence field configurations are in general θ-dependent.

Conditions (15.23) are consistent with the $\star$-gauge transformations. For example the field $CV_\theta(x)^T C$ can be shown to transform in the same way as $V^{-\theta}(x)$ [1].

For the component fields and gauge parameters the charge conjugation conditions imply that the components V^a, ω^{ab} are even in θ, while the components $\tilde{V}^a, \omega, \tilde{\omega}$ are odd

$$
V^a_\theta = V^a_{-\theta}, \quad \omega^{ab}_\theta = \omega^{ab}_{-\theta} \qquad (15.24)
$$

$$
\tilde{V}^a_\theta = -\tilde{V}^a_{-\theta}, \omega_\theta = -\omega_{-\theta}, \tilde{\omega}_\theta = -\tilde{\omega}_{-\theta}. \qquad (15.25)
$$

Similarly for the gauge parameters: $\varepsilon^{ab}_\theta = \varepsilon^{ab}_{-\theta}, \varepsilon_\theta = -\varepsilon_{-\theta}, \tilde{\varepsilon}_\theta = -\tilde{\varepsilon}_{-\theta}$. In particular, since the components $\tilde{V}^a$ are odd in θ we achieve their vanishing in the classical limit.

We can also conclude that the bosonic action is even in θ. Indeed (15.23) implies

$$
V^C = V_{-\theta}, \Omega^C = \Omega_{-\theta}, R^C = R_{-\theta}. \qquad (15.26)
$$

Hence the bosonic action $S_{bosonic}(\theta)$ is mapped into $S_{bosonic}(-\theta)$ under charge conjugation. Also for the fermionic action, $S_{fermionic}(\theta)$, we have $S_{fermionic}(\theta)^C = S_{fermionic}(-\theta)$ if the fermions are Majorana, i.e. if they satisfy $\psi^C = \psi$. From $S_{bosonic}(\theta)^C = S_{bosonic}(-\theta)$ we conclude that all noncommutative corrections to the classical action of pure gravity are even in θ; this is also the case if we couple noncommutative gravity to Majorana fermions

15.12 Seiberg-Witten map (SW map)

In Sect. 15.11 we have formulated a noncommutative gravity theory that in the classical limit $\theta \to 0$ reduces to usual vierbein gravity. In the full noncommutative regime it has however a doubling of the vierbein fields. We can insist on a noncommutative gravity theory that has the same degrees of freedom of the classical one. This is doable if we use the Seiberg-Witten map to express the noncommutative fields in terms of the commutative ones. In this way the gauge group is the Lorentz group $SL(2, C)$ and not the centrally extended one $GL(2, C)$, indeed also the noncommutative gauge parameters ε^{ab}, ε, $\tilde{\varepsilon}$ are expressed in term of the commutative ones ε^{ab}.

Because of the SW map the noncommutative fields can therefore be expanded in terms of the commutative ones, and hence the noncommutative gravity action can be expanded, order by order in powers of θ, in terms of the usual commutative gravity and spinor field as well as of the noncommutativity vector fields X_I. As we will see, we thus obtain a commutative gravity action that at zeroth order in θ is usual gravity (coupled to spinors) and at higher orders in θ contains higher derivative terms describing gravity and spinor fields coupled to the noncommutativity vector fields X_I.

The Seiberg-Witten map (SW map) relates the noncommutative gauge fields Ω to the ordinary Ω^0, and the noncommutative gauge parameters ε to the ordinary ε^0 and Ω^0so as to satisfy:

$$\Omega(\Omega^0) + \delta_\varepsilon \Omega(\Omega^0) = \Omega(\Omega^0 + \delta_{\varepsilon^0}\Omega^0) \tag{15.27}$$

With the noncommutative and ordinary gauge variation given by

$$\delta_\varepsilon \Omega = \mathrm{d}\varepsilon - \Omega \star \varepsilon + \varepsilon \star \Omega, \quad \delta_{\varepsilon^0}\Omega^0 = \mathrm{d}\varepsilon^0 - \Omega^0\varepsilon^0 + \varepsilon^0\Omega^0. \tag{15.28}$$

Equation (15.27) can be solved order by order in θ [13], yielding Ω and ε as power series in θ:

$$\Omega(\Omega^0, \theta) = \Omega^0 + \Omega^1(\Omega^0) + \Omega^2(\Omega^0) + \ldots + \Omega^n(\Omega^0) + \ldots \tag{15.29}$$

$$\varepsilon(\varepsilon^0, \Omega^0, \theta) = \varepsilon^0 + \varepsilon^1(\varepsilon^0, A^0) + \varepsilon^2(\varepsilon^0, \Omega^0) + \ldots + \varepsilon^n(\varepsilon^0, \Omega^0) + \ldots \tag{15.30}$$

where $\Omega^n(\Omega^0)$ and $\mathrm{E}^n(\varepsilon^0, \Omega^0)$ are of order n in θ. Note that ε depends on the ordinary ε^0 and also on Ω^0.

The Seiberg-Witten condition (15.27) states that the dependence of the noncommutative gauge field on the ordinary one is fixed by requiring that ordinary gauge variations of Ω^0 inside$\Omega(\Omega^0)$ produce the noncommutative gauge variation of Ω. This implies that once we expand, order by order in θ, the noncommutative action in terms of the commutative fields, the resulting action will be gauge invariant under ordinary local Lorentz transformations because the noncommutative action is invariant under the local noncommutative Lorentz transformations of Sect. 37.3. Following Ref. [2], up to first order in θ the solution reads:

$$\Omega = \Omega^0 + \frac{i}{4}\theta^{IJ}\{\Omega_I^0, \ell_J \Omega^0 + R_J^0\} \tag{15.31}$$

$$\varepsilon = \varepsilon^0 + \frac{i}{4}\theta^{IJ}\{\Omega_I^0, \ell_J \varepsilon^0\} \tag{15.32}$$

$$R = R^0 + \frac{i}{4}\theta^{IJ}\left(\{\Omega_I^0, (\ell_J + L_J)R^0\} - [R_I^0, R_J^0]\right) \tag{15.33}$$

$$\psi = \psi^0 + \frac{i}{4}\theta^{IJ}\Omega_I^0(\ell_J + L_J)\psi^0 \tag{15.34}$$

where Ω^0, R_A^0 are defined as the contraction along the tangent vector X_I of the exterior forms A $A\Omega^0$, R^0, i.e. $\Omega^0 \equiv i_I \Omega^0$, $R^0 \equiv i_I R^0$, (i_I being the contraction along X_I). We have also introduced the Lie derivative f_I along the vector field X_I, and the covariant Lie derivative L_I along the vector field X_I. L_I acts on R^0 and ψ^0 as $L_I R^0 = f_I R^0 - \Omega^0 \star R^0 + R^0 \star \quad \Omega^0$ and $L_I\, \psi^0 = f_I\, \psi^0 - \Omega^0\, \psi^0$. In *BI I* fact the covariant Lie derivative L_I has the Cartan form:

$$L_I = i_I D + D i_I$$

where D is the covariant derivative. We refer to [2] for higher order in θ expressions. All these formulae are *not SO*(1, 3)-gauge covariant, due to the presence of the "naked" connection Ω^0 and the non-covariant Lie derivative $f_I = i_I \mathrm{d} + \mathrm{d} i_I$. However, when inserted in the NC action the resulting action is gauge invariant order by order in θ. Indeed usual gauge variations induce the $\leftrightarrow$-gauge variations under which the NC action is invariant. Therefore the NC action, re-expressed in terms of ordinary fields via the SW map, is invariant under usual gauge transformations. Since these do not involve θ, the expanded action is invariant under ordinary gauge variations order by order in θ. Moreover the action, once re-expressed in terms of ordinary fields remains geometric, and hence invariant under diffeomorphisms. This is the case because the noncommutative action and the SW map are geometric, we indeed see that only coordinate independent operations like the contraction i_I and the Lie derivatives f_I and L_I appear in the SW map.

From (15.31) and (15.34) we also deduce

$$D\psi = D\psi^0 + \frac{i}{4}\theta^{IJ}\left(\Omega_I^0(\ell_J + L_J)D\psi^0 - 2R_I L_J \psi\right). \tag{15.35}$$

15.13 Action at First Order in θ

The expression of the gravity action, up to second order in θ, in terms of the commutative fields and of the first order fields (15.31)–(15.35) has been given in [2]. The action is gauge invariant even if the expression in [2] is not explicitly gauge invariant. We here present the explicit gauge invariant expression for the action up to first order in θ. We replace the noncommutative fields appearing in the action with their expansions (15.31)–(15.35) in commutative fields and integrate by parts in order to obtain an explicit $SO(3,1)$ gauge invariant action. We thus obtain the following gravity action coupled to spinors

$$
\begin{aligned}
S = \int & T\ r(iRVV\gamma_5) + \bar{\psi}V^3\gamma_5 D\psi + D\bar{\psi}V^3\gamma_5\psi \\
& + \frac{i}{4}\theta^{IJ}\left(\bar{\psi}\{V^3, r_{ij}\}\gamma_5 D\psi + D\bar{\psi}\{V^3, R_{IJ}\}\gamma_5\psi\right) \\
& + \frac{i}{2}\theta^{IJ}\Big(2L_I\bar{\psi}R_J V^3\gamma_5\psi - 2\bar{\psi}V^3 R_I\gamma_5 L_J\psi - L_1\bar{\psi}V^3\gamma_5 L_J D\psi \\
& - L_I D\bar{\psi}V^3\gamma_5 L_J\psi \\
& + \bar{\psi}(\{L_I V L_J V, V\} + L_I V\ \ V L_J V)\gamma_5 D\psi + D\bar{\psi}(\{L_I V\, L_J V, V\} \\
& + L_I V\ \ V L_J V)\gamma_5\psi\Big) + O(\theta^2)
\end{aligned}
\tag{15.36}
$$

where with obvious abuse of notation we have omitted the apex ^{0} denoting commutative fields, we also have omitted writing the wedge product, and $V^3 = V \wedge V \wedge V$.

15.14 Conclusions

We have constructed an extended Einstein gravity action that: (i) is explicitly invariant under local Lorentz transformations because expressed solely in terms of the gauge covariant operators L_I, i_I, D and fields R, V, ψ; (ii) is diffeomorphic invariant; (iii) has the same fields of classical gravity plus the noncommutative structure. This latter is given by the vector fields $\{X_I\}$ that choosing an appropriate kinetic term can become dynamical, the idea being that both spacetime curvature and noncommutativity should depend on matter distribution.

This extended action has been obtained from considering gravity on noncommutative spacetime, that as argued in the introduction is a very natural assumption at high energies, like those close to the inflationary epoch.

Acknowledgments We acknowledge the fruitful and pleasant atmosphere and the perfect organization enjoyed during the Udine Symposium.

15.15 A. Gamma Matrices in $D = 4$

We summarize in this Appendix our gamma matrix conventions in $D = 4$.

$$\eta_{ab} = (1, -1, -1, -1), \{\gamma_a, \gamma_b\} = 2\eta_{ab}, \quad [\gamma_a, \gamma_b] = 2\gamma_{ab}, \tag{15.37}$$

$$\gamma_5 \equiv i\gamma_0\gamma_1\gamma_2\gamma_3, \quad \gamma_5\gamma_5 = 1, \quad \varepsilon_{0123} = -\varepsilon^{0123} = 1, \tag{15.38}$$

$$\gamma_a^\dagger = \gamma_0\gamma_a\gamma_0, \quad \gamma_5^\dagger = \gamma_5 \tag{15.39}$$

$$\gamma_a^T = -C_{\gamma a}C^{-1}, \quad \gamma_5^T = C_{\gamma 5}C^{-1}, \quad C^2 = -1, \quad C^\dagger = C^T = -C \tag{15.40}$$

References

1. Aschieri P, Castellani L (2009a) Noncommutative D = 4 gravity coupled to fermions. JHEP 0906 086 [arXiv:0902.3817 [hep-th]]
2. Aschieri P, Castellani L (2011) Noncommutative gravity coupled to fermions: second order expansion via Seiberg-Witten map, arXiv:1111.4822 [hep-th]
3. Lizzi F, Mangano G, Miele G, Peloso M (2002) Cosmological perturbations and short distance physics from noncommutative geometry. JHEP 0206 049 [hep-th/0203099]
4. Aschieri P, Dimitrijevic M, Kulish P, Lizzi F, Wess J (2009b), Noncommutative spacetimes. Lecture notes in physics, vol 774, Springer
5. Aschieri P, Castellani L (2012a) Extended gravity from dynamical noncommutativity arXiv:1206.4096 [hep-th]
6. Aschieri P, Castellani L (2012b) Noncommutative gauge fields coupled to noncommutative gravity, arXiv:1205.1911 [hep-th]
7. Faddeev LD, Reshetikhin NYu, Takhtajan LA (1990) Quantization of lie groups and lie algebras. Algebra i Anal 1(1):178 (Leningrad Math. J. 1 193)
8. Manin Y (1988) Quantum groups and non-commutative geometry. Preprint Montreal University, Montreal, CRM-1561
9. Connes A (1994) Non-commutative geometry. Academic Press, San Diego, p 661 ISBN 0-12-185860-X
10. Kontsevich M (2003) Deformation quantization of poisson manifolds. Lett Math Phys 66: 157 [arXiv:q-alg/9709040 [q-alg]]
11. Aschieri P, Dimitrijevic' M, Meyer F, Wess J (2006) Noncommutative geometry and gravity. Class Quant Grav 23:1883–1912 [hep-th/0510059]
12. Chamseddine AH (2004), Sl(2, C) gravity with complex vierbein and its noncommutative extension. Phys Rev D69:024015 [arXiv:hep-th/0309166]
13. Ulker K, Yapiskan B (2008), Seiberg-witten maps to all orders. Phys Rev D77:065006[arXiv:0712.0506 [hep-th]]

Chapter 16
Quantum Gravity: A Heretical Vision

John Stachel

Abstract The goal of this work is to contribute to the development of a background-independent, non-perturbative approach to quantization of the gravitational field based on the conformal and projective structures of space-time. But first I attempt to dissipate some mystifications about the meaning of quantization, and foster an ecumenical, non-competitive approach to the problem of quantum gravity (QG), stressing the search for relations between different approaches in any overlapping regions of validity. Then I discuss some topics for further research based on the approach we call unimodular conformal and projective relativity (UCPR).

16.1 Only Theories

Perhaps it will be helpful if I recall a tripartite classification of theories that I proposed many years ago. The three categories are:

(1) *Perfectly perfect theories*: The range of these theories includes the entire universe: There is nothing In the world that these theories do not purport to explain, and they correctly explain all these phenomena. Today we call such theories TOEs— Theories of Everything.
(2) *Perfect theories*: These are more modest. They correctly explain all phenomena within their range of application, but there are phenomena that they do not purport to explain.
(3) Then there are just plain *Theories*: There are phenomena that they do not purport to explain, and there are phenomena that they do purport to explain, but do not explain correctly.

Both the history of science and my own experience have taught me that all we have now, ever have had in the past, or can hope to have in the future are just plain theories. This tale had two morals:

J. Stachel (✉)
Center for Einstein Studies, Boston University, Boston, MA, USA

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_16,

(1) Every theory has its range of validity and its limits; to understand a theory better we must find its limits. In this sense, we understand Newtonian gravity better than general relativity (GR).
(2) There will be theories with over-lapping ranges of validity; to understand each of these theories better we must explore the relations between them in the overlap regions. Some examples will be given in the next section.

16.2 What Is Quantum Theory? What Quantization Is and Is Not

A certain mystique surrounds the words "quantum theory." The very words conjure up visions of probing the depths of reality, exploring the paradoxical properties of the exotic building blocks of the universe: fundamental particles, dark matter, dark energy—dark thoughts.

> But the scope of the quantum mechanical formalism is by no means limited to such (presumed) fundamental particles. There is no restriction of principle on its application to any physical system. One could apply the formalism to sewing machines if there were any reason to do so! [17].

Then what is quantization? Quantization is just a way of accounting for the effects of the existence of *h*, the quantum of action, on any process undergone by some system–or rather on some theoretical model of such a system. This is the case whether the system to be quantized is assumed to be "fundamental" or "composite." That is, whether the model describes some (presumed) fundamental entities, or whether it describes the collective behavior of an ensemble of such entities.

> [T]he universal quantum of action … was discovered by Max Planck in the first year of this [20th] century and came to inaugurate a whole new epoch in physics and natural philosophy. We came to understand that the ordinary laws of physics, i.e., classical mechanics and electrodynamics, are idealizations that can only be applied in the analysis of phenomena in which the action involved at every stage is so large compared to the quantum that the latter can be completely disregarded [4].

We all know examples of the quantization of fundamental systems, such as electrons, quarks, neutrinos, etc.; so I shall just remind you of some examples of *non-fundamental quanta*, such as *quasi-particles*: particle-like entities arising in certain systems of interacting particles, e.g., phonons and rotons in hydrodynamics (see, e.g., [13]); and *phenomenological field quanta*, e.g., quantized electromagnetic waves in a homogeneous, isotropic medium (see, e.g., [11]).

So, successful quantization of some classical formalism does *not necessarily* mean that one has achieved a deeper understanding of reality—or better, an understanding of a deeper level of reality. What is does mean is that one has successfully understood the effects of the quantum of action on the phenomena (Bohr's favorite word), or processes (Feynman's favorite) described by the formalism being quantized.

Having passed beyond the quantum mystique, one is free to explore how to apply quantization techniques to various formulations of a theory without the need to single one out as the unique "right" one. One might say, with Jesus: "In my Father's house are many mansions" (John 14:2); or with Chairman Mao (in his more tolerant moments): "Let a hundred flowers blossom, let a hundred schools contend."

Three Morals of This Tale:

1. *Look for relations between quantizations*: If two such quantizations at *different*levels are carried out, one may then investigate the relation between them. *Example*: [7] has investigated the relation between microscopic and macroscopic quantizations of the electromagnetic field in a dielectric.
   If two such quantizations at the *same* level exist, one may investigate the relation between them. *Example:* [2] studied the relation between loop quantization and the usual field quantization of the electromagnetic field: If you "thicken" the loops, the two are equivalent.
2. *Don't Go "Fundamental"*: The search for a method of quantizing space-time structures associated with the Einstein equations is distinct from the search for an underlying theory of all "fundamental" interactions.

   > I see no reason why a quantum theory of gravity should not be sought within a standard interpretation of quantum mechanics (whatever one prefers). … We can consistently use the Copenhagen interpretation to describe the interaction between a macroscopic classical apparatus and a quantum-gravitational phenomenon happening, say, in a small region of (macroscopic) spacetime. The fact that the notion of spacetime breaks down at short scale within this region does not prevent us from having the region interacting with an external Copenhagen observer ([14], p. 370).

3. *Don't go "Exclusive"*: Any attempt, such as ours (see [5, 19]), to quantize the conformal and projective structures does not negate, and need not replace, attempts to quantize other space-time structures. Everything depends on the utility of the results of formal quantization in explaining some physical processes depending on the quantum of action.
   One should not look at different approaches to QG as "*either-or*" alternatives, but "*both-and*" supplements. The question to ask is not: "Which is right and which is wrong?" but: "In their regions of overlapping validity, what is the relation between these different models of quantization of some gravitational phenomena?".

16.3 Measurability Analysis

A physical theory consists of more than a class of mathematical models. Certain mathematical structures within these models must be singled out as corresponding to physically significant concepts. And these concepts must be in principle measurable. This is *not* operationalism: What is measurable is real. Rather, it is the opposite: What is real must be measurable by some idealized physical procedure that is consistent with the theory. This test of the physical validity of a theory is called *measurability analysis*

> Measurability analysis identifies those dynamic field variables that are susceptible to observation and measurement (“observables”), and investigates to what extent limitations inherent in their experimental determination are consistent with the uncertainties predicted by the formal theory [3].

16.4 Process is Primary, States are Secondary

I cannot put this point better than Lee Smolin has done:

> [R]elativity theory and quantum theory each …tell us– no, better, they scream at us- that our world is a history of processes. Motion and change are primary. Nothing is, except in a very approximate and temporary sense. How something is, or what its state is, is an illusion. … So to speak the language of the new physics we must learn a vocabulary in which process is more important than, and prior to, stasis [15].

Carlo Rovelli has helped us to develop that vocabulary for QG:

> The data from a local experiment (measurements, preparation, or just assumptions) must in fact refer to the state of the system on the entire boundary of a finite spacetime region. The field theoretical space …is therefore the space of surfaces Σ [a three-dimensional hypersurface bounding a finite four-dimensional spacetime region] and field configurations φ on Σ. Quantum dynamics can be expressed in terms of an [probability] amplitude $W[\Sigma, \varphi]$ [for some process].

Background dependence versus background independence:

> Notice that the dependence of $W[\Sigma, \varphi]$ on the geometry of Σ codes the spacetime position of the measuring apparatus. In fact, the relative position of the components of the apparatus is determined by their physical distance and the physical time elapsed between measurements, and these data are contained in the metric of Σ. Consider now a background independent theory. Diffeomorphism invariance implies immediately that $W[\Sigma, \varphi]$ is independent of Σ … Therefore in gravity W depends only on the boundary value of the fields. However, the fields include the gravitational field, and the gravitational field determines the spacetime geometry. Therefore the dependence of W on the fields is still sufficient to code the relative distance and time separation of the components of the measuring apparatus! ([14], p. 23).

16.5 Poisson Brackets Versus Peierls Brackets

One central method of taking into account the quantum of action is by means of introducing commutation relations between various particle or field quantities entering into the classical formalism. These commutation relations have more than a purely formal significance

> We share the point of view emphasized by Heisenberg and Bohr and Rosenfeld, that the limits of definability of a quantity within any formalism should coincide with the limits of measurability of that quantity for all conceivable (ideal) measurement procedures. For well-established theories, this criterion can be tested. For example, in spite of a serious challenge, source-free quantum electro-dynamics was shown to pass this test. In the case of quantum

> gravity, our situation is rather the opposite. In the absence of a fully accepted, rigorous theory, exploration of the limits of measurability of various quantities can serve as a tool to provide clues in the search for such a theory: If we are fairly certain of the results of our measurability analysis, the proposed theory must be fully consistent with these results ([1]).

It follows that one should replace canonical methods, based on the primacy of states, by some covariant method, based on the primacy of processes. As Bryce DeWitt emphasizes, Peierls found the way to do this:

> When expounding the fundamentals of quantum field theory physicists almost universally fail to apply the lessons that relativity theory taught them early in the twentieth century. Although they usually carry out their calculations in a covariant way, in deriving their calculational rules they seem unable to wean themselves from canonical methods and Hamiltonians, which are holdovers from the nineteenth century, and are tied to the cumbersome (3 + 1)-dimensional baggage of conjugate momenta, bigger-than-physical Hilbert spaces and constraints. One of the unfor-tunate results is that physicists, over the years, have almost totally neglected the beautiful covariant replacement for the canonical Poisson bracket that Peierls invented in 1952 ([8], Preface, p. *v*; see also Sect. 16.5, "The Peierls Bracket").

16.6 What Is Classical General Relativity?

GR is often presented as if there were only *one* primary space-time structure: the pseudo-Riemannian metric tensor g. Once one realizes that GR is based on *two* distinct space-time structures, the *chrono-geometry* (metric g) and the *inertio-gravitational field* (affine connection Γ), and the *compatibility conditions* between the two ($Dg = 0$), the question arises: What structure(s) shall we quantize and how?

Usually, it is taken for granted that all the space-time structures must be simultaneously quantized. Traditionally, one attempts to quantize the chrono-geometry, or some canonical (3 + 1) version of it, such as the first fundamental form of a Cauchy hypersurface; and introduces the inertia-gravitational field, again in canonical version as the second fundamental form of the hypersurface, disguised as the momenta conjugate to the first fundamental form (see, e.g., [20], pp. 160–170). More recently, the inverse approach has had great success in loop QG: One starts from a (3 + 1) breakup of the affine connection that makes it analogous to a Yang-Mills field, and introduces some (3 + 1) version of the metric as the momenta conjugate to this connection (see, e.g., [14]).

Both approaches have one feature in common: the (3 + 1) canonical approach adopted naturally favors states over processes, leading to a number of problems. In particular, the state variables (the "positions") are primary; their time derivatives (the "momenta") are secondary.

However, there is no need to adopt a canonical approach to GR, nor to initially conflate the two structures g and Γ. From the point of view of a first-order Palatini-type variational principle, *the compatibility conditions* between the two are just one of the two sets of dynamical field equations derived from the Lagrangian, linking g and Γ, which are initially taken to be independent of each other. The other set of field equations, of course, links the trace of the affine curvature tensor, the affine Ricci tensor, to the non-gravitational sources of the inertio-gravitational field. There is a sort of electromagnetic analogy: In the first order formalism, $\boldsymbol{G}^{\mu\nu}$ and $F_{\mu\nu}$ (or A_μ) are initially independent fields, which are then made compatible by the constitutive relations [16].

Both the canonical approach and the first-order Palatini-type approach take it for granted that the compatibility conditions must be preserved exactly, whether from the start or as a result of the field equations. As we shall see, in UCPR this is no longer the case.

16.7 The Newtonian Limit, Multipole Expansion of Gravitational Radiation

The remarkable accuracy of the Newtonian approximation for the description of so many physical systems suggest that the Newtonian limit of GR might provide a convenient starting point for a discussion of quantization of the gravitational field. In the version of Newtonian theory that takes into account the equivalence principle (see [18]), the chronometry (universal time) and the geometry (Euclidean in each of the preferred frame of reference picked out by the symmetry group, i.e., all frames of reference that are rotation-free, but linearly accelerated with respect to each other) are absolute, i.e., fixed background structures; while the inertia-gravitational field is dynamical and related by field equations relating the affine Ricci tensor to the sources of the field. The compatibility conditions between connection and chronometry and geometry allow just sufficient freedom to introduce a dynamical gravitational field. Thus, the quantum theory *must* proceed by quantization of the connection while leaving the chronometry and geometry *fixed* (see [6]).

This suggests the possibility of connecting the Newtonian near field and the far radiation field by the method of matched asymptotic expansions. Kip Thorne explained this approach:

> Previous work on gravitational-wave theory has not distinguished the local wave zone from the distant wave zone. I think it is useful to make this distinction, and to split the theory of gravitational waves Into two corresponding parts: Part one deals with the source's generation of the waves, and with their propagation into the local wave zone; thus it deals with … all of spacetime except the distant wave zone. Part two deals with the propagation of the waves from the local wave zone out through the distant wave zone to the observer … The two parts, wave generation and wave propagation, overlap in the local wave zone; and the two theories can be matched together there. … [F]or almost all realistic situations, wave propagation theory can do its job admrably well using the elementary formalism of geometric optics ([21], p. 316).

If one looks at this carefully, there are really three zones:

(1) Near zone, where the field is generated by the source.
(2) Intermediate zone, where the transition takes place between zones (1) and (3).

(3) Far zone, where the pure radiation field has broken free from the source.

But before proceeding any further with the discussion of quantization in this Newtonian limit, it will be helpful first to discuss UCPR.

16.8 Unimodular Conformal and Projective Relativity

Einstein was by no means wedded to general covariance when he started his search for a generalized theory of relativity that would include gravitation. The equivalence principle:

> made it not only probable that the laws of nature must be invariant with respect to a more general group of transformations than the Lorentz group (extension of the principle of relativity), but also that this extension would lead to a more profound theory of the gravitational field. That this idea was correct in principle I never doubted in the least. But the difficulties in carrying it out seemed almost insuperable. First of all, elementary arguments showed that the transition to a wider group of transformations is incompatible with a direct physical interpretation of the space-time coordinates, which had paved the way for the special theory of relativity. Further, at the outset it was not clear how the enlarged group was to be chosen [10].

He actually considered restricting the group of transformations to those that preserved the condition that the determinant of the metric be equal to -1, both when formulating GR and when investigating whether the theory could shed light on the structure of matter (see Einstein [9], the translation of his 1919 paper). So the choice of SL(4, R) as the preferred invariance group is actually in the spirit of Einstein's original work (see Stachel [19]).

I suspect that the restriction to such *unimodular diffeomorphisms*, which guarantees the existence of a volume structure, may be the remnant, at the continuum level, of a discrete quantization of four-volumes, which would form the fundamental space-time units, as in causal set theory. Quantization of three-volumes, etc., would be "perspectival" effects, dependent on the (3 + 1) breakup chosen for space-time. The fact that one can impose the unimodularity condition prior to, and independently of, any consideration of the conformal or projective structures lends some credence to this speculation.

If we confine ourselves to *unimodular diffeomorphisms*, we can easily go from compatible metric and connection to compatible conformal and projective structures. Many of the questions discussed above must then be reconsidered in this somewhat different light. One will now have to take into account both the conformal and projective connections and their compatibility conditions; and the conformal and projective curvature tensors.

Now we are ready to return to the Newtonian limit, and propose a *conjecture*:

In zone (1), the projective structure dominates; the field equations connect it with the sources of the field. In zone (3), the conformal structure dominates; the radiation field obeys Huygens' principle (see the next section). In zone (2), the compatibility conditions between the conformal and projective structures dominate, assuring that the fields of zones (1) and (3) describe the same field.

In order to verify these conjectures, we shall have to find the answers to the following *questions*: How do the field equations look in the near zone? Which projective curvature tensor is related to the sources in the near zone? In the far zone, which conformal curvature tensor obeys Huygens' principle? In the intermediate zone, which conformal and projective connections/curvatures should be made compatible?

16.9 Zero Rest Mass Radiation Fields, Huygens' Principle, and Conformal Structure

The name "Huygens' Principle" is given to several versions (see, e.g., [12]), but I shall consider only one. Let $u(x)$ be a function obeying some hyperbolic field equation on an n-dimensional differentiable manifold V_n, with a pseudo-Riemannian metric. As Hadamard showed, if the Cauchy problem is well-posed on some initial space-like hypersurface *S,* the solution at any future point $x_0 \in \mathrm{V}_n$, depends on some set of initial data given on the boundary and in the interior of the intersection of the retrograde characteristic conoid $u(x_0)$ with the initial surface *S*. If, for every Cauchy problem on any *S* and every x_0, the solution depends *only* on *the initial data on the boundary*, the equation is said to satisfy *Huygens' principle*.

Its importance for our purposes lies in the fact that, only if Huygens' principle holds for a solution to the field equations of massless fields, such as the electromagnetic and the gravitational, does geometrical optics, i.e., the null-ray representation of the field, make sense. In that case one may carry out the analysis of the radiation field in terms of the shear tensor of a congruence of null rays, the components of the conformal curvature tensor projected onto these rays, etc. Similarly, ideal measurement of these quantities become possible; for example, the shear by means of two screens: one with a circular hole and one behind it to register the distortion of the shadow cast by the first screen.

In an arbitrary space-time, whether it is a fixed background chrono-geometry or one that is interacting with the Maxwell field, solutions to either the empty space Maxwell or Einstein-Maxwell equations, respectively, do *not* obey Huygens' principle. However, in a conformally flat space-time they do; and the interacting Einstein-Maxwell plane wave metric, which is type *N* in the Pirani-Petrov classification (see, e.g.[20]), also does. And in all such cases, the conformal structure is all that is needed to carry out the conceptual analysis and the corresponding ideal measurements

I assume that asymptotically "free," locally plane-wave solutions to the Einstein-Maxwell equations that are regular at past or full null infinity (Penrose's scri-minus and scri-plus) do obey the Huygens condition. In addition to the above considerations, this condition is also necessary for an analysis of scattering in terms of the probability amplitude :<*incoming free wave* | *outgoing scattered free wave*> to be valid.

If these assumptions are correct, then the free radiation field can be analyzed and presumably quantized entirely in terms of the conformal structure. However, all of these assumptions must of course be carefully checked.

16.10 Zero Rest-Mass Near Fields and Projective Structure

Local massless fields, still tied to the sources, do not obey Huygens' principle, and hence cannot be so analyzed. However, the gravitational analogue of the Bohr-Rosenfeld method of measuring electromagnetic field averages over four-dimensional volumes should still hold in this case. In UCPR, four-volumes are invariantly defined independently of any other space-time structures. If we want the four volume elements to be parallel (i.e., independent of path), we introduce a one form related to the gradient of the four-volume field and require this to be the trace of the still unspecified affine connection. So we are still left with full freedom to choose the conformal and projective structures [5].

The so-called equation of "geodesic deviation" (it should really be called "autoparallel deviation" since it involves the affine connection) will ultimately govern this type of analysis. And if we abstract from the parameterization of the curves, the projective structure should govern the resulting equations for the autoparallel paths. And in terms of amplitudes connecting asymptotic in- and out-states, one would expect that projective infinity will take the place of conformal infinity. Again, these expectations, and their implications for quantization of the near fields and their sources must be carefully investigated.

For further details on many points, see the paper by Kaća Bradonjić, "Unimodular Conformal and Projective Relativity: an Illustrated Introduction," in this volume.

References

1. Amelino-Camelia G, Stachel J (2009) Measurement of the space-time interval between two events using the retarded and advanced times of each event with respect to a time-like world-line. GRG J Gen Relat Gravit 41:1107–1124
2. Ashtekar A, Rovelli C (1992) A loop representation for the Quantum Maxwell field. Class Quan Gravit 9:1121–1150
3. Bergmann P, Smith G (1982) Measurability analysis of the linearized gravitational field. Gen Relat Gravit 14:1131–1166
4. Bohr N (1957) Atoms and human knowledge. A public lecture delivered on December 13, Holmberg Hall, Norman, Oklahoma. http://www.nhn.ou.edu/Bohr-lecture-OU-1957.shtml
5. Bradonjić K, Stachel J (2012) Unimodular conformal and projective relativity. Euro Phys Lett 97:10001
6. Christian J (1997) Exactly soluble sector of quantum gravity, Phys Rev D56: 4844–4877
7. Crenshaw M, Bowden C (2002) On quantization of the fields in dielectrics. J Modern Optics 49:511–517
8. DeWitt B (2004) The global approach to quantum field theory, vol 1. Clarendon Press, Oxford

9. Einstein A (1952) Do gravitational fields play an essential part in the structure of the elementary particles of matter. The principle of relativity, Dover Publications, pp 191–198.
10. Einstein A (1956) Autobiographische Skizze. In: C. Seelig (ed) Helle Zeit-Dunkle Zeit, Euro Verlag, Zürich/Stuttgart/Wien, pp 9–17.
11. Ginzburg V (1989) Applications of electrodynamics in theoretical physics and astrophysics, Chapter 7. The Cerenkov and Doppler effects, Gordon and Breach, New York, pp 123–158
12. Hadamard J (1952) Lectures on Cauchy's problem in linear partial differential equations. Dover, New York
13. Mehra J, Rechenberg H (2001) The historical development of quantum theory, Superconductivity, Superfluidity, and Further Progress in Condensed Matter Physics (Sect. 2.3), Part 2, vol 6. Springer, New York 1947–1974:1159–1184
14. Rovelli C (2004) Quantum Gravity. Cambridge University Press, Cambridge, UK
15. Smolin L (2000) Three roads to quantum gravity. Oxford University Press, Oxford, UK
16. Stachel J (1984) The generally covariant form of Maxwell's equations. In: Berger MS (ed) J.C. Maxwell: the sesquicentennial symposium/ new vistas in mathematics, science and technology, North- Holland, 23–37.
17. Stachel J (1986) Do quanta need a new logic?. University of Pittsburgh Press, Pittsburgh, Philosophical problems of modern physics, From quarks to quasars, pp 349–385
18. Stachel J (2006) Einstein's intuition and the post-Newtonian approximation. In Topics in mathematical physics, general relativity and cosmology in honor of Jerzy Plebanski, World Scientific, Singapore, pp 453–467
19. Stachel J (2011) Conformal and projective structures in general relativity. Gen Relat Gravit 43:3399–4309
20. Stephani H (1990) General relativity: an introduction to the theory of the gravitational field, 2nd edn. Cambridge University Press, Cambridge, UK
21. Thorne K (1980) Multipole expansions of gravitational radiation. Rev. Mod. Phys. 52:299–339

Chapter 17
From Clock Synchronization to Dark Matter as a Relativistic Inertial Effect

Luca Lusanna

Abstract Clock synchronization leads to the definition of instantaneous 3-spaces (to be used as Cauchy surfaces) in non-inertial frames, the only ones allowed by the equivalence principle. ADM canonical tetrad gravity in asymptotically Minkowskian space-times can be described in this framework. This allows to find the York canonical basis in which the inertial (gauge) and tidal (physical) degrees of freedom of the gravitational field can be identified. A Post-Minkowskian linearization with respect to the asymptotic Minkowski metric (asymptotic background) allows to solve the Dirac constraints in non-harmonic 3-orthogonal gauges and to find non-harmonic TT gravitational waves. The inertial gauge variable York time (the trace of the extrinsic curvature of the 3-space) describes the general relativistic freedom in clock synchronization. After a digression on the gauge problem in general relativity, it is shown that dark matter, whose experimental signatures are the rotation curves and the mass of galaxies, may be described (at least partially) as an inertial relativistic effect (absent in Newton gravity) connected with the York time.

17.1 Introduction

The theory of global non-inertial frames in special relativity (SR) developed in Ref. [2] can be used in the family of globally hyperbolic, asymptotically Minkowskian space-times without super-translations (Beig and Ó Murchadha, 1997) in the framework of Einstein general relativity (GR), where also the space-time becomes dynamical with the 4-metric being determined (modulo the gauge freedom of 4-diffeomorphisms) by Einstein's equations [9]. The use of the ADM action allows to get the Hamiltonian formulation of GR in these space-times, which was studied in detail in Refs. [5, 7, 10] together with its extension to ADM tetrad gravity (needed

L. Lusanna (✉)
Sezione INFN di Firenze, Via Sansone 1, 50019 Sesto Fiorentino, Italy
e-mail: lusanna@fi.infn.it

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_17,

for the inclusion of fermions) by expressing the 4-metric in terms of tetrad fields. See Ref. [8] for a complete review of this framework both in SR and GR.

The replacement of clock synchronization with an admissible 3 + 1 splitting (a global non-inertial frame, the only one existing in the large in GR due to the equivalence principle) allows to introduce radar 4-coordinates $\sigma^A = (\tau; \sigma^u)$ centered on an arbitrary time-like observer endowed with an atomic clock. Without super-translations the allowed asymptotic symmetries form the ADM asymptotic Poincaré group, which reduces to the special relativistic Poincaré group of the existing matter when the Newton constant is switched off ($G = 0$). This allows to recover all the results of the standard model of elementary particles, at whose heart there are the representations of the Poincaré group in inertial frames of Minkowski space-time. Instead in the spatially compact without boundary space-times used in loop quantum gravity there is no trace of a realization of the Poincaré group. In absence of super-translations the non-Euclidean 3-spaces tend to an asymptotic inertial rest frame of the 3-universe at spatial infinity (where they are orthogonal to the conserved ADM 4-momentum [7]) in a direction-independent way. In this *non-inertial rest frame* there are asymptotic inertial observers whose spatial axes can be identified by means of the fixed stars of star catalogues. The 4-metric tends to an asymptotic Minkowski metric, to be used as an *asymptotic background metric* avoiding its splitting in a background plus perturbations like in the linearization leading to gravitational waves. In these space-times the canonical Hamiltonian is the ADM energy: therefore there is not a frozen picture like in the space-times used in loop quantum gravity.

17.2 ADM Tetrad Gravity and the York Canonical Basis: Identification of the Inertial and Tidal Variables and the Hamiltonian Post-Minkowskian Linearization

As shown in Ref. [10] in ADM tetrad gravity the configuration variables are cotetrads, which are connected to cotetrads adapted to the 3 + 1 splitting of space-time (so that the adapted time-like tetrad is the unit normal to the 3-space Σ_τ) by standard Wigner boosts for time-like vectors of parameters $\varphi_{(a)}(\tau, \sigma^r)$, $a = 1, 2, 3$: $E^{(\alpha)}_A = L^{(\alpha)}{}_{(\beta)}(\varphi_{(a)})\, \stackrel{o}{E}{}^{(\beta)}_A$. The adapted cotetrads have the following expression in terms of cotriads ${}^3e_{(a)r}$ on Σ_τ and of the lapse $N = 1+n$ and shift $n_{(a)} = N^r\, {}^3e_{(a)r}$ functions: $\stackrel{o}{E}{}^{(o)}_\tau = 1 + n$, $\stackrel{o}{E}{}^{(o)}_r = 0$, $\stackrel{o}{E}{}^{(a)}_\tau = n_{(a)}$, $\stackrel{o}{E}{}^{(a)}_r = {}^3e_{(a)r}$. The 4-metric becomes ${}^4g_{\tau\tau} = \epsilon\,[(1+n)^2 - \sum_a n^2_{(a)}]$, ${}^4g_{\tau r} = -\epsilon \sum_a n_{(a)}\, {}^3e_{(a)r}$, ${}^4g_{rs} = -\epsilon\, {}^3g_{rs} = -\epsilon \sum_a {}^3e_{(a)r}\, {}^3e_{(a)s}$. The 16 configurational variables in the ADM action are $\varphi_{(a)}$, $1+n$, $n_{(a)}$, ${}^3e_{(a)r}$. There are ten primary constraints (the vanishing of the 7 momenta of boosts, lapse and shift variables plus three constraints describing the rotation on the flat indices (a) of the cotriads) and four secondary ones (the super-Hamiltonian and super-momentum constraints): all of them are first class in the phase space spanned by 16 + 16 fields. This implies that there are 14 gauge variables describing *inertial* *effects* and 2 canonical pairs of physical degrees of freedom describing the *tidal effects* of the gravitational field (namely gravitational waves in the weak field limit). In this canonical basis only the momenta ${}^3\pi^r_{(a)}$ conjugated to the cotriads are not vanishing.

Then in Ref. [1] we have found a canonical transformation to a canonical basis adapted to ten of the first class constraints (not to the super-Hamiltonian and super-momentum ones), implementing the York map of Ref. [6] and diagonalizing the York-Lichnerowicz approach. In this York canonical basis 6 configuration variables are the 3 boosts $\varphi_{(a)}(\tau, \sigma^r)$ and the 3 angles $\alpha_{(a)}(\tau, \sigma^r)$ parametrizing the O(3,1) gauge freedom of tetrads (the gauge freedom for each observer to choose three gyroscopes as spatial axes and to choose the law for their transport along the world-line) with vanishing conjugate momenta. Other 4 gauge configuration variables are suitable lapse and shift functions $1+n(\tau, \sigma^r)$ and $\bar{n}_{(a)}(\tau, \sigma^r)$ with vanishing momenta. Three gauge configuration angles $\theta^i(\tau, \sigma^r)$ (i.e. the director cosines of the tangents to the three coordinate lines in each point of Σ_τ) describe the freedom in the choice of the 3-coordinates σ^r on each 3-space: their fixation implies the determination of the shift gauge variables $\bar{n}_{(a)}(\tau, \sigma^r)$, namely the appearances of gravito-magnetism in the chosen 3-coordinate system. Their conjugate momenta $\pi_i^{(\theta)}(\tau, \sigma^r)$ are the unknowns in the super-momentum constraints. The configuration variable $\tilde{\phi}(\tau, \sigma^r) = \sqrt{det\, {}^3g_{rs}(\tau, \sigma^r)}$ describes the 3-volume element and is the unknown in the super-Hamiltonian constraint. The final basic gauge variable is a momentum, namely the trace ${}^3K(\tau, \sigma^r)$ of the extrinsic curvature (also named the *York time*) of the non-Euclidean 3-space Σ_τ. The Lorentz signature of space-time implies that ${}^3K$ is a momentum variable: it is a time coordinate, while θ^i are spatial coordinates. Differently from SR ${}^3K$ is *an independent inertial gauge variable describing the remnant in GR of the freedom in clock synchronization*! The other components of the extrinsic curvature are dynamically determined. This gauge variable has no Newtonian counterpart (there Euclidean 3-space is absolute), because its fixation determines the final shape of the non-Euclidean 3-space. Moreover this gauge variable gives rise to a negative kinetic term in the weak ADM energy $\hat{E}_{ADM}$, vanishing only in the gauges ${}^3K(\tau, \vec{\sigma}) = 0$. The *tidal effects*, i.e. the physical degrees of freedom of the gravitational field, are described by the two canonical pairs $R_{\bar{a}}(\tau, \sigma^r)$, $\Pi_{\bar{a}}(\tau, \sigma^r)$, $\bar{a} = 1, 2$. The 4-metric is a function of n, $\bar{n}_{(a)}$, θ^i, $\tilde{\phi}$, $R_{\bar{a}}$.

In the York canonical basis the Hamilton equations generated by the Dirac Hamiltonian $H_D = \hat{E}_{ADM} + (constraints)$ are divided in four groups: (A) the contracted Bianchi identities, namely the evolution equations for $\tilde{\phi}$ and $\pi_i^{(\theta)}$ (they say that given a solution of the constraints on a Cauchy surface, it remains a solution also at later times); (B) the evolution equation for the four basic gauge variables θ^i and ${}^3K$: these equations determine the lapse and the shift functions once the basic gauge variables are fixed; (C) the evolution equations for the tidal variables $R_{\bar{a}}$, $\Pi_{\bar{a}}$; (D) the Hamilton equations for matter, when present. Once a gauge is completely fixed, the Hamilton equations become deterministic.

In the first paper of Ref.(Alba and Lusanna, 2011), we studied the coupling of N charged scalar particles plus the electro-magnetic field to ADM tetrad gravity in the York canonical basis. To regularize the self-energies both the electric charge and the sign of the energy of the particles are Grassmann-valued. We give the restriction of the Hamilton equations and of the constraints to the family of *non-harmonic 3-orthogonal* Schwinger time gauges, in which the 3-metric is diagonal. This family of gauges is determined by the gauge fixings $\varphi_{(a)}(\tau,\sigma^r) = \alpha_{(a)}(\tau,\sigma^r) = \theta^i(\tau,\sigma^r) \approx 0$ and ${}^3K(\tau,\sigma^r) \approx (arbitrary\, numerical\, function)$. Then in the second paper we defined a consistent *linearization* of ADM canonical tetrad gravity plus matter in the weak field approximation (with a ultraviolet cutoff on matter) to obtain a formulation of Hamiltonian Post-Minkowskian (PM) gravity with non-flat Riemannian 3-spaces and asymptotic Minkowski background. We can find the linearized 4-metric without Post-Newtonian (PN) approximations and all the relevant properties of gravitational waves can be recovered in these non-harmonic gauges. The non-Euclidean 3-spaces have a first order extrinsic curvature (with ${}^3K_{(1)}(\tau,\sigma^r)$ describing the clock synchronization convention) and a first order modification of Minkowski light-cone. The equations of motion for the particles and their PM limit are consistent with the equality of gravitational and inertial masses.

17.3 Post-Newtonian Particle Equations, Dark Matter as a Relativistic Inertial Effect Due to the York Time and the Gauge Problem in GR

In the third paper of Ref.(Alba and Lusanna, 2011) we disregarded electro-magnetism and we studied the non-relativistic PN limit of the PM particle equations. At the lowest order we recovered Newton gravitational forces. The 1PN forces together with the limit of the ADM Poincaré generators reproduce the known results about binaries.

The new result is the presence of 0.5PN inertial forces depending on the non-local function ${}^3\tilde{K}_{(1)}(\tau,\sigma^r) = \frac{1}{\triangle}\,{}^3K_{(1)}(\tau,\sigma^r)$ of the inertial gauge variable York time. This term in the non-local York time can be *re-interpreted* as the introduction of an *effective (time-, velocity- and position-dependent) inertial mass term* for the kinetic energy of each particle: $m_i \mapsto m_i\,\Big(1+\frac{1}{c}\,\frac{d}{dt}\,{}^3\tilde{\mathcal{K}}_{(1)}(t,\vec{\tilde{\eta}}_i(t))\Big) = m_i + \delta\, m_i({}^3\tilde{K}_{(1)})$ in each non-Euclidean 3-space, depending on its shape as a 3-sub-manifold of space-time. *It is the equality of the inertial and gravitational masses of Newtonian gravity to be violated due to the non-Euclidean nature of the 3-spaces implied by Einstein equivalence principle*. This opens the possibility to describe *dark matter as a relativistic inertial effect* implying that the effective inertial mass of particles in the 3-spaces is bigger of the gravitational mass because it depends on the non-local York time, namely the quantity $\delta\, m_i({}^3\tilde{K}_{(1)})$ can be interpreted as the mass of dark matter. As shown in the third paper of Ref.(Alba and Lusanna, 2011) the three main signatures of dark matter (the rotatin curves of galaxies and the masses of galaxies and clusters of galaxies from the virial theorem and from weak gravitational lensing) can be interpreted in this way. Due to our ignorance about the voids among the galaxies it is not yet possible to extract information on the York time from local fits to the non-local York time. This explanation of dark matter may look like a gauge artifact and not like an intrinsically defined observable effect.

Since in GR the gauge freedom is the arbitrariness in the choice of the 4-coordinates, the observables describing the tidal effects (and also the electromagnetic field) should be 4-scalars. However, at the experimental level the description of *macroscopic matter* (and also of the spectra of light from stars) is not based on 4-scalars but is *intrinsically coordinate-dependent*, namely is connected with the *metrological conventions* used by physicists, engineers and astronomers for the modeling of space-time. The basic conventions are: (a) an atomic clock as a standard of time; (b) the 2-way velocity of light in place of a standard of length; (c) a conventional reference frame centered on a given observer as a standard of space-time (GPS is an example of such a standard). The adopted astronomical reference frames are:

(A) The description of satellites around the Earth is done by means of NASA coordinates either in ITRS (the terrestrial frame fixed on the Earth surface) or in GCRS (the geocentric frame centered on the Earth center).
(B) The description of planets and other objects in the Solar System uses BCRS (a barycenter quasi-inertial Minkowski frame, if perturbations from the Milky Way are ignored; it is a PN Einstein space-time with 3-spaces having a very small extrinsic curvature of order c^{-2} and with a PN treatment of the gravitational field of the Sun and of the planets in a special harmonic gauge of Einstein GR), centered in the barycenter of the Solar System, and ephemerides.
(C) In astronomy the positions of stars and galaxies are determined from the data (luminosity, light spectrum, angles) on the sky as living in a 4-dimensional nearly-Galilei space-time with the celestial ICRS frame considered as a "quasi-inertial frame" (all galactic dynamics is Newtonian gravity), in accord with the assumed validity of the cosmological and Copernican principles. Namely one assumes a homogeneous and isotropic cosmological Friedmann-Robertson - Walker solution of Einstein equations (the standard ΛCDM cosmological model). In it the constant intrinsic 3-curvature of instantaneous 3-spaces is nearly zero as implied by the CMB data, so that Euclidean 3-spaces (and Newtonian gravity) can be used. However, to reconcile all the data with this 4-dimensional reconstruction one must postulate the existence of dark matter and dark energy as the dominant components of the classical universe after the recombination 3-surface! What is still lacking is a PM extension of the celestial frame such that the PM BCRS frame is its restriction to the solar system inside our galaxy. Hopefully this will be achieved with the ESA GAIA mission devoted to the cartography of the Milky Way. This metrological extension should produce 3-spaces whose associated York time is determined by the elimination of dark matter.

Finally this point of view could also eliminate all or part of *dark energy* if one relaxes the Killing symmetries (homogeneity and isotropy) of the standard FRW cosmological model (they imply that the York time is fixed to minus the Hubble constant). In inhomogeneous solutions like Szekeres space-times the York time is an inertial gauge variable and the luminosity distance to supernova's depends on it!

References

1. Alba D, Lusanna L (2007) The york map as a Shanmugadhasan canonical transformation in tetrad gravity and the role of non-inertial frames in the geometrical view of the gravitational field. Gen Rel Grav 39:2149 (gr-qc/0604086, v2).
2. Alba D, Lusanna L (2010a) Charged particles and the electro-magnetic field in non-inertial frames: I. Admissible 3 + 1 splittings of minkowski spacetime and the non-inertial rest frames. Int J Geom Methods Phys 7:33 (arXiv 0908.0213).
3. Alba D, Lusanna L (2010b) The Einstein-Maxwell-Particle System in the York canonical basis of ADM tetrad gravity: (I) the equations of motion in arbitrary Schwinger time gauges, (arXiv 0907.4087), (II) the weak field approximation in the 3-orthogonal gauges and hamiltonian post-minkowskian gravity: the N-body problem and gravitational waves with asymptotic background (arXiv 1003.5143), (III) the post-minkowskian N-Body problem, its Post-Newtonian Limit in non-harmonic 3-orthogonal gauges and dark matter as an inertial effect. (arXiv 1009.1794). To be published in Canad J Phys.
4. Beig R, Murchadha Ó (1987) The Poincaré group as the symmetry group of canonical general relativity. Ann Phys 174:463
5. DePietri R, Lusanna L, Martucci L, Russo S (2002) Dirac's observables for the rest-frame instant form of tetrad gravity in a completely fixed 3-orthogonal gauge. Gen Rel Grav 34:877 (arXiv gr-qc/0105084).
6. Isenberg J, Marsden JE (1984) J Geom Phys 1:85
7. Lusanna L (2001) The rest-frame instant form of metric gravity. Gen Rel Grav 33:1579 (arXiv gr-qc/0101048).
8. Lusanna L (2011) Canonical gravity and relativistic metrology: from clock synchronization to dark matter as a relativistic inertial effect (arXiv 1108.3224).
9. Lusanna L, Pauri M (2006) The physical role of gravitational and gauge degrees of freedom in general relativity. I: dynamical synchronization and generalized inertial effects; II: dirac versus Bergmann Observables and the objectivity of space-time. Gen Rel Grav 38:187, 229 (arXiv gr-qc/0403081 and 0407007).
10. Lusanna L, Russo S (2002) A new parametrization for tetrad gravity. Gen Rel Grav 34:189 (arXiv gr-qc/0102074).

Chapter 18
Experimental Tests of Quantum Mechanics: Pauli Exclusion Principle and Spontaneous Collapse Models

Catalina Curceanu Petrascu, Sergio Bartalucci, Mario Bragadireanu, Alberto Clozza, Carlo Guaraldo, Mihai Iliescu, Alessandro Rizzo, Antonio Romero Vidal, Alessandro Scordo, Diana Laura Sirghi, Florin Sirghi, Laura Sperandio, Oton Vazquez Doce, Angelo Bassi, Sandro Donadi, Edoardo Milotti, Matthias Laubenstein, Sergio Bertolucci, Mario Bragadireanu, Catalina Curceanu, Dorel Pietreanu, Titus Ponta, Michael Cargnelli, Tomoichi Ishiwatari, Johann Marton, Eberhard Widmann, Johann Zmeskal, Sergio di Matteo and Jean Pierre Egger

Abstract The Pauli exclusion principle (PEP), as a consequence or the spin-statistics connection, is one of the basic principles of the modern physics. Being at the very basis of our understanding of matter, it spurs a lively debate on its possible limits, deeply rooted as it is in the very foundations of Quantum Field Theory. The VIP (VIolation of the Pauli exclusion principle) experiment is searching for a possible small violation of the PEP for electrons, using the method of searching for Pauli Exclusion Principle forbidden atomic transitions in copper. We describe the experimental method and the obtained results; we briefly present future plans to go beyond the actual limit by upgrading the experiment using vetoed new spectroscopic fast

C. C. Petrascu (✉) · S. Bartalucci · M. Bragadireanu · A. Clozza · C. Guaraldo · M. Iliescu · A. Rizzo · A. R. Vidal · A. Scordo · D. L. Sirghi · F. Sirghi · L. Sperandio · O. V. Doce
INFN, Laboratori Nazionali di Frascati, CP 13 Via E. Fermi 40, 00044 Frascati, Roma, Italy

A. Bassi · S. Donadi · E. Milotti
Dipartimento di Fisica, Universita' di Trieste and INFN—Sezione di Trieste, Via A. Valerio, 2, 334127 Trieste, Italy

M. Laubenstein
INFN, Laboratori Nazionali di Gran Sasso, S.S. 17/bis, 67010 Assergi, AQ, Italy

S. Bertolucci
CERN, CH-1211 Geneva 23, Switzerland

M. Bragadireanu · C. Curceanu · D. Pietreanu · T. Ponta · D. L. Sirghi · F. Sirghi
"Horia Hulubei" National Institute of Physics and Nuclear Engineering, Atomistilor no. 407, P.O. Box MG-6, Magurele-Bucharest, Romania

M. Cargnelli · T. Ishiwatari · J. Marton · E. Widmann · J. Zmeskal
The Stefan Meyer Institute for Subatomic Physics, Boltzmanngasse 3, 1090 Vienna, Austria

S. Matteo
Institut de Physique UMR CNRS-UR1 6251, Universite' de Rennes1, 35042 Rennes, France

J. P. Egger
Institut de Physique, Universite' de Neuchatel, 1 rue A.-L. Breguet, 2000 Neuchatel, Switzerland

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_18,

Silicon Drift Detectors. We also mention the possibility of using a similar experimental technique to search for possible X-rays generated in the spontaneous collapse models of quantum mechanics.

18.1 Introduction

The Pauli Exclusion Principle (PEP), which plays a fundamental role in our understanding of many physical and chemical phenomena, is a consequence of the spin-statistics connection, [1], and, as such, it is intimately connected to the basic axioms of quantum field theory [2]. Although the principle has been spectacularly confirmed by the number and accuracy of its predictions, its foundation lies deep in the structure of quantum theory and has defied all attempts to produce a simple proof, as stressed for example by Feynman et al. [3]. Pauli himself in his Nobel lecture declared: "...Already in my original paper I stressed the circumstance that I was unable to give a logical reason for the exclusion principle or to deduce it from more general assumptions.....The impression that the shadow of some incompleteness (falls) here on the bright light of success of the new quantum mechanics seems to me unavoidable".

Given its basic standing in quantum theory, it seems appropriate to carry out precise tests of the PEP validity and, indeed, mainly in the last 15–20 years, several experiments have been performed to search for possible small violations [4–8] and [9]. Often, these experiments were born as by-products of experiments with a different objective (e.g., dark matter searches, proton decay, etc.), and most of the recent limits on the validity of PEP have been obtained for nuclei or nucleons.
In 1988 Ramberg and Snow [10] performed a dedicated experiment, searching for anomalous X-ray transitions, that would point to a small violation of PEP in a copper conductor. The result of the experiment was a probability ([11]) $\beta^2/2 < 1.7 \times 10^{-26}$ that the PEP is violated by electrons.

The VIP Collaboration set up a much improved version of the Ramberg and Snow experiment, with a higher sensitivity apparatus, VIP Proposal [12]. Our final aim is to improve the PEP violation limit for electrons by 3–4 orders of magnitude, by using high resolution Charge-Coupled Devices (CCDs) as soft X-rays detectors [13–17], and decreasing the effect of background by a careful choice of the materials and sheltering the apparatus in the LNGS underground laboratory of the Italian Institute for Nuclear Physics (INFN).

In the next sections we describe the experimental method and the experimental setup, the results of a first measurement performed in the Frascati National Laboratories (LNF) of INFN, along with results obtained by running VIP at the underground Gran Sasso National Laboratory (LNGS) of INFN.

We then briefly present future plans to go beyond the existing limit by using fast Silicon Drift Detectors (SDD) and a veto system.

We conclude the paper by presenting some ideas to use a similar experimental technique to perform measurements of X-rays predicted by spontaneous collapse models in quantum mechanics.

18.2 The VIP Experiment

VIP is a dedicated experiment for the measurement of the probability of the Pauli Exclusion Principle violation for electrons. The experiment uses the same method of the Ramberg and Snow experiment, with a much better soft X-ray detector in a low-background experimental area—the INFN Gran Sasso underground laboratory. The detector is an array of Charge-Coupled Devices (CCDs), characterized by the excellent background rejection capability, based on pattern recognition, and good energy resolution (320 eV FWHM at 8 keV in the present measurement).

18.2.1 The Experimental Method

The experimental method consists in the introduction of "fresh" electrons into a copper strip, by circulating a current, and in the search for the X-rays resulting from the forbidden radiative transitions that occur if one of these electrons is captured by a copper atom and cascades to a 1S state which is already filled by two electrons. In particular we are looking for the 2P to 1S transition.

The energy of this non-Paulian transition would differ from the normal transition energy by about 300 eV (7.729 keV instead of 8.040 keV), due to the additional screening effect given by the second electron on the 1S level, and was calculated using two different approaches [18], providing an unambiguous signal of PEP violation. The new value is more precise than the rough estimate given in paper of Ramberg and Snow, where the shift, about 600 eV in that case, was approximated as the difference between the normal Copper transition from 2P to 1S level and the corresponding Nickel (Z-1 with respect to Copper) one, no real calculation of the PEP violating transition being done. The measurement alternates periods without current in the copper strip, in order to evaluate the X-ray background in conditions where no PEP violating transitions are expected to occur, with periods in which current flows in the conductor, when we expect that the "fresh" electrons may undergo Pauli-forbidden transitions.

18.2.2 The VIP Setup

The VIP setup consists of an empty copper cylinder, 45 mm radius, 50 μm thickness, and 88 mm height, surrounded by 16 equally spaced "type 55" CCDs made by EEV. The CCDs are at a distance of 23 mm from the copper cylinder, and paired one above the other. The setup is enclosed in a vacuum chamber, and the CCDs are cooled to about 168 K by a cryogenic system. The current flows in the thin cylinder made of ultrapure 99.995 % copper foil from the bottom of the vacuum chamber. The CCDs surround the cylinder and are supported by cooling fingers which protrude from the cooling heads in the upper part of the chamber. The readout electronics is just behind the cooling fingers; the signals are sent to amplifiers on top of the chamber and the amplified signals are read out by ADC boards in the data acquisition computer. More details on CCD-55 performance, as well on the analysis method used to reject background events, can be found in Refs. [19–21], VIP improves very significantly on the Ramberg and Snow measurement, thanks to the following features:

- use of CCD detectors instead of gaseous detectors, having much better energy resolution (4–5 times better) and higher stability;
- experimental setup located in the clean, low-background, environment of the underground LNGS Laboratory;
- collection of much higher statistics (longer DAQ periods, thanks to the stability of CCDs).

18.3 The VIP Experimental Results

Before installation in the Gran Sasso laboratory, the VIP setup was prepared and tested at the LNF-INFN laboratory, where measurements were performed in the period 21 November–13 December 2005. Two types of measurements were performed:

- 14,510 min (about 10 days) of measurements with a 40 A current circulating in the copper target;
- 14,510 min of measurements without current.

CCDs were read-out every 10 min. The resulting energy calibrated X-ray spectra are shown in Fig. 18.1.

These spectra include data from 14 CCD's out of 16, because of noise problems in the remaining 2.

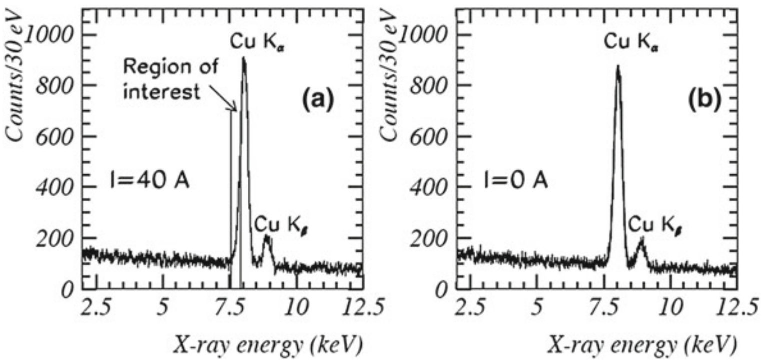


Fig. 18.1 Energy spectra with the VIP setup at LNF-INFN: **a** with current (I = 40 A); **b** without current

Both spectra, apart from the continuous background component, display clear Cu lines due to X-ray fluorescence caused by the cosmic ray background and natural radioactivity. No other lines are present and this reflects the careful choice of the materials used in the setup. The subtracted spectrum is structureless. This not only yields an upper bound for a violation of the Pauli Exclusion Principle for electrons, but also confirms the correctness of the energy calibration procedure and points to the absence of systematic effects.

To extract the experimental limit on the probability that PEP is violated for electrons, $\beta^2/2$, from our data, we used the same arguments of Ramberg and Snow. See details in [22]. The obtained value is:

$$\beta^2/2 < 4.5 \times 10^{-28} \tag{18.1}$$

Thus with this first measurement in an unshielded environment, we have improved the limit obtained by Ramberg and Snow by a factor about 40.

The experiment was installed at LNGS-INFN in Spring 2006 , and was in data taking until Spring 2010, alternating period with current on (signal) to periods with current off (background).

We have established a preliminary new limit on PEP violation by electrons from data taken at LNGS:

$$\beta^2/2 < 4.7 \times 10^{-29} \tag{18.2}$$

18.4 Future Perspectives

The presented VIP setup uses CCD detectors, which are integrating detectors (no timing capability), for the measurement of the X-rays. In the future we plan to switch to a new type of detectors, namely the triggerable Silicon Drift Detectors (SSD), which have a fast readout time (1 μs), a large collection area (1 cm^2) and an energy resolution a factor about 2 better than the one of the used CCDs.

These detectors were successfully used in the SIDDHARTA experiment, [23], for measurements of the kaonic atoms transitions at the DAFNE accelerator of LNF-INFN; using a proper trigger system a background rejection factor of the order of 10^{-4} was achieved in SIDDARTHA.

With these new detectors and with a more compact setup (higher acceptance) we expect a further reduction of the background produced by charged particles coming from the outside of the setup. It wasestimated that we can gain other 2–3 orders of magnitude in the $\beta^2/2$ factor.

Presently, the experimental setup is being under construction, with the aim to install it at LNGS in 2012.

Apart of the measurements of X rays related to the violation of PEP, we are presently considering the possibility to perform in the future measurements of X rays (exploiting these excellent X-ray detectors, the CCDs and SDDs) generated as spontaneous radiation predicted by (some) collapse models.

The collapse models deal with the "measurement problem" in quantum mechanics by introducing a new physical dynamics that naturally collapses the state vector. In the nonrelativistic collapse model developed by [24–26] and [27, 28] (see also [29] for a review), namely the continuous spontaneous localization (CSL) model, the state vector undergoes a nonunitary evolution in which particles interact with a fluctuating scalar field. This interaction has not only the effect of collapsing the state vector towards the particle number density eigenstates in position space, but it increases the expectation value of particle's energy as well. This means, for a free charged particle (as the electron) electromagnetic radiation. This type of phenomenon is predicted by the CSL and is totally absent in standard quantum mechanics.

In paper [30] a pioneering work on this spontaneous emission of radiation was performed - the author analyzed X-ray data measured in an underground experiment and interpreted them as a limit for the CSL parameter(s). It was shown that the highest sensitivity is for few keV X-rays, exactly in the range where our detectors are ideal.

We plan to perform a feasibility study to define a dedicated experiment to measure X-rays coming from the spontaneous collapse models. In this way the same experimental technique would test different aspects of fundamental aspects of quantum theory.

Acknowledgments The VIP Collaboration wishes to thank all the LNGS laboratory staff for the precious help and assistance during all phases of preparation, installation and data taking. The support from the HadronPhysics2 FP7 (227431), from the MIUR PRIN2008 2008LH2X28_004 and from COST1006, Fundamental Problems in Quantum Physics projects is acknowledged.

References

1. Pauli W (1949) Phys Rev 58:716
2. Luders G, Zumino B (1958) Phys Rev 110:1450–1453
3. Feynman RP, Leighton RB, Sands M (1963) The Feynman lectures on physics. Addison-Wesley, Reading.
4. Bernabei R et al (1997) Phys Lett B 408:439
5. Back HO et al (2005) Eur Phys J C37:421–431
6. Hilborn RC, Yuca CL (1996) Phys Rev Lett 76:2844–2847
7. Nemo Colaboration (2000) Nucl Phys B (Proc Suppl) 87:510–511
8. Nolte E et al (1991) J Phys G Nucl Part Phys 17:S355
9. Tsipenyuk Y, Barabash A, Kornoukhov V, Chapyzhnikov B (1998) Phys Chem 51:507
10. Ramberg E, Snow GA (1990) Phys Lett B 238:438
11. Yu Ignatiev A, Kuzmin VA (1987) Yad Fiz 46:786.
12. The VIP proposal, LNF-LNGS Proposal, september, 2004, http://www.lnf.infn.it/esperimenti/vip
13. Culhane JL (1990) Nucl Instrum Methods A 310:1
14. Egger JP, Chatellard D, Jeannet E (1993) Particle World 3:139
15. Fiorucci G et al (1990) Nucl Instrum Methods A 292:141
16. Varidel D et al (1990) Nucl Instrum Methods A 292:147
17. Kraft RP et al (1995) Nucl Instrum Methods A 372:372
18. di Matteo S, Sperandio L (2006) VIP Note, IR-04, 26 April 2006; the energy shift has been computed by Indelicato P, private communication.

19. Ishiwatari T et al (2004) Phys Lett B 593:48
20. Beer G et al (2005) Phys Rev Lett 94:212302
21. Ishiwatari T et al (2006) Nucl Instrum Methods Phys Res A 556:509
22. Bartalucci S (VIP Collaboration) et al. (2006) Phys Lett B 641:18.
23. Bazzi M et al (2009) Phys Lett B 681:310
24. Ghirardi GC, Rimini A, Weber T (1986) Phys Rev D 34:470
25. Ghirardi GC, Rimini A, Weber T (1987) Phys Rev 3287:470
26. Ghirardi GC, Rimini A, Weber T (1988) Found Phys 18:1
27. Pearle P (1990) Phys Rev A 39:2277
28. Ghirardi GC, Pearle P, Rimini A (1990) Phys Rev A 42:78
29. Bassi A (2007) J Phys 67(2007):012013
30. Fu Q (1997) Phys Rev A 56:1806

Chapter 19
CMB Anisotropy Computations Using Hydra Gas Code

Màrius Josep Fullana i Alfonso, Josep Vicent Arnau i Córdoba, Robert J. Thacker, Hugh M. P. Couchman and Diego P. Sáez Milán

Abstract From FFP6 to FFP11, we presented the advances in our Cosmic Microwave Background (CMB) anisotropy computations using N-body Hydra Codes. For such computations, codes without baryons were used: First sequential versions and afterwards parallel ones. With both of them we computed the weak lensing and the Rees-Sciama contributions to the CMB angular power spectrum. Using our numerical techniques, we reported a lensing effect higher than that estimated in previous papers (for very small angular scales). Our CMB computations require less interpolations and approximations than other approaches. This could explain part of our excess of power in lensing computations. Our higher time and angular resolutions could also contribute to this excess. Here, recent advances on previous computations are presented. Computations with baryons have been started. These calculations allow us to compute the Sunyaev-Zel'dovich contribution to the CMB angular power spectrum. We are also trying to compute the three effects—weak lensing, Rees-Sciama and

M. J. F. i Alfonso (✉)
Institut de Matemàtica Multidisciplinària, Universitat Politècnica de València,
Camí de Veras/n, 46022 València, Spain
e-mail: mfullana@mat.upv.es

J. V. A. i Córdoba
Departament de Matemàtica Aplicada, Universitat de València, Dr. Moliner 50,
46100 Burjassot, Spain

R. J. Thacker
Department Astronomy and Physics, St. Mary's University, Halifax ,
Nova Scotia B3H 3C3 , Canada

H. M. P. Couchman
Department of Physics and Astronomy, McMaster University, 1280 Main St. West,
Hamilton, Ontario L8S 4M1, Canada

D. P. S. Milán
Departament d'Astronomia i Astrofísica, Universitat de València, Dr. Moliner 50,
46100 Burjassot, Spain

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_19,

Sunyaev-Zel'dovich—at the same time, with the essential aim of seeing how their power spectra couple among them.

19.1 Introduction

Rough estimates of CMB anisotropies performed with PM codes may be found, e.g., in [1] and [2]. These calculations were improved by using AP3M codes with more resolution. We first used a Hydra AP3M sequential code [3], which was modified to move the CMB photons through the AP3M n-body simulations. Only dark matter was taken into account to evolve structures and compute Rees-Sciama (RS) and weak lensing (WL) CMB anisotropies. Afterwards, a Hydra AP3M parallel code was used to do the same estimates (RS and WL effects) with more resolution and bigger boxes [4]. Again only dark matter was taken into account in the evolution of structure.

Now, we are moving CMB photons along the simulation boxes of a Hydra AP3M parallel code with baryons. The RS and WL CMB anisotropies have been calculated again—with similar boxes and resolution— to compare with previous results obtained without baryons. The calculation of the Sunyaev-Zel'dovich (SZ) effect is being performed by using appropriate resolutions and our ray-tracing techniques, which were designed to move CMB photons thorough the simulated boxes while the code is running. The peculiar gravitational potential, its gradients, the electron number density, and other necessary quantities are calculated and used at every time step of the Hydra simulation. More details about previous work may be found in other FFP Proceedings [1], [3], [5], and [6].

19.2 Map Construction

WL deflections are given by the following formula:

$$\delta = -2\int W(\lambda)\nabla_{\perp}\phi \, \mathrm{d}\lambda, \tag{19.1}$$

where $\nabla_{\perp}\phi$ is the transverse gradient of the peculiar gravitational potential, $W(\lambda) = (\lambda_{em} - \lambda)/\lambda$, and

$$\lambda(a) = H_0^{-1} \quad \int (\Omega_m b + \Omega_\Lambda b^4)^{-1/2} db, \tag{19.2}$$

H_0, Ω_m, and Ω_Λ being the Hubble constant, and the density parameters of matter and vacuum, respectively. By using these formulas, the temperature contrasts $\Delta\, T/T$, and the C_l coefficients due to WL may be calculated [4]. The integral (19.2) is to be done in the interval [a,1].

The RS, thermal SZ, and kinetic SZ temperature contrasts are given by the expressions

$$\Delta T/T = 2 \int \partial\phi/\partial t dt, \tag{19.3}$$

$$\Delta T/T = (-2\sigma_T k/m_e c^2) \quad \int n_e(T_e - T_{CMB})dl, \tag{19.4}$$

and

$$\Delta T/T = -(\sigma_T/c) \int \mathrm{v}_r n_e dl, \tag{19.5}$$

respectively, where n_e, T_e, T_{CMB}, v_r, σ_T, k, m_e and c are the electron number density, the electron and CMB temperature, the radial component of the peculiar velocity, the Thompson cross section, the Boltzmann constant, the electron mass and the speed of light, respectively. Integrals (19.1), (19.3), (19.4) and (19.5) are to be calculated along the background null geodesics.

19.3 Evaluating Variables to Perform Integrations

In order to compute the integrals of Eqs. 19.1, 19.3, 19.4 and 19.5, we proceed as follows:

(1) Select the propagation directions of the CMB photons (ray-tracing).
(2) Assume the Born approximation, and use the photon step distance, Δ_{ps}, to determine all the evaluation positions on the background null geodesics, from the initial to the final redshift. Then, localize each of these positions inside one of the simulation boxes.
(3) Associate a test particle to each evaluation position. Times are obtained from the null geodesic equations in the background.
(4) At each time step of the N-body simulation (while it is running), determine which test particles require evaluations (peculiar potential, forces, temperature, and so on), and evaluate by using the long-range FFT component and short-range PP correction given by the Hydra algorithm (if it is necessary). The SZ effect, Eqs. 19.4 and 19.5, requires the number density and temperature of electrons on the test particle position, which may be estimated from the outputs of the Hydra code. The RS effect given by Eq. 19.3 uses the peculiar gravitational potential and, finally, the WL effect is obtained from the transverse component of the peculiar gravitational force.
(5) Avoid contributions to the peculiar gravitational potential due to scales larger than 42 h^{-1} Mpc, it may be done by removing the signal, in Fourier space, for wavenumbers $k < 0.15$ h Mpc^{-1}. The use of this cutoff in WL and RS calculations is justified in next section.
(6) If the evaluation time for a test particle lies between the two times defining a time step of the simulation, use linear interpolation between these two times to do the required evaluation.

19.4 Description of the Simulations

Simulations are done in the framework of the concordance model with the following parameters: reduced Hubble constant, $h = 0.7$; baryon, dark matter, and dark energy density parameters $\Omega_b = 0.046$, $\Omega_d = 0.233$ and $\Omega_\Lambda = 0.721$, respectively; optical depth for reionization, $\tau = 0.084$, and matter power spectrum normalisation parameter, $\sigma_8 = 0.817$. No tensor modes are considered. It has been verified that the hydra code produces the same structure without CMB photons (original tested version) and with them. The CMB photons are moved through the simulation box along specially chosen oblique paths and the boxes are not moved at all. In this way, the periodicity inside the box volume ensures that there are no discontinuities in the fields where photons cross from box to box. We take care to ensure that, for the chosen paths, periodicity effects become negligible. To minimize these effects, the photons must cross consecutive boxes through statistically independent regions, which require: (i) *preferred directions* leading to large enough distances between these regions and, (ii) a suitable cutoff avoiding large scale spatial correlations ($k < 0.15 hMpc^{-1}$) between the above distant regions. For the preferred directions (see [3], [7], and [8]), and box sizes of $L_{box} = 256\,\mathrm{h}^{-1}$Mpc, it can be easily verified that the CMB photons can travel from $z = 6$ to $z = 0$ ($\sim$5,900 h^{-1}Mpc) through different uncorrelated regions (without repetitions) located in successive simulation boxes. The total number of crossed boxes is *close to* 20. Our squared CMB maps are uniformly pixelised by choosing a certain number of pixels, N_{pix}, per edge. The angular resolution is then $\Delta_{ang} = \Phi_{map}/N_{pix}$ (Φ_{map} being the angular size of the map side). A preferred direction points toward the centre of any squared map. Since the maps are small, the directions of all the pixels are close to the central one and, consequently, they are also preferred directions. Therefore, lens deviations can be calculated for each pixel, with no significant periodic effects across the full map. For greater box sizes, the situation is better (see [4]).

This is notably different from other approaches using random translations and rotations [9] of the simulation box which, unavoidably, lead to discontinuities at crossing points between adjoining boxes.

Let us now list the parameters involved in our ray-tracing procedure: a number of directions, N_{dir}, per edge of the squared CMB map (one per pixel, $N_{dir} = N_{pix}$); an initial redshift, z_{in}, a photon step, Δ_{ps}, to perform the integrals in Eqs. 19.1, 19.3, 19.4 and 19.5; and the angles θ and φ defining the preferred direction. A simulation (with CMB photons) is characterized by the parameters and initial conditions required by the Hydra simulations (without CMB photons) together with the parameters of the ray-tracing.

For computations without baryons (WL and RS effects), the simulation parameters chosen in this paper are: $L_{box} = 512\,\mathrm{h}^{-1}$Mpc, number of particles $N_p = 512^3$, number of cells $N_c = 1,024^3$, softening $S_p = 12\,\mathrm{h}^{-1}$kpc, $N_{dir} = 512$, $z_{in} = 6$ and $\Delta_{ps} = 25\,\mathrm{h}^{-1}$kpc. The angular resolution of these simulations is $\Delta_{ang} \rhd 0.59'$ ($l \rhd$ 18,600). The map size is $\sim 5^\circ \Box\, 5^\circ$. The effective resolution is $E_{res} \sim 60\ \mathrm{h}^{-1}$kpc. N-body simulations of this type were presented in [4]. In our new simulations with baryons, the same parameters, except $N_p = 2 \square 512^3$, are first chosen to facilitate some necessary comparisons. Nevertheless, further applications of our simulations with baryons require more resolution, and the above simulation parameters must be changed (see below).

19.5 Previous Results and Perspectives

N-body simulations as those described in Sect. 19.4 were previously used [4] to estimate the angular power spectrum of the WL effect. The signal in the range 4,200 $< l <$ 7,000 was $2.0 \pm 0.4\,\mu$K, which is $\sim$1.4 μK higher than that found elsewhere [10]. Our direct estimate of the potential gradients at photon positions—using the same PP corrections as for dark matter particles—seems to be the main origin of the difference between our results and other researches based on projection planes and grid interpolations. These differences appear only at large $l >$ 2,000. Moreover, our method employs extremely fine time resolution (that of the AP3M simulation) and also a very good angular resolution. Our code was run for a wide range of the parameters involved in the N-body simulations and in the ray-tracing procedure (see Sect. 19.4), in this way, it was studied how the resulting angular power spectra depend on all these parameters. Since the dependence is weak, results appear to be robust [4], at least in the l-interval (4,200, 7,000).

On small scales, baryons do not follow the dark matter distribution. Therefore, while we have attempted to be as accurate as possible in our dark matter simulations (without baryons), we are probing scales where contributions from baryons are beginning to become significant. A study of the impact of baryons can be found in [11]. We are beginning to conduct simulations with baryons and feedback processes both to identify its impact on WL and RS signals, and also to systematically evaluate the combined impact of the SZ, WL, and RS effects.

The SZ effect dominates on total lensing for 4,200 $< l <$ 7,000. In [4], we discussed the implications of our estimates taking into account the CBI observations [12] and the BIMA measurements [13], which were the most significant anisotropy measurements (at very small angular scales) when that paper was written. The discussion took into account that the SZ power scales as $\sigma_8^{3.5}$. Then, the SZ effect associated to the σ_8 value measured by WMAP was too small to explain the high power observed by BIMA for the small angular scales under consideration. In that situation, our excess of WL power seemed to be an appropriate though small excess contributing to the explanation of the large BIMA power.

Afterwards, new data for small angular scale measurements of the CMB anisotropy changed the situation. Observations performed with the ACT [14] and the SPT [15] telescopes indicated that, for large l-values, the CMB power is much smaller than that previously reported by CBI and BIMA experiments. The new situation was discussed in [6]. For example, in the l-interval (5,000, 6,000), the ACT power is between $\sim$40 and $\sim$50 μK^2 (see Fig.4 of [14]). This power might be explained by the coupling of SZ, foreground radiation of Dusty Star Forming Galaxies (DSFG), plus almost negligible lensing (that predicted by previous simulations); however, our large WL effect of $\sim 6\,\mu K^2$ might complicate the explanation of the small ACT power. Anyway, our WL power may be compatible with the small ACT and SPT powers taking into account that (i) there are uncertainties in the radiation from DSFG, (2) the SZ effect should be studied in more detail, e.g., by using our ray-tracing procedure, and (3) the coupling between SZ, WL and RS effects might lead to a total power different from the simple addition of powers. Of course, the observed power must be explained by the correct coupling of WL, SZ, RS and DSFG contributions. It is worthwhile to emphasize that the WL and SZ effects are essentially produced by the same structure distributions (galaxy clusters and sub-structures involving dark matter and baryons) and, consequently, these effects must be strongly correlated. This implies that the spectra of these two effects must be superposed in an unknown way (not merely added). In practice, this superposition might be analyzed in detail with ray-tracing through hydra simulations including both baryons and dark matter.

19.6 Current Work and Projects

A version of a certain Hydra code with baryons has been modified to include CMB photons. Some subroutines and the computational load allocation configuration of the initial code have been remodelled. The resulting code has been tested. One of the most powerful tests is based on the comparison of the WL and RS power spectra obtained by using simulations with and without baryons (for the parameters given in Sect. 19.4). Results of both types of simulations are expected to be very similar. Small differences might appear as a result of various facts. Let us now list some of them: (i) simulations without baryons have been performed with a SGI Altix 3,700 computer and a Pathscale compiler, and the other type of simulations with an SGI Altix UV 1,000 computer and an Intel compiler, (ii) the initial Hydra codes we have modified –introducing CMB photons– exhibited some technical differences, and (iii) it is known that the presence of baryons alter the WL spectrum at very small angular scales [11]. The WL angular power spectra obtained with and without baryons are compared in Fig. 19.1, where the spectrum with baryons (point-dashed line) appears to be rather similar to that corresponding to dark matter only (dotted line). Although the slight differences are being studied yet, they strongly suggest that codes with and without baryons are working properly and also that new simulations with more resolution seem to be necessary to estimate the physical effect produced by the presence of baryons and to compare with [11]. We think that simulations based on the following parameters would be appropriate: $L_{box} = 200\,\mathrm{h}^{-1}\mathrm{Mpc}$, $N_p = 2 \square 640^3$, $N_c = 1{,}280^3$, $N_{dir} = 512$, $z_{in} = 6$, and S_p and Δ_{ps} to be adjusted to our new computational procedure. The angular resolution of these simulations is $\Delta_{ang} > 0.24'$, and the map size is $\sim 2^\circ \square 2^\circ$. These simulations would be also useful to develop our main projects: (1) the estimate of the SZ effect by using our ray-tracing procedure, (2) the calculation of the total anisotropy produced by the RS, WL, and SZ effects, whose superposition may be nonlinear, and (3) the comparison of the

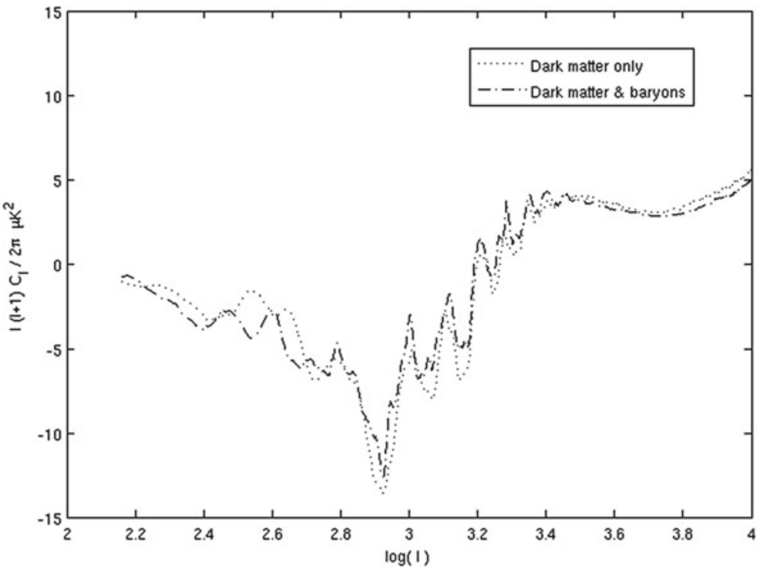


Fig. 19.1 Comparison of WL spectra corresponding to Hydra simulations with and without baryons. Parameters defining these simulations are given in the last paragraph of Sect. 19.4.

resulting total anisotropy with recent observations at very small angular scales [14], [15].

Acknowledgments This work has been partially supported by the Spanish Ministerio de Educación y Ciencia, MEC-FEDER project FIS2009-07705.

References

1. Fullana MJ, Sáez D (2004) Making Maps of the Rees-Sciama Effect. In: Proceedings of FFP6, Udine.
2. Puchades N, Fullana MJ, Arnau JV, Sáez D (2006) On the Rees-Sciama effect: maps and statistics. MNRAS 370:1849–1858
3. Fullana MJ, Arnau JV, Sáez D (2008) Weak lensing on the CMB: estimations based on AP3M simulations. In: Proceedings of FFP9, Udine.
4. Fullana MJ, Arnau JV, Thacker RJ, Couchman HMP, Sáez D (2010a) Estimating small angular scale CMB anisotropy with high resolution N-body simulations: weak lensing. ApJ 712:367–379
5. Fullana MJ, Sáez D (2007) Status of CMB Radiation. In: Sidhart B, Alfonso-Faus A, Fullana MJ (eds) AIP conference proceedings (905) of FFP9, Madrid, pp 13–22.
6. Fullana MJ, Arnau JV, Thacker RJ, Couchman HMP, Sáez D (2010b) Recent observations on CMB at high multipoles and AP3M computations at such scales. In: Proceedings of FFP11, Paris, 2010 (in press).
7. Antón L, Cerdá-Durán V, Quilis V, Sáez D (2005) Cosmic microwave background maps lensed by cosmological structures: simulations and statistical analysis. ApJ 628:1–13

8. Sáez D, Puchades N, Fullana MJ, Arnau JV (2006) Ray-Tracing through N-body simulations and CMB anisotropy estimations. In Proceedings of CMB and physics of the early universe, Ischia, pp 58–62
9. Carbone C, Springel V, Baccigalupi C, Baretelmann M, Matarrese S (2008) Full-sky maps for gravitational lensing of the cosmic microwave background. MNRAS 388:1618–1626
10. Das S, Bode P (2008) A large sky simulation of the gravitational lens- ing of the cosmic microwave background. ApJ 682:1–13
11. Jing YP, Zhang P, Lin WP, Gao L, Springel V (2006) The influence of baryons on the clustering of matter and weak-lensing surveys. ApJ 640:L119–L122
12. Bond JR et al (2005) The Sunyaev-Zel'dovich effect in CMB-calibrated theories applied to the Cosmic Background Imager anisotropy power at $l > 2000$. ApJ 626:12–30
13. Dawson KS, Holzapfel WL, Carlstrom JE, Joy M, LaRoque SJ (2006) Final results from the BIMA CMB anisotropy survey and search for a signature of the Sunyaev-Zel'dovich effect. ApJ 647:13–24
14. Fowler JW et al (2010) The atacama cosmology telescope: a measurement of the $600 < ell < 8000$ cosmic microwave background power spectrum at 148 GHz. ApJ 722:1148–1161
15. Lueker M et al (2010) Measurements of secondary cosmic microwave background anisotropies with the south pole telescope. ApJ 719:1045–1066

Chapter 20
Unimodular Conformal and Projective Relativity: An Illustrated Introduction

Kaća Bradonjić

Abstract This is an illustrated presentation of unimodular conformal and projective relativity, a formulation of unimodular relativity in terms of four independent fields with clear physical and geometric interpretations: conformal structure, four-volume element measure field, projective structure, and affine one-form. We present the motivation for the formalism, physical and geometrical interpretations of the independent fields, and briefly comment on its applications and prospects for quantization.

20.1 Introduction

Unimodular conformal and projective relativity (UCPR) is a formulation of unimodular relativity (UR) in terms of four independent fields with clear physical and geometric interpretation. It is, in a way, an extension of the Palatini approach, which takes the metric and the affine connection as fundamental fields. However, these two structures are not irreducible. In UR, the metric can be decomposed into a conformal metric (or *conformal structure*) and a four volume-element measure field, while an affine connection can be decomposed into a projective connection (or *projective structure*) and an affine one-form. UCPR goes a step further by treating the resulting four irreducible fields as independent dynamical fields.

UCPR was developed as a first step in a search for a background-independent, non-perturbative theory of quantum gravity. We approach this task by first identifying those space-time structures that could be used for measurability analysis of the full inertio-gravitational field. Measurability analysis identifies those concepts that are ideally measurable in the defining context (e.g. concept of hardness in the context of fluid and solid states of matter in classical thermodynamics) and determines the limits on their measurability and their definability in quantum theory (e.g. position

K. Bradonjić (✉)
Physics Department, Wellesley College, Wellesley, MA, USA
e-mail: kbradonjic@gmail.com

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_20,

and momentum in QM). The ideal measurement procedure, which may involve any devices and procedures that are consistent with the theory being studied in the regime in question, should yield the limits on the measurability of the quantities that are in agreement with the uncertainties formally predicted by the theory [1]. Any disagreement between the two indicates an internal inconsistency of the theory. Measurability analysis was successfully conducted by Bohr and Rosenfeld [2], and Bergmann and Smith [1] in the cases of quantum electro-dynamics and linearized general relativity (GR), respectively. Since we don't have a quantum theory of gravity, our goal is to use measurability analysis as a guide to a physically motivated formulation of such a theory. The first step in this approach is to formulate a classical theory of gravity that yields itself to measurability analysis. Our work in this direction has led us to UCPR. The choice of conformal and projective structures as fundamental fields is motivated by the work of Weyl and Ehlers, Pirani and Schild (E-P-S), to name a few, who emphasized the importance of these structures in GR. In a series of papers, E-P-S showed that conformal and projective structures, and consequently the pseudo-Riemannian space-time geometry in GR, can be axiomatically constructed by considering the propagation of massless and massive particles [3, 4]. Being so intimately related to physical fields, conformal and projective structures seem like a good place to start given that we have measurability analysis in mind.

This presentation of UCPR is mostly non-mathematical. It is focused on physical interpretations of the four fundamental fields, as well as various compatibility conditions that can be imposed on them. It is meant to provide a physical interpretation and visual representation of the structures involved. For a more technical introduction, the reader may look at an earlier publication [5]. All considerations of this paper deal with a differentiable manifold M, a symmetric pseudo-Riemannian metric $g_{\mu\nu}$ of signature $(+, -, -, -)$, and a symmetric affine connection $\Gamma^{\kappa}_{\mu\nu}$. Curvature tensors are formed from their respective connections by the following convention [6]: $R_{\dot{\mu}\dot{\nu}\dot{\sigma}^{\kappa}} = \partial_{\mu}\Gamma^{\kappa}_{\nu\sigma} - \partial_{\nu}\Gamma^{\kappa}_{\mu\sigma} + \Gamma^{\kappa}_{\mu\rho}\,\Gamma^{\rho}_{\nu\sigma} - \Gamma^{\kappa}_{\nu\rho}\,\Gamma^{\rho}_{\mu\sigma}$.

20.2 Unimodular Conformal and Projective Relativity

While GR is invariant under the full diffeomorphism group, which means that the symmetry group in the tangent space at each point is the general linear group GL$(4, \mathbb{R})$, UR is invariant under the unimodular diffeomorphisms, which means that the symmetry group in the tangent space at each point is the special linear group SL$(4, \mathbb{R})$ which consists of those point transformations that are volume-preserving. The framework of UR is instrumental in our approach for several reasons, which are discussed below. UCPR assumes four independent fields, illustrated in Fig. 20.1: conformal structure, four volume-element measure field, projective structure, and affine one-form.

Conformal structure, represented by a conformal metric tensor $\tilde{g}_{\mu\nu}$, determines a null cone at each point of M. It distinguishes among spacelike, timelike, and null directions, and determines the causal structure of space-time. Figure 20.1a shows that

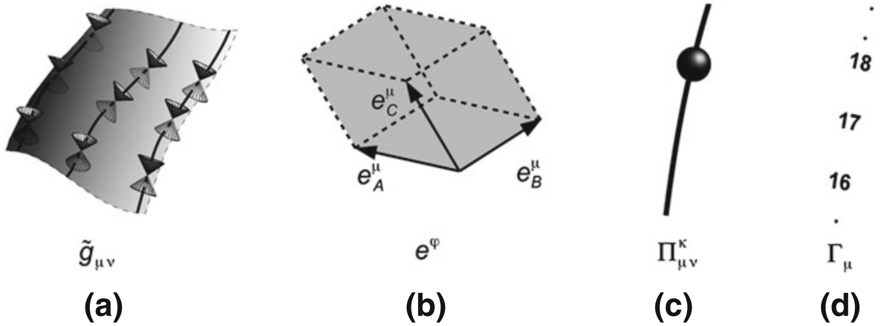


Fig. 20.1 UCPR deals with four independent fields: **a** conformal structure, **b** four volume–element (fourth dimension suppressed), **c** projective structure, and **d** affine one-form

these null cones globally determine characteristic null wave fronts (three-dimensional null hypersurfaces of constant phase) for any zero-rest-mass radiation field, including electromagnetic and gravitational.

In the same way that a metric can be used to construct Christoffel symbols $\{{}^{\kappa}_{\mu\nu}\}$ and the corresponding metric covariant derivative, $\tilde{g}_{\mu\nu}$ can be used to construct conformal Christoffel symbols $\{{}^{\widetilde{\kappa}}_{\mu\nu}\}$ and a conformal covariant derivative. Furthermore, we can use $\{{}^{\widetilde{\kappa}}_{\mu\nu}\}$ *to* construct the conformal-connection curvature tensor $\tilde{C}_{\dot{\mu}\dot{\nu}\dot{\sigma}}{}^{\kappa}$. This curvature tensor is distinct from the Weyl curvature tensor $C_{\dot{\mu}\dot{\nu}\dot{\sigma}}{}^{\kappa}$, which is the conformally invariant part of the metric curvature $K_{\dot{\mu}\dot{\nu}\dot{\sigma}}{}^{\kappa}$.

Four volume-element measure field e^φ is a scalar quantity that is independent of any other space-time structure. As shown in Fig. 20.1b, one can choose a tetrad of basis vectors in the tangent space at each point of the manifold. Such a tetrad forms a parallelepiped at each point and provides a natural four-volume element. The four-volume element measure field e^φ is a weighting of this natural four-volume determined by the basis tetrad. Under SL(4, $\mathbb{R}$) it transforms as a scalar, which ensures that the four-volume at a point is an invariant of the theory. This doesn't mean that e^φ is a fixed non-dynamical field, as it is sometimes assumed in some UR theories, for example [7], but that it is same in all allowed frames of reference and independent of any other space-time structures.

Having an invariant e^φ is advantageous for a couple of reasons. First, a four volume-element is essential for measurability analysis because, as Bohr and Rosenfeld showed, only field averages over space-time regions are measurable. The invariance of e^φ guarantees that we can perform space-time integration of all other fields. Second, some approaches to quantum gravity, such as causal set theory [8], postulate a discrete structure of space-time at high energies and attempt to recover the classical manifold in the low energy limit. In UCPR, quantization of the four-volume may arise from some dynamical procedure, rather than be simply postulated.

A four volume-element measure field and a conformal metric can be combined to form a metric tensor [6],

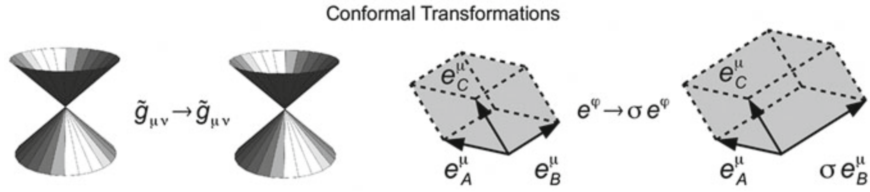


Fig. 20.2 Conformal transformations leave $\tilde{g}_{\mu\nu}$ unchanged and rescale e^{φ}

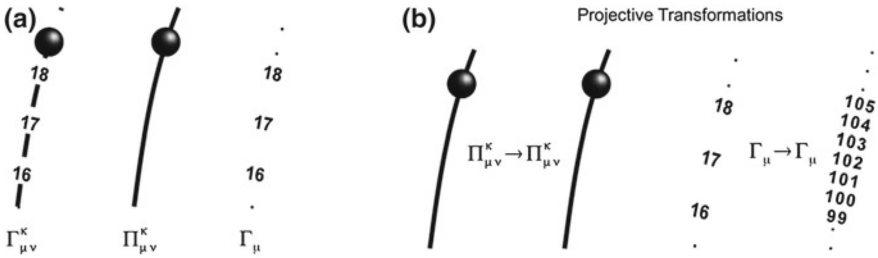


Fig. 20.3 **a** $\Gamma^{\kappa}_{\mu\nu}$ determines curves of freely falling massive particles, $\Pi^{\kappa}_{\mu\nu}$ their unparameterized paths, and Γ_{μ} the parameterization along the paths. **b** Projective transformations leave $\Pi^{\kappa}_{\mu\nu}$ invariant, but change Γ_{μ}.

$$g_{\mu\nu} = e^{\varphi} \tilde{g}_{\mu\nu}. \tag{20.1}$$

Under conformal transformations, $\tilde{g}_{\mu\nu}$ remains invariant, while e^{φ} gets multiplied by the conformal factor (Fig. 20.2). Consequently, we recover the expected conformal transformation of the metric. These transformation properties allow us to treat the conformal structure as an equivalence class of conformally related metrics.

Projective structure determines the auto-parallel paths of freely falling massive particles (Fig. 20.1c). Here we note the important distinction between a path and a curve. One can think of a path as an *unparameterized* trajectory through space-time. A curve, on the other hand, is a *parameterized* trajectory. Mathematically, projective structure is represented by projective parameters $\Pi^{\kappa}_{\mu\nu}$. Under GL(4, $\mathbb{R}$), these parameters have complicated transformations laws. T. Y. Thomas first noted that under SL(4, $\mathbb{R}$) $\Pi^{\kappa}_{\mu\nu}$ transform as components of a symmetric, traceless connection [9]. This allows us to define a projective covariant derivative and the projective-connection curvature tensor $\Pi_{\dot{\mu}\dot{\nu}\dot{\sigma}^{\kappa}}$. $\Pi_{\dot{\mu}\dot{\nu}\dot{\sigma}^{\kappa}}$ is distinct from the usual projective curvature tensor $P_{\dot{\mu}\dot{\nu}\dot{\sigma}^{\kappa}}$, which is the projectively invariant part of the affine curvature tensor $R_{\dot{\mu}\dot{\nu}\dot{\sigma}^{\kappa}}$.

The affine one-form Γ_{μ} determines a parameterization along a path. As shown in Fig. 20.3a, if we strip a curve of its parameterization, we are left with a path. On the other hand, a choice of $\Pi^{\kappa}_{\mu\nu}$ and Γ_{μ} determines a unique affine connection $\Gamma^{\kappa}_{\mu\nu}$ [6],

$$\Gamma^{\kappa}_{\mu\nu} = \Pi^{\kappa}_{\mu\nu} + 1/5\,(\delta^{\kappa}_{\mu}\,\Gamma_{\nu} + \delta^{\kappa}_{\nu}\,\Gamma_{\mu}). \tag{20.2}$$

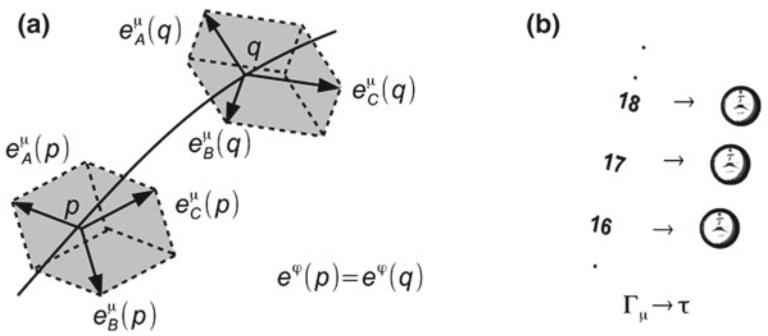


Fig. 20.4 Equi-affine condition (**a**) demands that the parallel transport e^{φ} of is path-independent, and (**b**) allows us to identify the parameterization determined by Γ_{μ} to metrical proper time τ

As shown in Fig. 20.3b, $\Pi^{\kappa}_{\mu\nu}$ is invariant under projective transformations, while Γ_{μ} is not. The projective connection can be thought of as an equivalence class of projectively related affine connections.

20.3 Compatibility Conditions

In the usual formulation of GR, compatibility between $g_{\mu\nu}$ and $\Gamma^{\kappa}_{\mu\nu}$ is imposed by a single condition: $\nabla_{\kappa}g_{\mu\nu} = 0$. UCPR deals with four independent space-time structures, so we can approach the full metric-affine compatibility in steps, as well as impose some new compatibility conditions.

Equi-affine condition: The paralellopiped formed by the basis tetrad is in general transformed by the parallel transport in a path-dependent way.

The equi-affine condition ensures that such transformations are *path-independent*. The usual formulation of this condition in GR demands that the affine covariant derivative of the metric determinant vanishes, and as such imposes a relation between e^{φ} and $\Gamma^{\kappa}_{\mu\nu}$, and consequently $\Pi^{\kappa}_{\mu\nu}$ and Γ_{μ}. In UCPR we can impose the equi-affine condition without imposing any restrictions $\Pi^{\kappa}_{\mu\nu}$ on by simply demanding that

$$\Gamma_{\mu} = 2\partial_{\mu}\varphi. \tag{20.3}$$

This condition also allows us to identify the parameterization of timelike paths determined by Γ_{μ} with the metrical proper time τ (Fig. 20.4b).

Weyl condition as it is formulated in GR demands that the affine covariant derivative of $g_{\mu\nu}$ be proportional to $g_{\mu\nu}$. As such, it imposes a condition on all four fields of UCPR and, by default, relates the parameterization of auto-parallel curves to proper time τ. Physically, this leads to the possibility of the second clock effect: the proper time and the ticking rate of a clock depend on its history (Fig. 20.5). In UCPR, this

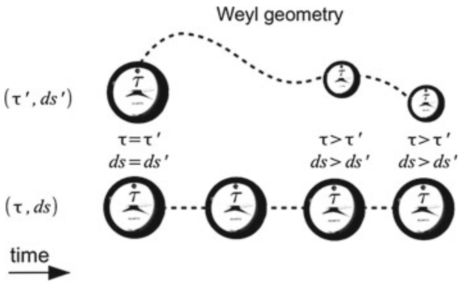


Fig. 20.5 Weyl condition in GR leads to the possibility of the second clock effect in which the ticking rate of the clock depends on its history

unnecessary result is avoided by expressing the Weyl condition in terms of $\Pi^{\kappa}_{\mu\nu}$ and $\{\widetilde{\kappa \atop \mu\nu}\}$, without imposing any relation between Γ_μ and e^{φ},

$$\Pi^{\kappa}_{\mu\nu} - \left\{ \widetilde{\kappa} \atop \mu\nu \right\} = \frac{1}{4}\left[\frac{1}{5}(\delta^{\kappa}_{\mu}\,\Gamma_\nu + \delta^{\kappa}_{\nu}\,\Gamma_\mu) - \tilde{g}_{\mu\nu}\,\tilde{g}^{\kappa\sigma}\,\Gamma_\sigma\right]. \tag{20.4}$$

Full conformal-projective compatibility demands that the conformal geodesics are also projective auto-parallel paths, and is satisfied if $\{\widetilde{\kappa \atop \mu\nu}\} = \Pi^{\kappa}_{\mu\nu}$. This condition also puts no restriction on e^{φ} and Γ_μ.

Full metric-affine compatibility in UCPR holds if *both* equi-affine and Weyl condition hold. However, having four independent fields allows us to choose which conditions to impose, and naturally allows us to consider theories with intermediate conditions.

20.4 Conclusion

UCPR provides the flexibility to construct a variety of theories which dynamize one or more of the four independent fields. The Lagrangian formulation of UCPR with the standard GR action, but now four independent fields, yields the breakup of the usual field equations into an irreducible set of equations that is equivalent to those of GR [5]. Moreover, in UCPR we can construct many different Lagrangians that may provide better descriptions of various physical systems.

The next step in the investigation is to determine which space-time structures of UCPR are good candidates for quantization. While there is much to be done in this direction, there are several approaches that seem to be most promising. The obvious candidates for measurability are the various curvature tensors that can be constructed in UCPR. While it is known that the conformally invariant curvature tensors $\tilde{C}_{\dot\mu\dot\nu\dot\sigma^\kappa}$ and $C_{\dot\mu\dot\nu\dot\sigma^\kappa}$ can be probed with zero rest-mass fields, an analogous analysis of the projectively invariant curvature tensors $\Pi_{\dot\mu\dot\nu\dot\sigma^\kappa}$ and $P_{\dot\mu\dot\nu\dot\sigma^\kappa}$ by using massive particles remains to be done. Furthermore measurability analysis of these fields can be discussed in terms of fields still attached to its sources, and free fields, both in the near and far zones [10]. UCPR also opens numerous possibilities for further investigation. It allows for a construction of a range of theories differing in their choice to dynamize some rather than all four fields, and to impose compatibility conditions that are less restrictive than the full metric-affine compatibility.

References

1. Bergmann PG, Smith GJ (1982) Measurability analysis of the linearized gravitational field. General Relativity and Gravitation 14:1131–1166
2. Bohr N, Rosenfeld L (1933) Zur frage der Meßbarkeit der elektromagnetischen Feldgrößen, Mat-fys. Medd. Dan. Vid. (12), 1; English translation in Selected Papers of Leon Rosenfeld (1979) Cohen R S and Stachel J (ed) Reidel, Dordrecht/Boston/London 357–400
3. Ehlers J, Schild A (1973) Geometry on a manifold with projective structure. Communications in Mathematical Physics 32:119–146
4. Pirani FAE (1973) Building space-time from light rays and free particles. Symposia Mathematica 12:67–83
5. Bradonjić K, Stachel J (2012) Unimodular conformal and projective relativity. Europhysics Letters 97:10001
6. Schouten JA (1954) Ricci-calculus. Springer-Verlag, Berlin
7. Anderson JL, Finkelstein D (1971) Cosmological constant and fundamental length. American Journal of Physics 39:901–904
8. Sorkin RD (1997) Forks in the road, on the way to quantum gravity. International Journal of Theoretical Physics 36:2759–2781
9. Thomas TY (1925) On the projective and equi-projective geometries of paths. In: Proceedings of the National Academy of Sciences (11) 199–203
10. Stachel J (2014) Quantum Gravity: A Heretical Vision, Frontiers of Fundamental Physics and Physics Education. In: Burra S et al (ed) Research Springer Proceedings in Physics (145) 149–158

Part V
Condensed Matter Physics

Chapter 21
Universality of Charge Transport Across Disordered Nanometer-Thick Oxide Films

Mikhail Belogolovskii and Vincenzo Lacquaniti

Abstract Theoretical and experimental analysis of electron transport across ultra-thin, homogeneously disordered oxide layers is presented with particular regard to the question of how much the effects are universal. We show that (i) distribution of transparencies across dirty subnanometer-thick insulating films is bimodal and (ii) conductance-voltage characteristics of oxide layers with thicknesses increased up to several nanometers are power functions with an index near 1.3. The universality of transport properties is explained as an effect of strong local barrier-height fluctuations generated by the presence of oxygen vacancies.

21.1 Introduction

Transport characteristics of a mesoscopic system are usually linked to complicated dynamical processes such as impurity scattering, inter-particle interactions, etc. and, in general, are determined by the system dimensionality, geometry, as well as by other sample-specific parameters. Universal transport properties, if they are, should be independent on microscopic details of particular materials and may include only a limited number of characteristics averaged over the sample. Inter alia, the universal behavior of physical quantities can be estimated for homogeneously disordered films with a very large spread of microscopic parameters.

Increasing interest in the ultra-thin amorphous oxide layers is motivated by their promising applications as a gate dielectric in metal-oxide-semiconductor transistors with higher dielectric constant than that of SiO_2, as well as a blocking dielectric for new-generation flash memory cells [1, 2]. Another field of their applications relates multilayered junctions with quantum-mechanical tunneling as the main physical

M. Belogolovskii (✉)
Donetsk Institute for Physics and Engineering, Donetsk 83114, Ukraine

V. Lacquaniti
National Institute of Metrological Research, 10135 Turin, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_21,

mechanism for electron transport across them [3]. Most such devices are fabricated using aluminum due its superconducting properties and tendency to form a native oxide AlO_x that can be employed as a tunnel barrier. But in both cases significant leakage current, which strongly limits applications of ultra-thin oxide films in semi- and superconducting devices, has been found [1, 4]. Thus, identification of the physical origin of the extra current across ultra-thin oxide films is of great scientific and practical importance.

The needed information can be obtained from measurements of current (I)—voltage (V) characteristics of tri-layered structures with an ultra-thin insulating (I) layer placed between two metallic (M) electrodes. The tunnel current across such a system often exhibits unconventional behavior which does not fit into any theoretical picture. The best way to show it is to transform one or both electrodes from a normal (N) state to a superconducting (S) one. In NIS and even more in SIS trilayers the shape of quasiparticle I–V curves with a single quantum channel is extremely sensitive to the transmission probability D [3]. Because of it, such experiments can provide valuable knowledge concerning the distribution of transparencies $\rho(D)$ in the samples studied. First analysis of the $\rho(D)$.function in subnanometer-thick AlO_x layers was done in the paper [5] by measuring current-versus-voltage and differential conductance-versus-voltage characteristics of planar highly conductive Nb-AlO_x-Nb trilayers at 1.8 K. In spite of the presence of an Al-oxide interlayer, the qquasiparticluasiparticle I–V curves did not exhibit typical for conventional superconducting tunnel junctions subgap resistance R_{sg} (it is measured at voltages $|V| < 2\Delta/e$, Δ is the superconducting energy gap of the order of 1 meV) much greater than the normal-state resistance R_N. Moreover, R_{sg} was of the order of R_N [5]. Another unexpected finding is the shape of I–V characteristics measured for normal-state tunnel junctions with a several nm-thick oxide interlayer in a voltage range of several hundred millivolts. It was found to be a power function with a power index near 7/3.

The figures of merit that we address in this work are the origin of unusual I–V characteristics in heterostructures with disordered ultra-thin oxide films with the thickness d and the universality of the phenomena discussed.

21.2 Universal Distribution of Transparencies in Dirty Subnanometer-Thick Oxide Films

Let us start with a subnanometer-thick oxide barrier which can be modeled by a set of disordered short-range scatterers with strength γ_k at position vectors $\boldsymbol{\rho}_k$ randomly distributed within a plane interface between the two metallic electrodes which is perpendicular to the transport direction. The scattering characteristic of an ultra-thin interface can be calculated from the standard Schrŏdinger equation with a localized potential $V(\mathbf{r}) = V(x, \boldsymbol{\rho}_k) = \sum_k \gamma_k \delta(x)\delta(\boldsymbol{\rho} - \boldsymbol{\rho}_k)$ where the x-axis is orthogonal to the interface and, hence, parallel to the current direction. After some algebra we obtain that the probability of an electron to be transmitted through the disordered ultra-thin interface for states at the Fermi energy E_F is a sum of local transparencies $D_k = (1 + Z_k^2)^{-1}$ with $Z_k = Z(\boldsymbol{\rho}_k) = k_F \int_0^d V(x, \boldsymbol{\rho}_k)dx/E_F$, k_F is the Fermi wave vector. If the parameter Z_k is a uniform random variable ranging from zero to infinity, then its distribution function $\rho(Z) = 2\hbar\overline{G}/e^2 = \text{const}$. Here $\overline{G} = \int_0^\infty \rho(Z)G(Z)dZ$ is the disorder-averaged macroscopic conductance, $G(Z) = \frac{2e^2}{h}D = \frac{2e^2}{h}\frac{1}{1+Z^2}$ is the conductance of a normal-state one-dimensional tunnel junction with a scattering parameter $Z = Hdk_F/E_F$, where H is the barrier height of an ultra-thin potential barrier [6]. With the parametrization $D = (1 + Z^2)^{-1}$ we can transfer to the distribution function of local transparencies $\rho(D)$ which is bimodal with two peaks at $D = 0$ and $D = 1$ $\rho(D) = \hbar\overline{G}\left[e^2 D^{3/2}(1 - D)^{1/2}\right]^{-1}$. This result was obtained, first, by Melsen and Beenakker for a three-dimensional clean M-I-M-I-M structure [7], then by Schep and Bauer for a dirty interface in an M-I-M trilayer [8] and its universality was many times questioned. It can be shown analytically that the transparency of the double-barrier system is also a Lorentzian $D(\theta) = \left[1 + \tilde{Z}(\theta)^2\right]^{-1}$ with a single parameter $\tilde{Z}(\theta)$, a rapidly oscillating function, which changes periodically from zero (for resonance conditions) to very high values during incident angle θ variations from $-\pi/2$ to $-\pi/2$. This behavior which is very similar to that in disordered ultra-thin insulating films discussed above is just the reason why the two distributions for physically different systems do coincide.

The $\rho(D)$ formula contains only a single parameter $\overline{G}$. Our aim was to compare the theory with experimental data using no free parameters. It can be realized by transferring metallic electrodes in the tunnel junction into a superconducting state and dividing the measured I–V curves over those obtained in normal-state experiments. Whereas G-versus-Z dependence for an N-I-N junction is very simple (see above), it is not so for S-I-S devices due to Andreev reflections at the S-I interface when an incident electron (hole) with a probability amplitude $a(\varepsilon)$ is retroreflected into a hole (electron) of the same energy ε (εwill be calculated from the Fermi energy E_F) and almost the same momentum which is travelling in the opposite direction to the incoming charge [3]. Our experimental situation was even more complicated since one of the electrodes was a Nb/Al bilayer and, hence, we were dealing with an asymmetric S_1-I-S_2 junction where S_1 stands for an S/N bilayer. In proximity with Nb a nano-scaled Al layer becomes superconducting and it modifies the standard equation for a homogeneous superconductor (for example, Nb) where $a_{\text{Nb}}(\omega) = i\left(\omega - \sqrt{\omega^2 + \Delta_{\text{Nb}}^2}\right)/\Delta_{\text{Nb}}$ (with ω, the Matsubara frequency) to a more general case with a function $\Phi(\omega)$, which is the ratio of a modified and normal Green's functions in a superconductor $a_{\text{Al}}(\omega) = i\left(\omega - \sqrt{\omega^2 + \Phi_{\text{Al}}^2(\omega)}\right)/\Phi_{\text{Al}}(\omega)$. In the calculations we have used a simplest approximation for $\Phi_{\text{Al}}(\omega)$ in the S/N bilayer derived by Golubov et al. [9] $\Phi_{\text{Al}}(\omega) = \Delta_{\text{Nb}}/\left(1 + C\sqrt{\omega^2 + \Delta_{\text{Nb}}^2}/\Delta_{\text{Nb}}\right)$ and they

Table 21.1 Comparison of theoretical and experimental data for superconducting heterostructures

	Nb/Al-AlO_x-Nb(exper)	NIN (calc)	NIS (calc)	S/NIS (calc)	SIS (calc)
R_{sg}/R_N	1.26	1.0	1.40	1.17	0.68

were based on a numerical method which uses the bimodal distribution $\rho(D)$ and was developed earlier in a few publications; see, for example, the paper [10].

Our experimental results were obtained on asymmetric Nb/Al-AlO_x-Nb junctions developed at INRiM [11] with the Al-interlayer thicknesses d_{Al} ranged from 40 to 150 nm and the exposure dose, the product of the oxygen pressure and the oxidation time, from 150 to 500 Pa·s. Electrical measurements were performed below critical temperatures of Nb/Al bilayers about 8–9 K for different d_{Al} with a conventional four-terminal dc technique. The samples exhibited supercurrents with values of 6–15 mA at 1.7 K and for the measurements of quasiparticle I–V curves we have applied magnetic fields B up to 50 mT through a suitable coil. Normal-state resistance R_N were determined from a linear fit to I–V curves with and without supercurrents at voltages above 1 mV and the results of both estimations were in a good agreement with each other. Subgap Ohmic resistances R_{sg} were extracted from experimental data as a slope of a best-fit linear regression line for quasiparticle curves in the interval from 0 to 0.2 mV where the subgap current increases linearly with V. The R_{sg}/R_N ratio obtained after averaging over five Nb/Al-AlO_x-Nb samples with different d_{Al} and R_N (measured at 1.7 K) is given in Table 21.1 together with theoretical outputs for N-I-N, N-I-S, S-I-S, and S/N-I-S devices.

Reasonable agreement between ratio R_{sg}/R_N calculated for an S/N-I-S structure and that measured experimentally proves that, independently on the Al interlayer thickness, the distribution of transparencies across the disordered oxide layer is universal and quantitatively well describes the experimental results.

21.3 Universal Current-Voltage Characteristic for Disordered Several Nm-Thick Insulating Layers

When the thickness of a dirty insulting layer is increased up to several nanometers, the tunnel barrier in an M-I-M trilayer cannot be more described with a delta function in the transport direction and its internal structure should be taken into account in order to explain unusual bias dependence of the tunnel current which is markedly different from that predicted by the standard tunneling model [12]. This observation is ascribed to the presence of localized states within noncrystalline materials. Assuming that in this case the dominant mechanism for electronic conduction is hopping, Glazman and Matveev [13] proposed a microscopic model for charge transport across two and more localized states forming optimal conduction chains. In some experiments (see, e.g., [14], [15] it was found that the Glazman-Matveev theory does well describe experimental data for tunneling into complex oxides like manganites, independently on whether a barrier was due to an oxygen-depleted layer at the oxide surface [14], or thin insulating layers inside manganite single crystals resulted from a percolative nature of the transition between charge ordered insulating and metallic ferromagnetic states [15]. At the same time it should be noticed that the Glazman-Matveev theory which takes into account inelastic tunneling via pairs of the localized states was based on a sound-like approximation for the phonon dispersion relation valid only for extremely small wave vectors in the complex oxides. It raises the natural questions about why the theory works in the materials studied and why it is so universal. In this section we answer the questions by developing a general theoretical framework for inelastic processes arising when an electron is hopping across localized states inside thin amorphous films.

If a tunneling charge transfers classically forbidden region elastically (without energy loss), the probability of such process exponentially depends on the tunneling distance l: $D_{\text{el}} \propto \exp(-2\kappa l)$, where κ^{-1} is the localization length, and the differential conductance $G(V) = dI(V)/dV$ is proportional to V^2 [3]. But as the barrier thickness d increases, hopping along chains containing localized states is favored, since in this case it is not necessary to transfer quantum-mechanically the whole distance between the electrodes, but rather to jump from one of them to a first nano-island, then transfer to the second one and, after all hopping events, to jump to the opposite electrode. The electron jumps can be as elastic, as inelastic, with emitting a phonon of the energy ε. Due to the strong electron-phonon interaction for localized states, the latter processes which reduce the electron energy from E_1 to $E_2 = E_1 - \varepsilon$ are very important just in amorphous semiconductors. For fixed E_1 and E_2 the number of the inelastic tunneling events is proportional to the electron-phonon interaction function $\alpha^2 F(\varepsilon)$ whose amplitude is determined by α^2, a characteristic of the interaction strength, whereas the shape of the function resembles that of the phonon density of states $F(\varepsilon)$ (Wolf 2011). Then the total probability of electron inelastic tunneling through the distance l with the energy decrease from zero to ε is proportional to $\exp(-2\kappa l) \int_0^{\varepsilon} \alpha^2 F(\omega) d\omega$. Taking into account that optimal conductance chains correspond to the case when transmission probabilities of all hoppings are almost identical [12, 13], we can analytically calculate I–V curves for any microstructure of the insulating nano-scaled layer without any assumption concerning the phonon spectrum. As an example, it can be a process of inelastic tunneling across the dielectric through two localized states coupled elastically to the nearest electrodes discussed earlier [12, 13]. For this configuration we find that the inelastic contribution to the differential conductance $G_{\text{inel}}(V) \sim \exp\left(-\frac{2}{3}\kappa d\right) \int_0^{eV} \left(\int_0^{\varepsilon} \alpha^2 F(\omega) d\omega \right)^{1/3} d\varepsilon$. The factor $\exp(-2\kappa d/3)$ reflects the presence of two-step tunneling events across the barrier.

To go further, we need an exact dependence of the phonon spectrum on energy. In general, it is very complicated but for complex oxides of transition metals we are interested in, it can be approximated as $F(\varepsilon) \approx \text{const}\, \theta(\bar{\varepsilon} - \varepsilon)$ with $\bar{\varepsilon}$, the cut-

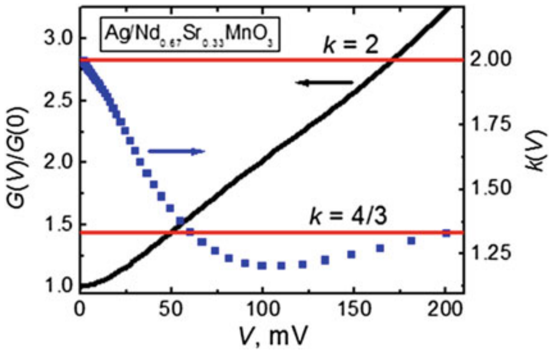


Fig. 21.1 Differential conductance and the power index k dependencies on voltage for an $Ag/NSMO$ point contact

off phonon energy (see, for example, neutron scattering data for $La_{0.625}$ $Ca_{0.375}$ MnO_3 polycrystalline samples [16]. With this approximation for a double-state configuration of the defects inside the disordered complex-oxide layer we obtain the following conductance-versus-voltage dependence $G_{\mathrm{inel}}(V) \sim V^{4/3}$ which should be universal for different complex-oxide materials. To check it, we suggest the following procedure which has no fitting parameters. If the differential conductance of an inhomogeneous thin insulating layer is a power function of the voltage bias $G(V) = G_0 + \mathrm{const} \cdot V^k$, we can find the index k and, hence, to distinguish between different transport mechanisms by calculating the function $k(V) = d \ln(G(V) - G_0)/d \ln V$.

In Fig. 21.1 we have applied the proposed procedure to experimental data for point-contact junctions formed by a sharp Ag tip with $Nd_{0.67}$ $Sr_{0.33}$ MnO_3(NSMO) thin films [14] and obtained a clear transition from elastic tunneling behavior with $k = 2$ at very low biases to inelastic one with $k \simeq 4/3$ for voltages increased up to several tens of millivolts.

21.4 Conclusions

Resuming, we have analyzed two phenomena in transport characteristics of strongly disordered oxide films: universal bimodal distribution of transparencies across subnanometer-thick layers for voltages about several millivolts and universal current-voltage characteristic of layers with increased thicknesses up to several nanometers and voltages up to several hundreds of millivolts. We believe that the universalities arise due to strong local barrier-height fluctuations caused by oxygen vacancies [17].

Just these fluctuations generate huge variations of the parameter Z in ultra-thin oxide films and formation of localized states inside thicker layers. If the transport property depends only on a single parameter and the corresponding analytical relation is mathematically simple (like Lorentzian in subnanometer-thick interlayers), the result obtained after averaging over the uniform random variable is of a general character bearing no relation to sample-specific details of particular objects and including only a limited number of macroscopic parameters or, by other words, is universal.

References

1. Kingon AI, Maria JP, Streiffer SK (2000) Alternative dielectrics to silicon dioxide for memory and logic devices. Nature 406(6799):1032–1038
2. Robertson J (2006) High dielectric constant gate oxides for metal oxide Si transistors. Rep Prog Phys 69(2):327–396
3. Wolf EL (2011) Principles of electron tunneling spectroscopy, 2nd edn. Oxford University Press, Oxford
4. Greibe T, Stenberg MPV, Wilson CM, Bauch T, Shumeiko VS, Delsing P (2011) Are "pinholes" the case of excess current in superconducting tunnel junctions? A study of Andreev current in highly resistive junctions. Phys Rev Lett 106(9):097001
5. Naveh Y, Patel V, Averin DV, Likharev KK, Lukens JE (2000) Universal distribution of transparencies in highly conductive Nb/AlO_x/Nb junctions. Phys Rev Lett 85(25):5404–5407
6. Blonder GE, Tinkham M, Klapwijk TM (1982) Transition from metallic to tunneling regimes in superconducting microconstrictions: excess current, charge imbalance, and supercurrent conversion. Phys Rev B 25(7):4515–4532
7. Melsen JA, Beenakker CWJ (1994) Reflectionless tunneling through a double-barrier NS junction. Physica B 203(3–4):219–225
8. Schep KM, Bauer GEW (1997) Transport through dirty interfaces. Phys Rev Lett 56(24):15860–15872
9. Golubov AA, Houwman EP, Gijsbertsen JG, Krasnov VM, Flokstra J, Rogalla H, Kupriyanov MY (1995) Proximity effect in superconductor-insulator-superconductor Josephson tunnel junctions: theory and experiment. Phys Rev B 51(2):1073–1089
10. Averin D, Bardas A (1995) ac Josephson effect in a single quantum channel. Phys Rev Lett 75(9):1831–1834
11. Lacquaniti V, De Leo N, Fretto M, Maggi S, Sosso A (2007) $Nb/Al-AlO_x$/Nb overdamped Josephson junctions above 4.2 K for voltage metrology. Appl Phys Lett 91(25):252505
12. Xu Y, Ephron D, Beasley MR (1995) Directed inelastic hopping of electrons through metal-insulator-metal tunnel junctions. Phys Rev B 52(4):2843–2859
13. Glazman LI, Matveev KA (1988) Inelastic tunneling across thin amorphous films. Sov Phys JETP 67(6):1276–1282
14. Svitunov VM, Leonova VN, Belogolovskii MA, Medvedev YuV, Revenko YuF, Strzhemechny YM, Hui D, Endo T (2008) Tunneling spectroscopy of manganites with nanoscale structural non-uniformities. Mod Phys Lett B 22(29):2811–2819
15. Belogolovskii M, Jung G, Markovich V, Dolgin B, Wu XD, Yuzhelevski Y (2011) Bias dependent 1/ f conductivity fluctuations in low-doped $La_{1-x}Ca_xMnO_3$manganite single crystals. J Appl Phys 109(7):073920
16. Adams CP, Lynn JW, Smolyaninova VN, Biswas A, Greene RL, Ratcliff W, Cheong S-W, Mukovskii YM, Shulyatev DA (2004) First-order nature of the ferromagnetic phase transition in (La-Ca)MnO_3near optimal doping. Phys Rev B 70(13):134414
17. Kim DJ, Choi WS, Schleicher F, Shin RH, Boukari S, Davesne V, Kieber C, Arabski J, Schmerber G, Beaurepaire E, Jo W, Bowen M (2010) Control of defect-mediated tunneling barrier heights in ultrathin MgO films. Appl Phys Lett 97(26):263502

Part VI
Statistical Physics

Chapter 22
Pursuit and Evasion with Temporal Non-locality and Stochasticity

Toru Ohira

Abstract We discuss a new aspect of an old mathematical problem of chase and escape. We consider one group chases another, called "group chase and escape", by presenting simple models. We have found that even a simple model can exhibit rich and complex behavior. The model has been extended to investigate the effects of (a) stochasticity in chasing and escaping movements, (b) reaction delays (temporal non-locality) when chasing, and (c) the conversion of caught escapees to new chasers. We show that these effects can add further complexity and result in unexpected behaviors.

22.1 Introduction

"Pursuit and Evasion" (or " Chases and Escapes ") is a traditional mathematical problem [1]. Typical questions include "How much time is needed for a chaser to catch a target?" and "What is the best escaping strategy?" There has been much mathematical interest in obtaining analytical results, so the majority of the questions have dealt with cases in which one chaser is pursuing a single escapee.We recently proposed a simple extended model called "Group Chase and Escape" [2] in which one group chases another group. This extension connects the traditional problem of "Chases and Escapes" with current interest in the collective motions of self-driven particles such as animals, insects, cars, etc [4–6].

In this chapter, we briefly present our basic model and its rather complex behaviors. Each chaser approaches its nearest escapee while each escapee moves away from its nearest chaser. Although there is no communications within groups, aggregate formations are observed both for both chasers and escapees. How these behaviors appear as a function of parameters, such as densities will be discussed.

In addition, we have extended our models in three main ways. First, we introduced a stochasticity. Players now make errors in which direction they step with some

T. Ohira (✉)
Sony Computer Science Laboratories, Inc., Tokyo, Japan

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_22,

probability. It turns out that some levels of fluctuations work better for more effective capturing. Second, we introduced a temporal non-locality in the form of a reaction delay in a chaser who is pursuing an escapee that is moving with a uniform speed in a circular path. We did not observe a complex chaser's trajectory with constant reaction delay, but distance–dependent reaction delays can cause quite complex behaviors. Finally, we report briefly on the effect of the probabilistic conversion of the captured escapees into new chasers.

22.2 Basic Model

Here, we describe our basic "Group Chase and Escape" model [2]. Essentially, it is a chase and escape problem in which one group chases another. In order to keep our extension simple, we made each chaser in a chasing group take one step toward its nearest escapee, while each escapee takes one step away from its nearest chaser. They do this independently of each other, meaning there is no communication or direct interaction among members within either the chasing or the escaping groups. We also decreed that a caught escapees be removed from the field, so that gradually the number of escapees decreases. The chase and escape finishes when all the escapees are caught and removed from the field ("complete capture").

There are various possible implementations of this conceptual model. To start with, we considered a square lattice with a periodic boundary condition and discrete step and time movements of the players. We also introduced an exclusion volume property: they cannot move if another of the same type (chaser or escapee) occupies the next location of intended motion. Also, when there are multiple choices (typically only two, due to the square lattice) for the next step, one of them is chosen with equal probability.

We have simulated the above model under various conditions [2, 3]. One of the interesting qualitative behaviors observed is a formation of aggregates by both chasers and escapees in spite of the fact that there is no direct interaction among the members of each group. In a related matter, given the initial size of the number of escapees, there exists an optimal number of chasers for effective capture. This can be seen by increasing the number of chasers with a given number of escapee, and noting the time taken to finish capturing all of the escapees. This complete capture time will decrease at a rather fast pace until it reaches the optimal number of chasers, after which it changes at a much slower rate. One of the reasons for this is the excluding volume effect (mentioned above): chasers get in each other's way. However, this is not the only cause—the very act of chasing and escaping is also a crucial factor, as such an effect is not seen if we set both chasers and escapees as groups of random walkers.

22.3 Extended Model

In this section, we discuss extension of the basic model to include stochasticity (fluctuations), temporal non-locality (delay), and conversions.

22.3.1 Effects of Stochasticity

We have extended our study to examine the effects of fluctuation in the above basic model [2, 10–12]. Specifically, we introduced errors in taking steps by players of both sides. With some probability, a chaser now takes a step in the wrong direction, thus increasing its distance from the nearest escapee. This error probability, which is also introduced in the steps of the escapees, is designed so that with the maximal error both sides become groups of random walkers, while with zero errors the model is reduced to the basic model described in the previous section.

We simulated the model with the above fluctuation error with varying probabilities and different ratios of the numbers of chasers and escapees. Increasing the error rate naturally led to a longer time for complete capture, and this is what happened when the number of chasers was relatively large. However, a rather interesting situation was observed when there were small number of chasers and escapees compared to the size of the square grid field. In this case, there exists the optimal level of fluctuation with which the time for complete capture became minimal—indeed less than not only the case in which both sides were randomly walking, but also in the case of the basic model mentioned above.

Cases that exhibit an appropriate level of fluctuation leading to "better" effects are being studied in various fields, including biology, material science, engineering and so on, under the name of "stochastic resonance" [7–9]. Our observation here can be considered as one of such examples.

22.3.2 Effects of Temporal Non-locality

Next, we consider the effects of temporal non-locality in the form of delayed reaction time on the part of the chasers [13, 14]. To examine this, we go back to one of the original one-to-one chase and escape problems in which the escapee moves in a circular path at a constant speed while the chaser moves with its velocity vector pointing to the current position of the escapee. We know that if the speed of the chaser is not as fast as the escapee, the capture is not possible, and the path of the chaser will approach to a "limit circle". The center of the two circles is the same and the ratio of the radii is the same as the ratio of the speeds of the chaser and the escapee.

Now let us consider a case in which the speeds of the chaser and escapee are the same. In this case, the chaser moves behind the escapee with the same uniform distance on the same circle. We now introduce a delay to the reaction time of the chaser. Specifically, the velocity vector of the chaser points to the past position of the escapee. We can immediately see that if we introduce this reaction time as fixed constant (fixed delay), the qualitative behavior of the chaser's motion does not change: the effect is merely an increase of the distance between the chaser and the escapee on the same circle.

However, if we set the reaction time to be proportional to the distance between the chaser and the escapee (distant dependent delay: the longer the distance, the longer the delay), the path of the chaser deviates from the circle. As we increase the rate of this proportionality, the path will become quite complex.

Delay-induced complex behaviors have been studied in various contexts. The most notable examples are the dynamical trajectories given by delay differential equations, such as the MacKay—Glass model [15]. In this model, a very simple first order differential equation with an external force term, which is a function of the delayed past state of the dynamical variable, can show various dynamics from a stable fixed point, limit cycles, and further create complex chaotic trajectories as we increase the value of the delay. Our observation here can be considered another example of delay-induced complexities.

22.3.3 Effects of Conversion

Finally, we briefly discuss the effect of conversions [16]. We extend the basic model in such a way that the caught escapee becomes a new chaser with a certain probability, while escapees can proliferate with some probability as well. The balance between these two factors again produces a non-monotonic change in the time it takes for complete capture with varying parameters. For example, if we fix the number of chasers and escapees and change the proliferation probability, there exists an optimal value to have the longest time for complete capture for a smaller value range of conversion probabilities. This finding has also been reported in a separate publication [16].

Using our model with this extension has potential for application to studies on the spread of certain epidemics such as rabies.

22.4 Discussion

We have described our recent proposal and investigation of group-based chases and escapes and are faced with the following tasks. First, we should make our model more realistically by including communication within groups or more complex chasing and escaping strategies. They reflect such cases of one group of animals chasing another, e.g., wolves hunting deer [17]. Second, we need to consider a possible application to distributed robotics or other engineering systems. For example, the human immune system includes neutrophil granulocytes that chase foreign external viruses or chemicals. Could we implement such a defense system against attacks in cyberspace by adapting the concept of chase and escape included? Finally, this type of chase and escape interactions among the constituents in groups has not been addressed in studies of physical theories. Extension of many-body theories of statistical physics for the purpose of adapting them to chase and escape-type interactions will likely pose interesting challenges in the future.

Acknowledgments The author would like to thank A. Kamimura, S. Matsumoto, and R. Nishi for their collaborative works. Thanks also go to T. Nogawa, N. Ito, K. Nishinari and J. Milton for their fruitful discussions.

References

1. Nahin PJ (2007) Chases and escapes: the mathematics of pursuit and evasion. Princeton University Press, Princeton
2. Kamimura A, Ohira T (2010a) Group chase and escape. New J Phys 12:053013
3. Matsumoto S, Nogawa T, Kamimura A, Ito N, Ohira T (2011) Dynamical aspect of group chase and escape, in nonequilibrium statistical physics today. In: Garrido PL, Marro J, Santos F (eds) Proceedings of the 11th granada seminar on computational and statistical physics. AIP Conference Proceedings, vol 1332. La Herradura, pp 226–227 (September 2010).
4. Chowdhury D, Santen L, Schadschneider A (2000) Statistical physics of vehicular traffic and some related systems. Phys Rep 329:199
5. Helbing D (2001) Traffic and related self-driven many-particle systems. Rev Mod Phys 73:1067–1141
6. Vicsek T, Zafiris A (2010a) Collective motion, ArXiv. cond-mat:1010.5017.
7. Wisenfeld K, Moss F (1995) Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs. Nature 373:33–36
8. Bulsara AR, Gammaitoni L (1996) Tuning in to noise. Phys Today 49:39
9. Gammaitoni L, Hanggi P, Jung P, Marchesoni F (1998) Stochastic resonance. Rev Mod Phys 70:223
10. Kamimura A, Matsumoto S, Ito N, Ohira T (2010b) Chase and escape in groups. In: Bandini S, Manzoni S, Umeno H, Vizzari G (eds) Proceedings of the international conference on cellular automata for research and industry. Ascoli Piceno, September 2010. Lecture notes in computer sciences, vol 6350, pp 570–579.
11. Kamimura A, Matsumoto S, Nogawa T, Ito N, Ohira T (2011) Stochastic resonance with group chase and escape. In: Deen MJ, Chen CH (eds) Proceedings of the 21st international conference on noise and fluctuations. Toronto, June 2011, pp 200–203.
12. Matsumoto S, Nogawa T, Kamimura A, Ito N, Ohira T (2010) Fluctuation effect for group chase and escape. In: Proceedings of the 16th symposium on simulations of traffic flow. Nagoya, November 2010, pp 1–4.
13. Ohira T, Kamimura A, Milton J (2011) Pursuit-escape with distance-dependent delay. In: Proceedings of 7th European nonlinear dynamics conference. Rome, July 2011, MS-11 4.
14. Milton J, Lippai J, Bellows R, Blomberg A, Kamimura A, Ohira T (2011) Visuomotor tracking tasks with delayed pursuit and escape. In: Proceedings of 8th international conference on multibody systems, nonlinear dynamics and control. Washington DC, August 2011, DETC2011-47312.

15. Mackey M, Glass L (1977) Oscillation and chaos in physiological control systems. Science 197:287–289
16. Nishi R, Kamimura A, Nishinari K, Ohira T (2012) Group chase and escape with conversion from targets to chasers. Physica A 391:337–342
17. Vicsek T (2010) Closing in on evaders. Nature 466(43):2010

Part VII
Theoretical Physics

Chapter 23
Behaviour at Ultra High Energies

Burra G. Sidharth

Abstract The Large Hadron Collider has already attained an unprecedented energy of 7 TeV. By 2013 it is expected to reach its peak energy of double this figure. We can hope that many surprises and discoveries are waiting to happen in the years to come. In this context we explore the behaviour of particles, particularly fermions at these Ultra High Energies. In particular two aspects will be touched upon: The Feshbach-Villars formulation for high energies and also considerations at the Planck length. Some new insights are explored thereby.

23.1 Introduction

The LHC in Geneva is already operating at a total energy of 7 TeV and hopefully after a pause in 2012, it will attain its full capacity of 14 TeV in 2013. These are the highest energies achieved todate in any accelerator. It is against this backdrop that it is worthwhile to revisit very high energy collisions of Fermions. We will in fact examine their behaviour at such energies.

To get further insight, let us consider the so called Feshbach-Villars formulation [4] and analyze the problem from this point of view rather than that of conventional Field theory. In this case with an elementary transformation, the equations for the components φ and χ of the Dirac wave function can be written as

B. G. Sidharth (✉)
International Institute for Applicable Mathematics and Information Sciences, Hyderabad, India

B. G. Sidharth
International Institute for Applicable Mathematics and Information Sciences, Udine, Italy

B. G. Sidharth
B.M. Birla Science Centre, Adarsh Nagar, Hyderabad 500 063, India

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_23,

$$
\begin{aligned}
i\hbar\,(\partial\phi/\partial t) &= (1/2m)\,(\hbar/i\backslash\nabla - eA/c)^2\,(\phi+\chi) \\
&\quad \left(e\phi + mc^2\right)\chi \\
i\hbar\,(\partial\chi/\partial t) &= -\,(1/2m)\,(\hbar/i\nabla - eA/c)^2\,(\phi+\chi) \\
&\quad \left(e\phi - mc^2\right)\phi
\end{aligned}
\tag{23.1}
$$

What Feshbach and Villars did was give a particle interpretation to the Klein-Gordon and Dirac equations without invoking field theory or the Dirac sea. In this case ϕ represents the "low energy" solutions, that is the normal solution and χ represents the "high energy" solutions. It must be remembered that at our usual energies it is the wave function ϕ, the so called positive energy solution that dominates, χ being of the order of v^2/c^2 of ϕ. On the other hand at very "high energies" χ the so called negative energy solution dominates. Feshbach and Villars identified these two solutions with particles and antiparticles respectively. We have

$$
\varphi = \begin{pmatrix}\phi_0 \\ \chi_0\end{pmatrix} e^{i/\hbar(p\times Et)}, \quad \phi = \phi_0(p)e^{(i/\hbar(p\times - Et)} \tag{23.2}
$$

We consider separately the positive and negative values of E (coming from (23.2)), viz.,

$$
E = \pm E_p; \quad E_p = \left[(cp)^2 + \left(mc^2\right)^2\right]^{\frac{1}{2}} \tag{23.3}
$$

The solutions associated with these two values of E are

$$
\phi_0^{(+)} = \frac{E_p + mc^2}{2(mc^2E_p)^{1/2}} \quad \chi_0^{(+)} = \frac{mc^2 - E_p}{2(mc^2E_p)^{1/2}}
$$

$$
\phi_0^{(-)} = \frac{mc^2 - E_p}{2(mc^2E_p)^{1/2}} \quad \chi_0^{(-)} = \frac{E_p + mc^2}{2(mc^2E_p)^{1/2}}
$$

For $E = \pm E_p$.

As is well known the positive solution ($E = E_p$) and the negative solution ($E = -E_p$) represent solutions of opposite charge. We also mention the well known fact that a meaningful subluminal velocity operator can be obtained only from the wave packets formed by positive energy solutions. However the positive energy solutions alone do not form a complete set, unlike in the non relativistic theory. This also means that a point description in terms of the positive energy solutions alone is not possible for the K-G (or the Dirac) equation, that is for the position operator,

$$
\delta\left(\vec{X} - \vec{X}_0\right)
$$

In fact the eigen states of this position operator include both positive and negative solutions. All this is well known (Cf.ref.(Feshback 1958)).

This matter was investigated earlier by Newton and Wigner too (Newton and Wigner [8] from a slightly different angle. Some years ago the author revisited this aspect from yet another point of view [20] and showed that this is symptomatic of noncommutativity which is exhibited by

$$[x_\imath,\ x_j] = O(l^2) \cdot \Theta_{\imath j}$$

and is related to spin and extension. The noncommutative nature of spacetime has been a matter of renewed interest in recent years particularly in Quantum Gravity approaches. At very high energies, it has been argued that [25] there is a minimum fuzzy interval, symptomatic of a non commutative spacetime, so the usual energy momentum relation gets modified and becomes [26]

$$E^2 = p^2 + m^2 + \alpha l^2 p^4 \tag{23.4}$$

the so called Snyder-Sidharth Hamiltonian [5, 11, 27]. It has been argued that for fermions $\alpha > 0$ while α can be <0 for bosons. Using (23.4) it is possible to deduce the ultrarelativistic Dirac equation [17, 21]

$$(D + \beta l p^2 \gamma^5)\psi = 0 \tag{23.5}$$

$\beta = \sqrt{\alpha}$. In (23.5) D is the usual Dirac operator while the extra term appears due to the new dispersion relation (23.4).

As indicated above α is positive. It is known that [29], in this case Eq. (23.5) can be written in Hamiltonian form

$$-\gamma_0 p^0 \psi = (\bar{D} + \imath \alpha l p^2 \gamma_5)\psi \tag{23.6}$$

where $\bar{D} \equiv \sum_\imath \gamma^\imath p_\imath$. Further it is well known that the Hamiltonian is given by

$$H = \imath \gamma_5 \overrightarrow{\sum} \cdot \vec{p} = \imath \gamma_5 |\vec{p}||s(\vec{p})| \tag{23.7}$$

It can be seen from (23.7) that the Dirac particle acquires an additional mass. However what is very interesting is that the extra mass term is not invariant under parity owing to the presence of γ_5. Indeed as we know from the theory of Dirac matrices

$$P\gamma_5 = -P\gamma_5 \tag{23.8}$$

In the case of a massless Dirac particle, it was argued that this leads to the mass of the neutrino [28].

Thus the mass m gets split into $m + m'$ and $m - m'$ with two states, Ψ_L and Ψ_R. Remembering that a dominant ϕ and a dominant χ respectively represent particle and antiparticle in this Feshbach-Villars formulation and also remembering that under reflection, as is well known,

$$\phi \to \phi, \quad \chi \to -\chi \tag{23.9}$$

we can see that this means that the particle and antiparticle have different masses, namely $m + m'$ and $m - m'$. Indeed this conclusion was anticipated earlier [30]. The difference would be minute but in principle can be observed. Already there have been reports of such mass asymmetry being observed in the MINOS Fermi Lab experiment with neutrinos and anti neutrinos [13]. What the MINOS team recorded was a difference in the Δm^2 value for neutrinos and anti neutrinos by as much as forty percent. It is expected that more definitive results would be available by 2012.

It has been pointed out that the fact that equations like (23.7) and the following applied to neutrinos which are massless suggests one (or more) neutrinos. This is brought out more clearly in the above. Remarkably there seems to be very recent confirmation of such an extra or sterile neutrino (Roe).

The above discussion brings out ultra high energy effects in Fermionic behavior. Already Eq. (23.4) shows modifications to Lorentz symmetry, as has been discussed in detail in several places, for example (Cf.ref. [7, 26] and references therein). This exposes the limits of strict special relativistic considerations.

23.2 Extra Relativistic Effects

It has just been announced that the OPERA (Oscillation Project with Emulsion Tracking Apparatus) experiment, 1,400 m underground in the Gran Sasso National Laboratory in Italy has detected neutrinos travelling faster than the speed of light, which has been a well acknowledged speed barrier in physics. This limit is 299,792,458 m/s, whereas the experiment has detected a speed of 299,798,454 m/s. In this experiment neutrinos from the CERN Laboratory 730 km away in Geneva were observed. They arrived 60 ns faster than expected, that is faster than the time allowed by the speed of light. The experiment has been measured to 6σ level of confidence, which makes it a certainty (Adam) and has been repeated again. However it is such an astounding discovery that the OPERA scientists would like further confirmation from other parts of the world. In 2007 the MINOS experiment near Chicago did find hints of this superluminal effect.

It must be reported that the author had predicted such deviations from Einstein's Theory of Relativity, starting from 2,000 (Cf.eq.(23.4)).

(23.4) shows that the energy at very high energies for fermions is greater than that given by the relativity theory so that effectively the speed of the particle is slightly greater than that of light. For example, if in the usual formula, we replace c by $c + c'$, then, comparing with the above we would get:

$$c' = \alpha l^2 p^4 \cdot [4m^2 c^3 + 2p^2 c]^{-1}$$

The difference is slight, but as can be seen is maximum for the lightest fermions, viz., neutrinos which are in any case already travelling with the velocity c. We could also argue that the extra term in the Snyder-Sidharth Hamiltonian can contribute partly to an oscillating mass of the neutrino oscillations and partly to a fluctuating super luminal velocity.

23.3 Ultra High Energy Particles

Let us look at all this differently. Following Weinberg [32] let us suppose that in one reference frame S an event at x_2 is observed to occur later than one at x_1, that is, $x_2^0 > x_1^0$ with usual notation. A second observer S' moving with relative velocity $\vec{v}$ will see the events separated by a time difference

$$x_2^0 - x_1^0 = \Lambda_\alpha^0(v)(x_2^\alpha - x_1^\alpha)$$

where $\Lambda_\alpha^\beta(v)$ is the "boost" defined by or,

$$x_2'^0 - x_1'^0 = \gamma(x_2^0 - x_1^0) + \gamma\vec{v} \cdot (x_2 - x_1)$$

and this will be negative if

$$v \cdot (x_2 - x_1) < -(x_2^0 - x_1^0) \tag{23.10}$$

We now quote from Weinberg [32]:

"Although the relativity of temporal order raises no problems for classical physics, it plays a profound role in quantum theories. The uncertainty principle tells us that when we specify that a particle is at position x_1 at time t_1, we cannot also define its velocity precisely. In consequence there is a certain chance of a particle getting from x_1 to x_2 even if $x_1 - x_2$ is spacelike, that is, $|x_1 - x_2| > |x_1^0 - x_2^0|$. To be more precise, the probability of a particle reaching x_2 if it starts at x_1 is non negligible as long as

$$(x_1 - x_2)^2 - (x_1^0 - x_2^0)^2 \leq \frac{\hbar^2}{m^2} \tag{23.11}$$

where $\hbar$ is Planck's constant (divided by 2π) and m is the particle mass. (Such space-time intervals are very small even for elementary particle masses; for instance, if m is the mass of a proton then $\hbar/m = w \times 10^{-14}$ cm or in time units 6×10^{-25} s. Recall that in our units $1s = 3 \times 10^{10}$) cm. We are thus faced again with our paradox; if one observer sees a particle emitted at x_1, and absorbed at x_2, and if

$(x_1 - x_2)^2 - (x_1^0 - x_2^0)^2$ is positive (but less than $\hbar^2/m^2$), then a second observer may see the particle absorbed at x_2 at a time t_2 before the time t_1 it is emitted at x_1".

To put it another way, the temporal order of causally connected events cannot be inverted in classical physics, but in Quantum Mechanics, the Heisenberg Uncertainty Principle leaves a loop hole.

As can be seen from the above, the two observers S and S' see two different events, viz., one sees, in this example the protons while the other sees neutrons. Moreover, this is a result stemming from (23.11), viz.,

$$0 < (x_1 - x_2)^2 - (x_1^0 - x_2^0)^2 (\leq \frac{\hbar^2}{m^2}) \tag{23.12}$$

The inequality (23.12) points to a reversal of time instants $(t_1, \ t_2)$ as noted above. However, as can be seen from (23.12), this happens within the Compton wavelength.

We now observe that in the above formulation for the wave function

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix},$$

where, as noted, ϕ and χ are, for the Dirac equation, each two spinors. ϕ (or more correctly ϕ_0) represents a particle while χ represents an antiparticle. So, for one observer we have

$$\chi \sim 0 \tag{23.13}$$

and for another observer we can have

$$\phi \sim 0 \tag{23.14}$$

that is the two observers would see respectively a particle and an antiparticle. This would be the same for a single observer, if for example the particle's velocity got a boost so that (23.14) rather than (23.13) would dominate after sometime.

Interestingly, just after the Big Bang, due to the high energy, we would expect, first (23.14) that is antiparticles to dominate, then as the universe rapidly cools, particles and antiparticles would be in the same or similar number as in the Standard Model, and finally on further cooling (23.13) that is particles or matter would dominate.

Finally we now make two brief observations, relevant to the above considerations. Latest results in proton-antiproton collisions at Fermi Lab have thrown up the Bs mesons which in turn have decayed exhibiting CP violations in excess of the predictions of the Standard Model, and moreover this seems to hint at a new rapidly decaying particle. Furthermore, in these high energy collisions particle to antiparticle and vice versa transformations have been detected.

23.4 Ultra High Energy Particles

Some years ago [24], we explored some intriguing aspects of gravitation at the micro and macro scales. We now propose to tie up a few remaining loose ends. At the same time, this will give us some insight into the nature of gravitation itself and why it has defied unification with other interactions for nearly a century. For this, our starting point is an array of n Planck scale particles.

As discussed in detail elsewhere, such an array would in general be described by (Jack Ng and Van Dam [7]

$$l = \sqrt{n \Delta x^2} \tag{23.15}$$

$$ka^2 \equiv k \Delta x^2 = \frac{1}{2} k_B T \tag{23.16}$$

where k_B is the Boltzmann constant, T the temperature, r the extent and k is the analogues of the spring constant given by

$$\omega_0^2 = \frac{k}{m} \tag{23.17}$$

$$\omega = \left(\frac{k}{m} a^2 \right)^{\frac{1}{2}} \frac{1}{r} = \omega_0 \frac{a}{r} \tag{23.18}$$

We now identify the particles with Planck masses and set $\Delta x \equiv a = l_P$, the Planck length. It may be immediately observed that use of (23.17) and (23.16) gives $k_B T \sim m_P c^2$, which ofcourse agrees with the temperature of a black hole of Planck mass. Indeed, Rosen [15] had shown that a Planck mass particle at the Planck scale can be considered to be a Universe in itself with a Schwarzchild radius equalling the Planck length.

Whence the mass of the array is given by

$$m = m_P / \sqrt{n} \tag{23.19}$$

while we have,

$$l = \sqrt{n} l_P, \tau = \sqrt{n} \tau_P, \tag{23.20}$$

$$l_P^2 = \frac{\hbar}{m_P} \tau_P$$

In the above $m_P \sim 10^{-5}$ g, $l_P \sim 10^{-33}$ cm and $\tau_P \sim 10^{-42}$ s, the original Planck scale as defined by Max Planck himself. We would like the above array to represent a typical elementary particle. Then we can characterize the number n precisely. For this we use in (23.19) and (23.20)

$$l_P = \frac{2' G m_P}{c^2} \tag{23.21}$$

which expresses the well known fact that the Planck length is the Schwarzchild radius of a Planck mass black hole, following Rosen. This gives

$$n = \frac{lc^2}{Gm} \sim 10^{40} \tag{23.22}$$

where l and m in the above relations are the Compton wavelength and mass of a typical elementary particle and are respectively $\sim 10^{-12}$ cm and 10^{-25} g respectively.

Before coming to an interpretation of these results we use the well known result alluded to that the individual minimal oscillators are black holes or mini Universes as shown by Rosen [15]. So using the Beckenstein temperature formula for these primordial black holes [16], that is

$$kT = \frac{\hbar c^3}{8\pi G m}$$

we can show that

$$Gm^2 \sim \hbar c \tag{23.23}$$

We can easily verify that (23.23) leads to the value $m = m_P \sim 10^{-5}$ g. In deducing (23.23) we have used the typical expressions for the frequency as the inverse of the time—the Compton time in this case and similarly the expression for the Compton length. However it must be reiterated that no specific values for l or m were considered in the deduction of (23.23).

We now make two interesting comments. Cercignani and co-workers have shown [2, 3] that when the gravitational energy becomes of the order of the electromagnetic energy in the case of the Zero Point oscillators, that is

$$\frac{G\hbar^2\omega^3}{c^5} \sim \hbar\omega \tag{23.24}$$

then this defines a threshold frequency ω_{max} above which the oscillations become chaotic. In other words, for meaningful physics we require that

$$\omega \leq \omega_{max}.$$

where ω_{max} is given by (23.24). Secondly as we can see from the parallel but unrelated theory of phonons [6, 12], which are also bosonic oscillators, we deduce a maximal frequency given by

$$\omega_{max}^2 = \frac{c^2}{l^2} \tag{23.25}$$

In (23.25) c is, in the particular case of phonons, the velocity of propagation, that is the velocity of sound, whereas in our case this velocity is that of light. Frequencies greater than ω_{max} in (23.25) are again meaningless. We can easily verify that using (23.24) in (23.25) gives back (23.23). As $\hbar c = 137e^2$, in a Large Number sense, (23.23) can also be written as,

$$Gm_P^2 \sim e^2$$

That is, (23.23) expresses the known fact that at the Planck scale, electromagnetism equals gravitation in terms of strength.

In other words, gravitation shows up as the residual energy from the formation of the particles in the universe via Planck scales particles.

The scenario which emerges is the following. Analogous to Prigogine cosmology [10, 31], from the dark energy background, in a phase transition Planck scale particles are suddenly created. These then condense into the longer lived elementary particles by the above process of forming arrays. But the energy at the Planck scales manifests itself as gravitation, thereafter.

We will further discuss this in the next section.

23.5 Discussion

Equation (23.22) can also be written as

$$\frac{Gm}{lc^2} \sim \sqrt{N} \tag{23.26}$$

where $N \sim 10^{80}$ is the Dirac Large Number, viz., the number of particles in the universe. There are two remarkable features of (23.22) or (23.26) to be noted. The first is that it was deduced as a consequence in the author's 1997 cosmological model [22]. In this case, particles are created fluctuationally from the background dark energy. The model predicted a dark energy driven accelerating universe with a small cosmological constant. It may be recalled that at that time the prevailing paradigm was exactly opposite—that of a dark matter constrained decelerating universe.

As is now well known, shortly thereafter this new dark energy driven accelerating universe with a small cosmological constant was confirmed conclusively through the observations of distant supernovae. It may be mentioned that the model also deduced other inexplicable relations like the Weinberg formula that relates the microphysical constants with a large scale parameter like the Hubble Constant:

$$m \approx \left(\frac{H\hbar^2}{Gc}\right)^{\frac{1}{3}} \tag{23.27}$$

While (23.27) has been loosely explained away as an accidental coincidence Weinberg [32] himself emphasized that the mysterious relation is in fact unexplained. To quote him, "In contrast (this) relates a single cosmological parameter (the Hubble Constant) to the fundamental constants $\hbar$, G, c and m and is so far unexplained."

The other feature is that (23.26) like (23.27) expresses a single large scale parameter viz., the number of particles in the universe or the Hubble constant in terms of purely microphysical parameters.

As we saw the scenario is similar to the Prigogine cosmology in which out of what Prigogine called the Quantum Vacuum, or what today we may call Dark Energy background, Planck scale or Planck mass are created in a phase transition, very similar to the formation of Benard cells [9]. The energy at the Planck scale, given by (23.24) then gets distributed in the universe—amongst all the particles, as the Planck particles form these various elementary particles according to Eqs. (23.15) to (23.20). This is brought out by the fact that Eq. (23.26) can also be written as the well known Eddington formula:

$$Gm^2/e^2 \sim \frac{1}{\sqrt{N}} \tag{23.28}$$

which was believed to be another ad hoc coincidence unrelated to (23.27). Equation (23.28) shows how the gravitational force over the cosmos is weak compared to the electromagnetic force. In other words the initial "gravitational energy" on the formation of the Planck scale particles, that is (23.23) is distributed amongst the various particles of the universe [23]. From this point of view while l, m, c etc. are indeed microphysical constants as Dirac characterized them, G is not. It is related to the Large Scale cosmos through the Dirac Number N of particles in the universe. This would also explain the Weinberg puzzle: In this case in Eq. (23.27), there are the large scale parameters namely G and H on right side of the equation.

Once we recognize this, we can easily see that unlike what was thought previously, the Weinberg formula (23.27) is in fact the same as the Dirac formula (23.28). To see this, we use in (23.27) two well known relations from cosmology (Cf.eg. [32]), viz.,

$$R \sim \frac{GM}{c^2} \text{ and } M = Nm$$

where R is the radius of the universe $\sim 10^{28}$ cm, M its mass $\sim 10^{55}$ g and m is as before the mass of a typical elementary particle. Then (23.27) will reduce to (23.28). Thus, there is only one relation—(23.27) or (23.28), and they express the fact that rather than being a microphysical parameter, G rather than representing a fundamental interaction is related to the large scale cosmos via either of these equations.

It must be observed that this conclusion resembles that of Sakharov [18], for whom Gravitation was a secondary force like elasticity.

References

1. Adam T et al, arXiv 1109.4897
2. Cercignani C, Galgani L, Scotti A (1972) Phys Lett 38A:403
3. Cercignani C (1998) Found Phys Lett 11(2):189–199
4. Freshbach H, Villars F (1958) Rev Mod Phys 30(1):24–45
5. Glinka LA (2008) Apeiron 2, arXiv hep-ph 0812.0551.
6. Huang K (1975) Statistical mechanics. Wiley Eastern, New Delhi, p 75ff.
7. Jack Ng Y, Van Dam H (1994) Mod Phys Lett A 9(4):335–340.
8. Newton TD, Wigner EP (1949) Reviews of modern physics 21(3):400–405
9. Nicolis G, Prigogine I (1989) Exploring complexity. W. H, Freeman, New York, p 10
10. Prigogine I (1996) The End of Certainty. The Free Press, New York, p172ff.
11. Raoelina A, Rakotonirina C (2011) EJTP 8(25):691–696
12. Reif F (1965) Fundamentals of statistical and thermal physics. McGraw-Hill Book Co., Singapore
13. Report (2010) Physics World, IOP June 18, 2010.
14. Roe B http://physlinks.com, to appear in Phys Rev Lett
15. Rosen N (1993) Int J Th Phys. 32(8):1435–1440
16. Ruffini R, Zang LZ (1983) Basic concepts in relativistic astrophysics world scientific, Singapore, p 111ff.
17. Sahoo S (2010) Ind J Pure Appld Phys 48, October 2010, pp 691–696.
18. Sakharov AD (1968) Soviet physics-doklady. 12(11):1040–1041
19. Schweber SS (1961) An introduction to relativistic quantum field theory harper and row, New York.
20. Sidharth BG (2002) Foundation of physics letters 15:(5) pp 501ff.
21. Sidharth BG (2005a), Int J Mod, Phys E 14(6):927ff.
22. Sidharth BG (2005b) The Universe of Fluctuations Springer, Netherlands.
23. Sidharth BG (2005) Found. Phys Lett 18(4):393ff.
24. Sidharth BG (2006) Int J Mod PhysA. 21(31):6315
25. Sidharth BG (2008a) Thermodynamic Universe World Scientific, Singapore and other references therein.
26. Sidharth BG (2008b) Foundation of Physics 38(8):695–706
27. Sidharth BG (2008c), Foundation of, Physics 38(1):89ff.
28. Sidharth BG (2010b) Int J Mod Phys E. 19(1):79–91
29. Sidharth BG (2010c) Int J Mod PhysE. 19(11):1–8
30. Sidharth BG (2010d) EJTP 7(24), July 2010.
31. Tryon EP (1973) Is the Universe a Vacuum Fluctuation? in. Nature. 246:396–397
32. Weinberg S (1972) Gravitation and Cosmology John Wiley & Sons, New York, p 61ff.

Chapter 24
Toward "Ghost Imaging" with Cosmic Ray Muons

Milena D'Angelo, Augusto Garuccio, Franco Romano, Francesco Di Lena, Marco D'Incecco, Roberto Moro, Antonietta Regano and Giuliano Scarcelli

Abstract Optical ghost imaging is a remote imaging technique that exploits either the correlations between light beams/entangled photon pairs, or the Hanbury-Brown Twiss [1, 2] effect typical of chaotic light sources. Is it possible to implement ghost imaging with massive particles? The Extreme Energy Events (EEE) project [3] offers a platform for attempting to answer this question. Our analysis is based on the experimental data taken in L'Aquila by two distant EEE muon telescopes [4, 5]. Interestingly, muons from cosmic ray showers exhibit spatio-temporal correlations that offer the possibility to evaluate the feasibility of ghost imaging with massive particle.

24.1 Introduction

Ghost imaging is an optical technique that aims at gathering information on a distant object without necessity of employing imaging optics or high-resolution detectors near the object [6–10]. This is achieved by using two correlated beams of light:

M. D'Angelo (✉) · A. Garuccio · F. Romano
Dipartimento Interateneo di Fisica "M. Merlin" dell'Universita degli Studi e del Politecnico di Bari and INFN Sezione di Bari, Bari, Italy

F. Di Lena
Dipartimento Interateneo di Fisica "M. Merlin" dell'Universita degli Studi e del Politecnico di Bari, Bari, Italy

M. D'Incecco
INFN Laboratori Nazionali del Gran Sasso, Assergi, AQ, Italy

R. Moro · A. Regano
Museo Storico della Fisica, Centro Studi e Ricerche E. Fermi, Roma, Italy

G. Scarcelli
Harvard Medical School, Wellman Center for Photomedicine, Massachusetts General Hospital, Boston, MA, USA

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_24,

one (probe) interacts with the distant object and is revealed by a "bucket" detector, with no spatial resolution; the other (reference), in a local laboratory, goes through imaging optics and is detected by a high spatial resolution detector. The correlation between the two beams allows revealing the structure of the object, remotely, from a coincidence measurement between the two detectors. Due to this property, optical ghost imaging has recently emerged as a promising low-light-level remote sensing tool [11–13].

The main idea of the present paper is to trace the way toward the implementation of the ghost imaging protocol using correlated massive particles instead of light beams. Muons generated by cosmic rays are an interesting source to study in the context of ghost imaging: As naturally available deeply penetrating particles characterized by an extremely small De Broglie wavelength, they are promising candidates for long distance high resolution ghost imaging.

The project Extreme Energy Events (EEE) [3] offers a platform for studying the feasibility of ghost imaging with cosmic ray muons. Muon "telescopes" composed of three Multigap Resistive Plate Chambers (MRPC) [4, 5, 14], have been built and installed in many high schools across Italy. Some of them are already in operation and coincidence detection between muons have been measured in two neighbor schools of L'Aquila, 180 m apart [15, 16]. Our analysis is based on the experimental data taken for 9 days in L'Aquila.

In this paper, we start with a brief review of quantum imaging, by introducing the ghost imaging experiments based on both entangled [6–8] and separable systems of photons [17–21] as well as the ones based on chaotic light [9, 10, 18, 22–27]. We then introduce some basics elements of the EEE project—namely, the detectors employed and the coincidence data so far collected [15, 16]—and enter in the core of our analysis: Extending the quantum imaging schemes to muons from cosmic rays. After studying the spatio-temporal correlation characterizing the detected muon pairs, we present a preliminary study of the feasibility of muon ghost imaging; this is an essential step toward understanding the potentialities of this naturally available source for practical applications such as remote sensing. The presented results represent the first analysis of ghost imaging with massive particles and may pave the way for the extension to particles of many other intriguing quantum optical phenomena involving classical and non-classical correlations. From a practical standpoint, the natural abundance of cosmic ray muons on earth, their long-range correlation, extremely small De Broglie wavelength and high penetrating ability would suggest this protocol as a viable way to perform very long-distance high-resolution remote imaging.

24.2 Ghost Imaging with Photons

The first ghost imaging experiment was realized in the mid 1990s by Pittman et al. [8], following the theoretical proposal of Klyshko [6, 7]. By taking advantage of the strong correlation characterizing signal-idler photon pairs generated by Spontaneous

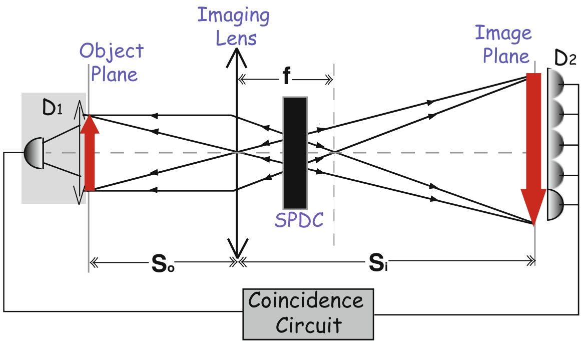


Fig. 24.1 Schematic representation of the unfolded experimental setup to observe ghost imaging with SPDC photon pairs: Object and imaging lens are illuminated by the signal photons, while the idlers propagate in free space; a ghost image appears when counting coincidences between the fixed bucket detector D_1, placed behind the object, and the scanning point-like detector D_2, placed in the "ghost" image plane as defined by the Gaussian two-photon thin lens equation

Parametric Down-Conversion (SPDC) [28], Pittman et al. proved the possibility of reproducing the ghost image of an object, remotely. The expression "ghost" was introduced to emphasize the very peculiar nature of the phenomenon: A mask (object) is inserted in front of a bucket detector, which simply counts the SPDC signal photons transmitted by the object; the image of the mask is retrieved by recording the joint detection events of the signal-idler pairs while scanning a distant photon counting detector in the two-photon image plane, as defined on the idler beam side by the two-photon Gaussian thin lens equation:

$$\frac{1}{s_o} + \frac{1}{s_i} = \frac{1}{f}, \tag{24.1}$$

where f is the focal length of the lens, s_o the object-lens distance, and s_i is the "ghost" lens-image distance, given by the sum of the distances from the image to the source (i.e., the SPDC crystal) and from the source to the lens, as shown in the unfolded experimental setup of Fig. 24.1. The single counting rates at both detectors are always fairly constant.

Seven years after its first discovery, an intense debate [17–20, 29], was opened following an experimental work by Bennink et al. [21], who raised the question whether or not ghost imaging could be reproduced by classically correlated beams of light; Ref. [18] contains a summary of this debate. A schematic representation of the experimental setup employed by Bennink, et al. [21] is reported in Fig. 24.2: Pairs of light beams classically correlated in momentum are focused by two separate lenses, the object is inserted in the focal plane of one lens, and coincidence counts are recorded between a bucket detector behind the object and a high-resolution detector placed in the focal plane of the other lens.

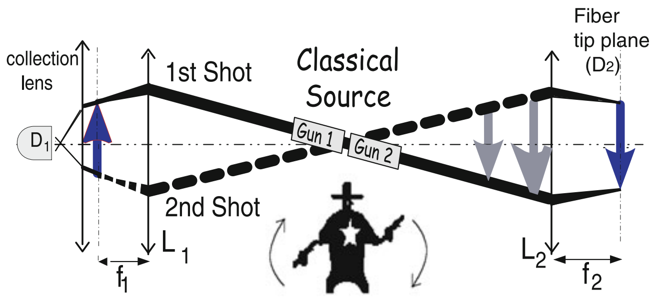


Fig. 24.2 Schematic representation of the unfolded experimental setup for simulating ghost imaging with pairs of light beams classically correlated in momentum; the focal planes of the two lenses are required to transform the momentum correlation into the required "position" correlation

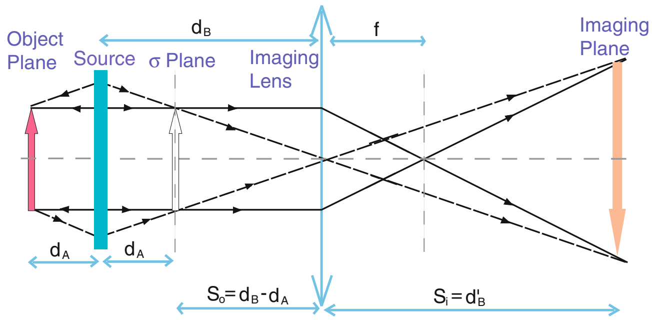


Fig. 24.3 Schematic representation of the unfolded experimental setup for observing ghost imaging with chaotic light. A lens-less ghost image can be obtained by placing the high-resolution detector in the σ plane, whose distance from the source is equal to the object-source distance

One of the most interesting results that came out of this discussion is the possibility of producing ghost images by replacing SPDC with chaotic/thermal radiation [9, 10]: Similar to the entangled two-photon case, a two-photon Gaussian thin lens equation was found for this source [10]. The experimental setup is shown in Fig. 24.3. The most evident difference with respect to the entangled case was the existence, for chaotic light, of a constant background noise accompanying the ghost image. The interpretation of the effect was soon found in terms of quantum interference, namely, coherent superposition of indistinguishable two-photon probability amplitudes [24, 27]: the thermal ghost image was shown to exist not only when employing an imaging lens, but also in a lens-less setup, provided the object-source distance is exactly equal to the image-source distance.

Beside its fundamental interest, quantum imaging has inspired several practical applications, from metrology [30] to low-light-level remote sensing [11–13].

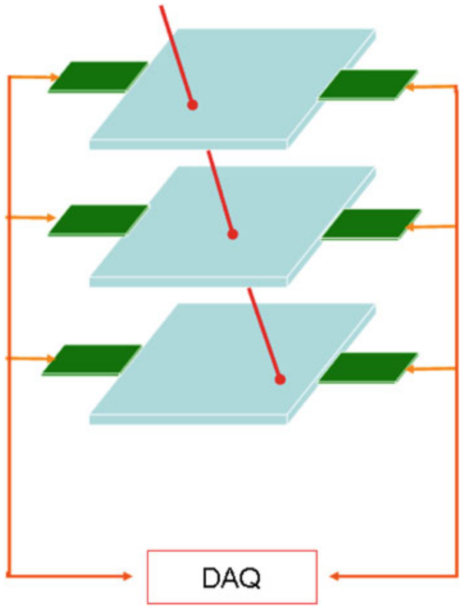


Fig. 24.4 Sketch of the telescope employed in the EEE project. Each plane represents a MRPC; the hit points on the MRPCs allow reconstructing the track of the detected muon

24.3 Cosmic Ray Muons and the EEE Telescopes

The Extreme Energy Events (EEE) Project [3] aims at studying the extremely high energy cosmic rays by means of muon detectors (also called telescopes) distributed over an area of about 10^6 km^2, in Italy. In fact, the goal is to detect the muon component of Extensive Air Showers (EAS) by measuring coincidence events between distant telescopes. Each telescope consists of three 50 cm apart Multigap Resistive Plate Chamber (MRPC) [14], gas detectors with an active area of about 2 m^2 characterized by both time and position resolution, as well as tracking capability. As described in Ref. [15], the electric signal generated by the passage of a particle through the detector is collected by one of the 24 pick-up electrodes (160 cm-long copper strips with a pitch of 3.2 cm) mounted on each MRPC. The MRPC efficiency is around 95 % and the time resolution is about 100 ps; the absolute time is recorded by a GPS, whose resolution is around 60 ns ($\sim\sqrt{2}\sigma_{GPS}$). The hit strip and the difference between the signal arrival times at the strip ends enable reconstruction of the particle impact point (i.e., its $x - y$ coordinates), with a spatial resolution of about 2 cm [4, 5]. The signals detected by each MRPC are collected only when a triple coincidence of the MRPCs occurs; a triple coincidence event in the telescope thus identifies the track of the detected particle, as shown in Fig. 24.4. Data processing allows reconstructing the muon direction with an angular resolution of about 2° [4, 5]; the acceptance of the EEE telescopes is 39° in the plane perpendicular to the copper strips and 58° in the orthogonal direction.

The present paper is based on the data taken for 9 effective days by two telescopes installed in two neighbor High Schools in L'Aquila, 180 m apart [15, 16]. The data presented in Ref. [15] are divided in three sets: (1) single track-single track (S-S) coincidences, where only a single track is present in each telescope; (2) single track-multihit (S-M) coincidences, where a single track in one telescope is accompanied by a multiple event, with a number of hits ≥ 4, in the other telescope; (3) multihit-multihit (M-M) coincidences, where high-multiplicity events occur at both telescopes. In this paper we only consider the S-S events, namely, muon pairs detected at the two Schools, which we shall indicate as station A and B.

24.4 Toward Ghost Imaging with Cosmic Ray Muons

The existence of temporal correlation is at the heart of any ghost imaging, but of course either angular or momentum correlation are also required for guarantying the position correlation implicit in ghost images.

Let us start by analyzing the temporal correlation between muons detected at the two Schools. The histogram of the time differences $(t_A - t_B)$ between muons detected by the two telescopes is characterized by a constant background and a peak centered around $t_A - t_B = 0$. Both the width of the peak and the constant background can be optimized by correcting the time differences for the average inclination $(\theta' = (\theta_A + \theta_B)/2)$ of the detected muon pairs with respect to the line joining the two telescopes (having length L); the correction is implemented by replacing the detection time differences $t_A - t_B$ with [15, 16]: $\Delta t = t_A - t_B \pm L\cos(\theta')/c$. The temporal distribution of the corrected coincidence counts is reported in Fig. 24.5; the Gaussian fit $a\exp[-(x-\mu)^2/2\sigma^2] + b$ gives for the peak visibility $V = 28\,\%$, for the peak width $\sigma = 220\,\text{ns}$, in agreement with the results presented in [15, 16].

The events falling within the observed temporal correlation peak are supposed to be mostly due to muon pairs coming from the same air shower, while the constant background is due to muons coming from different air showers. This hypothesis is strengthen by the high visibility ($V = 93\,\%$) acquired by the temporal correlation peak as soon as one selects the events characterized by almost parallel muon tracks, namely, $\alpha < 5°$ where α is the angle between muons detected at the two stations, as shown in Fig. 24.6. Also the peak width is somewhat reduced by the parallel track condition ($\sigma = 180\,\text{ns}$). The highly improved visibility of this new temporal peak indicates the existence of a strong angular correlation between muon pairs belonging to the same air shower. In addition, this result indicates that the nature of the observed correlation is certainly not predominantly chaotic; in fact, chaotic identical muons propagating in the same spatial mode (such as the one selected by imposing $\alpha < 5°$) would produce a fermionic HBT-type dip in the temporal histogram.

In order to study the angular correlation between cosmic ray muons, we define a coincidence time window centered around $\Delta t = 0$ and having total width approximately equal to the peak base (i.e., $4\sigma \approx 900\,\text{ns}$), and compare the angular distribution of coincident muon pairs with the angular distribution of muon pairs detected

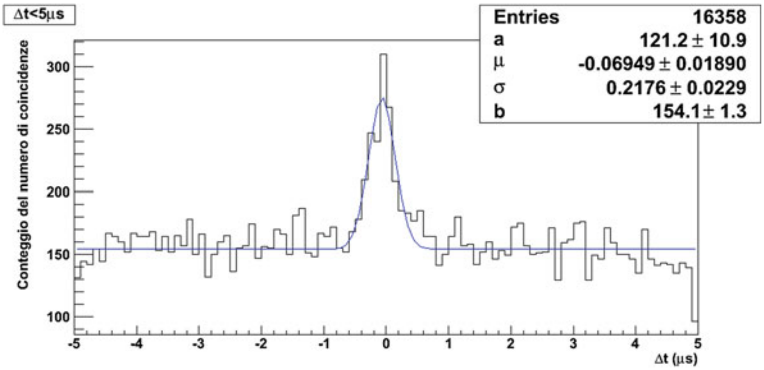


Fig. 24.5 Distribution of the muon pairs detected at the two Schools, A e B, as a function of the “corrected” detection time difference Δt. Channels are 100 ns wide

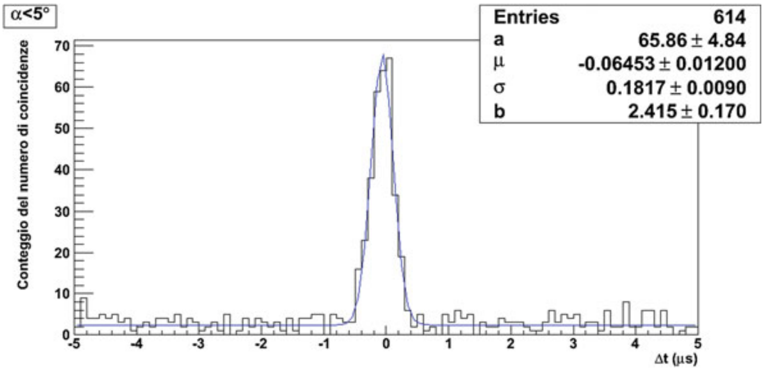


Fig. 24.6 Distribution of parallel muons detected at the two Schools, A e B, as a function of the “corrected” detection time difference Δt. Channels are 100 ns wide. The condition of parallel tracks is imposed by selecting the angle between the tracks (α) to be smaller than 5°

outside the coincidence window (i.e., pairs of independent muons). The results are shown in Fig. 24.7, where $\theta_{A,B}$ represents the detection angle, at station A and B, respectively, with respect to the vertical direction; a common reference frame (x, y) has been defined for the two stations A and B in a plane parallel to the MRPC planes. The angular correlation (as opposed to anti-correlation) gives rise to the distribution of the coincident muon pairs around the diagonal (as opposed to the anti-diagonal) of the $(\theta_{Ai}, \theta_{Bi})$ planes, with $i = x, y$, as clearly appears from the two plots in the left column of Fig. 24.7. The symmetric distribution of muon pairs detected outside the coincidence window (right column in Fig. 24.7) indicates that independent muon pairs are neither correlated nor anti-correlated; the angular correlation characterizing coincident muon pairs is thus a pure second order effect (i.e., it is not a trivial projection of the angular distribution of independent muon pairs).

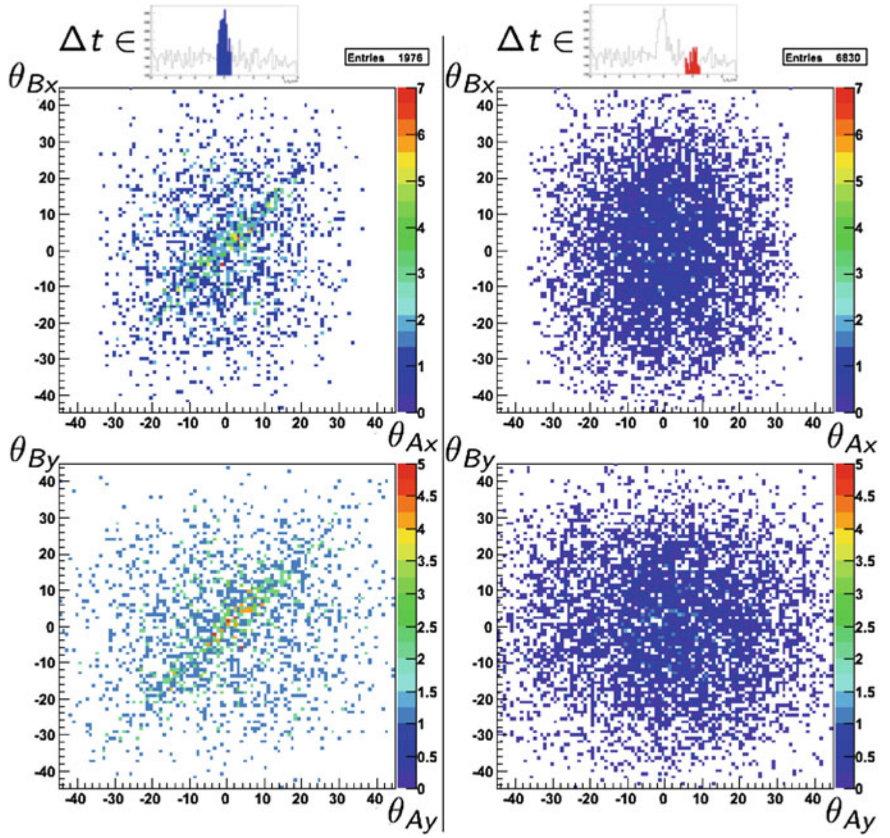


Fig. 24.7 *Left* Distribution of the coincident muon pairs as a function of their track inclinations with respect to the vertical direction, in two orthogonal directions, x (*up*) and y (*down*), in a common reference frame for the two stations A and B. *Right* Same plot for independent muon pairs, namely, muon pairs detected at the two stations far away from the temporal coincidence window. Both histograms are made of 100×100 channels

The next step of is to study the possibility of extending quantum imaging schemes to cosmic ray muons. In this perspective, we exploit the discovered angular correlation to perform a preliminary feasibility study of ghost imaging with cosmic ray muons: We simulate the presence of two lenses placed on top of the two telescopes (A and B), as schematically represented in Fig. 24.8, and study the position-position correlation between the focal planes of the two lenses. The possibility of simulating the presence of the two lenses comes from the knowledge of the reconstructed muon tracks: The ability of the EEE telescopes to reconstruct the tracks of the detected muons gives the angular information required to simulate the deviation a muon would have experienced if a lens was put along its path toward the telescope. In particular, the positions of the detected muons in the focal planes of the simulated

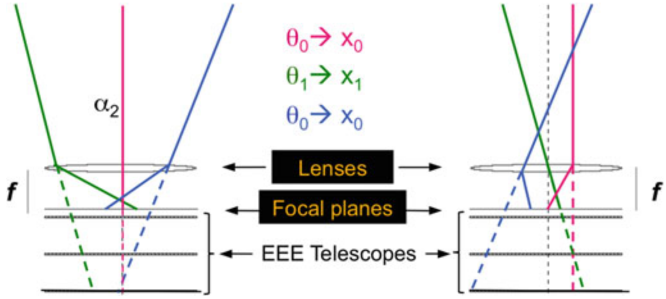


Fig. 24.8 Schematic representation of the setup employed to study the position-position correlation: The presence of a lens of focal length f is simulated on top of each telescope, in such a way that all detected muons coming at a given angle are collected in a given point of the focal plane, independent of their position of incidence. The *dashed line* is the reconstructed track of detected muons; the *continuous line* is the track due to the simulated lens

lenses is simply given by $x_{A,B} = f tan(\theta_{A,B})$, where f is the focal length of the two simulated lenses, and $\theta_{A,B}$ is the incidence angle of the muons detected at station A and B, respectively; this result holds independently of the incidence positions of the detected muons, as depicted in Fig. 24.8.

The results obtained by simulating two lenses of focal length $f = 10$ cm are shown in Fig. 24.9, where we plot the distribution of the muon pairs detected both within (blue) and outside (red) the coincidence window as a function of their relative distance $|\rho_A - \rho_B|$ in the focal planes of the two simulated lenses. The high-visibility peak characterizing muon pairs detected in coincidence indicates that the angular correlation is naturally transformed into a position-position correlation between the focal planes of two lenses, as expected. In order to quantify such position-position correlation we normalize the distribution of the coincident muon pairs with respect to the distribution of the independent pairs (e.g., we divide the blue by the red curve) and evaluate the width of the resulting distribution by performing a gaussian fit, as shown in Fig. 24.9; the visibility of the resulting correlation peak is 90 % and the spatial correlation in the focal plane of the simulated lenses is $\sigma = 0.6$ cm (corresponding to an angular correlation of 3.6°, in agreement with an analysis performed without lenses).

The existence of the position correlation between muons from extensive air shower is the building block to implement a ghost imaging scheme based on such massive particles. The result of Fig. 24.9 represents the point-spread-function of a ghost imaging scheme analogous to the classical version of optical ghost imaging (Fig. 24.2); the main difference is that we are dealing with correlated massive particles rather than with anti-correlated light beams.

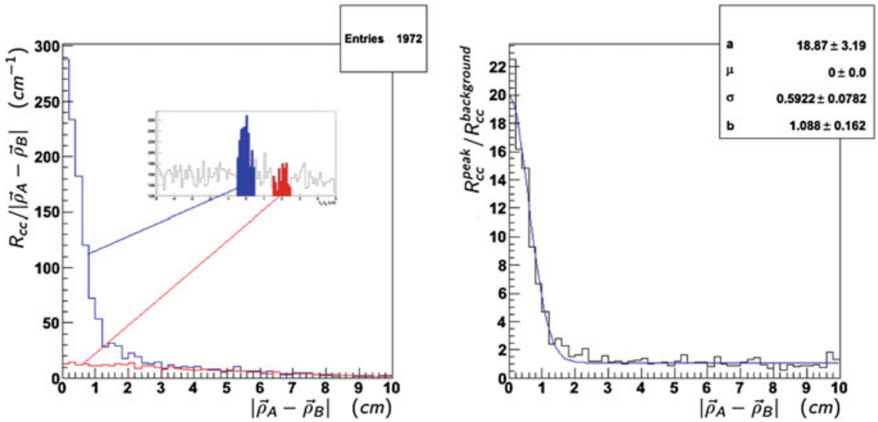


Fig. 24.9 *Left* Distribution of coincident muon pairs (*blue*) and independent muon pairs (*red*) as a function of their relative distance $|\rho_A - \rho_B|$ in the focal planes of the two simulated lenses, having focal length $f = 10$ cm. *Right* Normalized distribution of the coincidence muon pairs as a function of their relative position in the focal planes of the simulated lenses; the continuous curve is the corresponding gaussian fit

24.5 Conclusion

We have reviewed both the concept of optical ghost imaging and the preliminary results of the feasibility of ghost imaging with massive particles, thus investigating the potentialities of cosmic ray muons for long-distance ghost imaging applications. The analysis has been based on the data taken in L'Aquila within the EEE project [15].

We are currently working on the simulation of more sophisticated ghost imaging schemes, in line with the optical experiments schematically drawn in Figs. 24.1 and 24.3, by employing the same exact set of available data. This will allow both a deeper comprehension of the correlations characterizing this naturally available source of massive particles, and will indicate its potentialities in view of remote sensing applications; in fact, the results will automatically pave the way toward the extension to massive particles of many other intriguing quantum optical phenomena involving classical and non-classical correlations.

In this perspective, we are also extending the present analysis to both energy and momentum; in fact, real objects will have the potentials to be "ghostly imaged" by means of cosmic ray muons only if the strong angular correlation exploited so far is accompanied by energy correlation, thus resulting in sufficient momentum-momentum correlation between muon pairs detected in coincidence in two distant locations.

The Authors sincerely thank the EEE collaboration and the "Centro Studi e Ricerche E. Fermi", for giving us permission to analyze the data taken in L'Aquila within the EEE project; the availability of these data has been essential for conducting the present feasibility study. The Authors are particularly thankful to Marcello Abbrescia for interesting insights about the working principle of the EEE telescopes and the general setup of the EEE experiment, particularly useful in the start-up phase of the present research.

References

1. Hanbury-Brown R (1974) Intensity Interferometer. Taylor and Francis Ltd, London
2. Hanbury-Brown R, Twiss RQ (1956) Nature 177, 27 and 178, 1046.
3. Zichichi A (2005) La Scienza nelle Scuole, EEE-Extreme Energy Events, oral presentation at Congresso Nazionale, Societa' Italiana di Fisica (SIF). Bologna; also at www.centrofermi.it/eee.
4. Abbrescia M et al (2008) Nucl Instrum Methods A 593:263
5. Abbrescia M et al. (2008) Proceedings of the Cosmic Ray International Seminar (CRIS), Malfa, September 15–19; to be published in Nucl Phys B.
6. Belinskii AV, Klyshko DN (1994) JETP 105:487
7. Klyshko DN (1988) Sov Phys Usp 31:74
8. Pittman TB, Shih YH, Strekalov DV, Sergienko AV (1995) Phys Rev A 52:R3429
9. Gatti A et al (2004) Phys Rev A 70:013802
10. Valencia A, Scarcelli G, D'Angelo M, Shih Y (2005) Phys Rev Lett 94:063601
11. Meyers R et al (2008) Phys Rev A 77:041801
12. Scarcelli G (2009) Nat Phys 5:252
13. Brida G, Genovese M, Ruo Berchera I (2010) Nat Photon 4:227.
14. Akindinov A et al (2000) Nucl Instrum Methods A 456:16
15. Abbrescia M et al. (2010) Il Nuovo Cimento 125 B:243.
16. Regano A et al. (2009) oral presentation at XCV Congresso Nazionale, Societa' Italiana di Fisica (SIF). Bari.
17. D'Angelo M, Kim YH, Kulik SP, Shih YH (2004) Phys Rev Lett 92:233601
18. D'Angelo M, Valencia A, Rubin MH, Shih YH (2005) Phys Rev A 72:013810
19. Bennink RS, Bentley SJ, Boyd RW, Howell JC (2004) Phys Rev Lett 92:033601
20. Howell JC, Bennink RS, Bentley SJ, Boyd RW (2004) Phys Rev Lett 92:210403
21. Bennink RS, Bentley SJ, Boyd RW (2002) Phys Rev Lett 89:113601
22. Ferri F et al (2005) Phys Rev Lett 94:183602
23. Zhang D et al (2005) Opt Lett 30:2354
24. Scarcelli G, Berardi V, Shih Y (2006) Appl Phys Lett 88:061106
25. Scarcelli G, Valencia A, Shih Y (2004) Europhys Lett 68:618
26. D'Angelo M, Shih YH (2005) Laser Phys Lett 2:567
27. Scarcelli G, Berardi V, Shih Y (2006) Phys Rev Lett 96:063602
28. Klyshko DN (1988) Photon and Nonlinear Optics. Gordon and Breach Science, New York
29. Gatti A, Brambilla E, Lugiato LA (2003) Phys Rev Lett 90:133603
30. Migdall A (1999) Phys Today 52:41

Chapter 25
The Dark Energy Universe

Burra G. Sidharth

Abstract Some 75 years ago, the concept of dark matter was introduced by Zwicky to explain the anomaly of galactic rotation curves, though there is no clue to its identity or existence to date. In 1997, the author had introduced a model of the universe which went diametrically opposite to the existing paradigm which was a dark matter assisted decelarating universe. The new model introduces a dark energy driven accelarating universe though with a small cosmological constant. The very next year this new picture was confirmed by the Supernova observations of Perlmutter, Riess and Schmidt. These astronomers got the 2011 Nobel Prize for this dramatic observation. All this is discussed briefly, including the fact that dark energy may obviate the need for dark matter.

25.1 Introduction

By the end of the last century, the Big Bang Model had been worked out. It contained a huge amount of unobserved, hypothesized "matter" of a new kind—dark matter. This was postulated as long back as the 1930s to explain the fact that the velocity curves of the stars in the galaxies did not fall off, as they should. Instead they flattened out, suggesting that the galaxies contained some undetected and therefore non-luminous or **dark matter**. The identity of this dark matter has been a mat-

B. G. Sidharth (✉)
International Institute for Applicable Mathematics and Information Sciences,
Hyderabad, India

B. G. Sidharth
International Institute for Applicable Mathematics and Information Sciences, Udine, Italy

B. G. Sidharth
B.M. Birla Science Centre, Adarsh Nagar, Hyderabad 500 063, India

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_25,

ter of guess work, though. It could consist of Weakly Interacting Massive Particles (WIMPS) or Super Symmetric partners of existing particles. Or heavy neutrinos or monopoles or unobserved brown dwarf stars and so on.

In fact Prof. Abdus Salam speculated some two decades ago [21] "And now we come upon the question of dark matter which is one of the open problems of cosmology". This is a problem which was speculated upon by Zwicky 50 years ago. He showed that visible matter of the mass of the galaxies in the Coma cluster was inadequate to keep the galactic cluster bound. Oort claimed that the mass necessary to keep our own galaxy together was at least three times that concentrated into observable stars. And this in turn has emerged as a central problem of cosmology. "You see there is the matter which we see in our galaxy. This is what we suspect from the spiral character of the galaxy keeping it together. And there is dark matter which is not seen at all by any means whatsoever. Now the question is what does the dark matter consist of? This is what we suspect should be there to keep the galaxy bound. And so three times the mass of the matter here in our galaxy should be around in the form of the invisible matter. This is one of the speculations."

The universe in this picture, contained enough of the mysterious dark matter to halt the expansion and eventually trigger the next collapse. It must be mentioned that the latest WMAP survey [13], in a model dependent result indicates that as much as twenty three percent of the Universe is made up of dark matter, though there is no definite observational confirmation of its existence.

That is, the Universe would expand up to a point and then collapse.

There still were several subtler problems to be addressed. One was the famous **horizon problem**. To put it simply, the Big Bang was an uncontrolled or random event and so, different parts of the Universe in different directions were disconnected at the very earliest stage and even today, light would not have had enough time to connect them. So they need not be the same. Observation however shows that the Universe is by and large uniform, rather like people in different countries showing the same habits or dress. That would not be possible without some form of faster than light intercommunication which would violate Einstein's Special Theory of Relativity.

The next problem was that according to Einstein, due to the material content in the Universe, space should be curved whereas the Universe appears to be **flat**.

There were other problems as well. For example astronomers predicted that there should be **monopoles** that is, simply put, either only North magnetic poles or only South magnetic poles, unlike the North South combined magnetic poles we encounter. Such monopoles have failed to show up even after 75 years.

Some of these problems were sought to be explained by what has been called **inflationary cosmology** whereby, early on, just after the Big Bang the explosion was super fast [5, 28].

What would happen in this case is, that different parts of the Universe, which could not be accessible by light, would now get connected. At the same time, the super fast expansion in the initial stages would smoothen out any distortion or curvature effects in space, leading to a flat Universe and in the process also eliminate the monopoles.

Nevertheless, inflation theory has its problems. It does not seem to explain the cosmological constant observed since. Further, this theory seems to imply that the fluctuations it produces should continue to indefinite distances. Observation seems to imply the contrary.

One other feature that has been studied in detail over the past few decades is that of **structure formation** in the Universe. To put it simply, why is the Universe not a uniform spread of matter and radiation? On the contrary it is very lumpy with planets, stars, galaxies and so on, with a lot of space separating these objects. This has been explained in terms of fluctuations in density, that is, accidentally more matter being present in a given region. Gravitation would then draw in even more matter and so on. These fluctuations would also cause the cosmic background radiation to be non uniform or anisotropic. Such anisotropies are in fact being observed. But this is not the end of the story. The galaxies seem to be arranged along two dimensional structures and filaments with huge separating voids.

From 1997, the conventional wisdom of cosmology that had concretized from the mid sixties onwards, began to be challenged. It had been believed that the density of the Universe is near its critical value, separating eternal expansion and ultimate contraction, while the nuances of the dark matter theories were being fine tuned. But that year, the author proposed a contra view, which we will examine.

25.2 Cosmology

To proceed, as there are $N \sim 10^{80}$ such particles in the Universe, we get, consistently,

$$Nm = M \tag{25.1}$$

where M is the mass of the Universe. It must be remembered that the energy of gravitational interaction between the particles is very much insignificant compared to electromagnetic considerations.

In the following we will use N as the sole cosmological parameter.
We next invoke the well known relation [3, 8, 20]

$$R \approx \frac{GM}{c^2} \tag{25.2}$$

where M can be obtained from (25.1). We can arrive at (25.2) in different ways. For example, in a uniformly expanding Friedman Universe, we have

$$\dot{R}^2 = 8\pi G\rho R^2/3$$

In the above if we substitute $\dot{R} = c$ at R, the radius of the universe, we get (25.2). We now use the fact that given N particles, the (Gaussian) fluctuation in the particle number is of the order $\sqrt{N}$ [3, 4, 14–17], while a typical time interval for the fluctuations is $\sim \hbar/mc^2$, the Compton time, the fuzzy interval within which there is no meaningful physics as argued by Dirac and in greater detail by Wigner and Salecker. So particles are created and destroyed—but the ultimate result is that $\sqrt{N}$ particles are created just as this is the nett displacement in a random walk of unit step. So we have,

$$\frac{dN}{dt} = \frac{\sqrt{N}}{\tau} \tag{25.3}$$

whence on integration we get, (remembering that we are almost in the continuum region that is, $\tau \sim 10^{-23} sec \approx 0$),

$$T = \frac{\hbar}{mc^2}\sqrt{N} \tag{25.4}$$

We can easily verify that the Eq. (25.4) is indeed satisfied where T is the age of the Universe. Next by differentiating (25.2) with respect to t we get

$$\frac{dR}{dt} \approx HR \tag{25.5}$$

where Hin (25.5) can be identified with the Hubble Constant, and using (25.2) is given by,

$$H = \frac{Gm^3c}{\hbar^2} \tag{25.6}$$

Equations (25.1), (25.2) and (25.4) show that in this formulation, the correct mass, radius, Hubble constant and age of the Universe can be deduced given N, the number of particles, as the sole cosmological or large scale parameter. We observe that at this stage we are not invoking any particular dynamics—the expansion is due to the random creation of particles from the quantum vacuum background. Equation (25.6) can be written as

$$m \approx \left(\frac{H\hbar^2}{Gc}\right)^{\frac{1}{3}} \tag{25.7}$$

Equation (25.7) has been empirically known as an "accidental" or "mysterious" relation. As observed by Weinberg [26], this is unexplained: it relates a single cosmological parameter H to constants from microphysics. In our formulation, Eq. (25.7) is no longer a mysterious coincidence but rather a consequence of the theory.

As (25.6) and (25.5) are not exact equations but rather, order of magnitude relations, it follows, on differentiating (25.5) that a small cosmological constant $\wedge$ is allowed such that

$$\wedge \leq 0(H^2)$$

This is consistent with observation and shows that $\wedge$ is very small—this has been a puzzle, the so called cosmological constant problem because in conventional theory, it turns out to be huge [27]. But it poses no problem in this formulation. This is because of the characterization of the ZPF or quantum vacuum as independent and primary in our formulation this being the mysterious dark energy. Otherwise we would encounter the cosmological constant problem of Weinberg: a $\wedge$ that is some 10^{120} orders of magnitude of observable values!

To proceed we observe that because of the fluctuation of $\sim \sqrt{N}$ (due to the ZPF), there is an excess electrical potential energy of the electron, which in fact we identify as its inertial energy. That is [3, 14],

$$\sqrt{N}e^2/R \approx mc^2.$$

On using (25.2) in the above, we recover the well known Gravitation- Electromagnetism ratio viz.,

$$e^2/Gm^2 \sim \sqrt{N} \approx 10^{40} \tag{25.8}$$

or without using (25.2), we get, instead, the well known so called Weyl-Eddington formula,

$$R = \sqrt{N}l \tag{25.9}$$

(It appears that (25.9)) was first noticed by H. Weyl [24]. Infact (25.9) is the spatial counterpart of (25.4). If we combine (25.9) and (25.2), we get,

$$\frac{Gm}{lc^2} = \frac{1}{\sqrt{N}} \propto T^{-1} \tag{25.10}$$

where in (25.10), we have used (25.4). Following Dirac (cf.also [6]) we treat G as the variable, rather than the quantities m, l, c and $\hbar$ which we will call micro physical constants because of their central role in atomic (and sub atomic) physics.
Next if we use G from (25.10) in (25.6), we can see that

$$H = \frac{c}{l} \quad \frac{1}{\sqrt{N}} \tag{25.11}$$

Thus apart from the fact that H has the same inverse time dependence on T as G, (25.11) shows that given the microphysical constants, and N, we can deduce the Hubble Constant also, as from (25.11) or (25.6).

Using (25.1) and (25.2), we can now deduce that

$$\rho \approx \frac{m}{l^3} \quad \frac{1}{\sqrt{N}} \tag{25.12}$$

Next (25.9) and (25.4) give,

$$R = cT \tag{25.13}$$

Equations (25.12) and (25.13) are consistent with observation.
Finally, we observe that using M, G and H from the above, we get

$$M = \frac{c^3}{GH}$$

This relation is required in the Friedman model of the expanding Universe (and the Steady State model too). In fact if we use in this relation, the expression,

$$H = c/R$$

which follows from (25.11) and (25.9), then we recover (25.2). We will be repeatedly using these relations in the sequel.
As we saw the above model predicts a dark energy driven ever expanding and accelerating Universe with a small cosmological constant while the density keeps decreasing. Moreover mysterious large number relations like (25.6), (25.12) or (25.9) which were considered to be miraculous accidents now follow from the underlying theory. This seemed to go against the accepted idea that the density of the Universe equalled the critical density required for closure and that aided by dark matter, the Universe was decelerating.
However, as noted, from 1998 onwards, following the work of **Perlmutter**, **Schmidt and Riess**, these otherwise apparently heretic conclusions have been vindicated by observation.
It may be mentioned that the observational evidence for an accelerating Universe was the American Association for Advancement of Science's Breakthrough of the Year, 1998 while the evidence for nearly seventy five percent of the Universe being Dark Energy, based on the Wilkinson Microwave Anisotropy Probe (WMAP) and the Sloan Sky Digital Survey was the Breakthrough of the Year, 2003 [12, 13]. The trio got the 2011 Nobel for Physics.

25.3 Discussion

1. We observe that in the above scheme if the Compton time $\tau \to \tau_P$, we recover the Prigogine Cosmology [7, 25]. In this case there is a **phase transition** in the background ZPF or Quantum Vacuum or Dark Energy and Planck scale particles are produced.

On the other hand if $\tau \to 0$ (that is we return to point spacetime), we recover the Standard Big Bang picture. But it must be emphasized that in neither of these two special cases can we recover the various so called Large Number coincidences for example Eqs. like (25.4) or (25.6) or (25.8) or (25.9).
2. The above ideas lead to an important characterization of gravitation. This also explains why it has not been possible to unify gravitation with other interactions, despite nearly a century of effort.
Gravitation is the only interaction that could not be satisfactorily unified with the other fundamental interactions. The starting point has been a diffusion equation

$$|\Delta x|^2 = < \Delta x^2 > = \nu \cdot \Delta t$$

$$\nu = \hbar/m, \nu \approx lv \tag{25.14}$$

This way we could explain a process similar to the formation of Benard cells [7, 22]—there would be sudden formation of the "cells" from the background dark energy, each at the Planck Scale, which is the smallest physical scale. These in turn would be the underpinning for spacetime.
We could consider an array of N such Planckian cells [23]. This would be described by

$$r = \sqrt{N \Delta x^2} \tag{25.15}$$

$$ka^2 \equiv k \Delta x^2 = \frac{1}{2} k_B T \tag{25.16}$$

where k_B is the Boltzmann constant, T the temperature, r the extent and k is the spring constant given by

$$\omega_0^2 = \frac{k}{m} \tag{25.17}$$

$$\omega = \left(\frac{k}{m} a^2 \right)^{\frac{1}{2}} \frac{1}{r} = \omega_0 \frac{a}{r} \tag{25.18}$$

We now identify the particles or cells with Planck masses and set $\Delta x \equiv a = l_P$, the Planck length. It may be immediately observed that use of (25.17) and (25.16) gives $k_B T \sim m_P c^2$, which ofcourse agrees with the temperature of a black hole of Planck mass. Indeed, Rosen [10] had shown that a Planck mass particle at the Planck scale can be considered to be a Universe in itself with a Schwarzchild radius equalling the Planck length. We also use the fact alluded to that a typical elementary particle like the pion can be considered to be the result of $n \sim 10^{40}$ Planck masses.
Using this in (25.15), we get $r \sim l$, the pion Compton wavelength as required. Whence the pion mass is given by

$$m = m_P/\sqrt{n}$$

which of course is correct, with the choice of n. This can be described by

$$l = \sqrt{n} l_P, \tau = \sqrt{n} \tau_P, \tag{25.19}$$

$$l_P^2 = \frac{\hbar}{m_P} \tau_P$$

The last equation is the analogue of the diffusion process seen, which is in fact the underpinning for particles, except that this time we have the same Brownian process operating from the Planck scale to the Compton scale (cf. also [18, 19]).

We now use the well known result alluded to that the individual minimal oscillators are black holes or mini Universes as shown by Rosen [10]. So using the Beckenstein temperature formula for these primordial black holes [11], that is

$$kT = \frac{\hbar c^3}{8\pi G m}$$

we can show that

$$Gm^2 \sim \hbar c \tag{25.20}$$

We can easily verify that (25.20) leads to the value $m \sim 10^{-5} gms$. In deducing (25.20) we have used the typical expressions for the frequency as the inverse of the time—the Compton time in this case and similarly the expression for the Compton length. However it must be reiterated that no specific values for l or m were considered in the deduction of (25.20).

We now make two interesting comments. Cercignani and co-workers have shown [1, 2] that when the gravitational energy becomes of the order of the electromagnetic energy in the case of the Zero Point oscillators, that is

$$\frac{G\hbar^2\omega^3}{c^5} \sim \hbar\omega \tag{25.21}$$

then this defines a threshold frequency ω_{max} above which the oscillations become chaotic. In other words, for meaningful physics we require that

$$\omega \leq \omega_{max}.$$

Secondly as we can see from the parallel but unrelated theory of phonons [4, 9], which are also bosonic oscillators, we deduce a maximal frequency given by

$$\omega_{max}^2 = \frac{c^2}{l^2} \tag{25.22}$$

In (25.22) c is, in the particular case of phonons, the velocity of propagation, that is the velocity of sound, whereas in our case this velocity is that of light. Frequencies greater than ω_{max} in (25.22) are again meaningless. We can easily verify that using (25.21) in (25.22) gives back (25.20).

In other words, gravitation shows up as the residual energy from the formation of the particles in the universe via Planck scales (Benard like) cells.

3. It has been mentioned that despite nearly 75 years of search, Dark Matter has not been found. More recently there is evidence against the existence of Dark Matter or its previous models. The latest LHC results for example seem to rule out SUSY.

On the other hand our formulation obviates the need for Dark Matter. This follows from an equation like (25.10) which shows a gravitational constant decreasing with time. Starting from here it is possible to deduce not just the anomalous rotation curves of galaxies which was the starting point for Dark Matter; but also we could deduce all the known standard results of General Relativity like the precession of the perihelion of mercury, the bending of light, the progressive shortening of the time period of binary pulsars and so on (Cf.ref. [22]).

References

1. Cercignani C, Galgani L, Scotti A (1972) Phys Lett 38A:403
2. Cercignani C (1998) Found. Phys Lett 11(2):189–199
3. Hayakawa S (1965) Suppl of PTP commemmorative issue pp. 532–541.
4. Huang K (1975) Statistical Mechanics. Wiley Eastern, New Delhi, pp. 75ff.
5. Linde AD (1982) Phys Lett 108B:389
6. Melnikov VN (1994) Int J Theor Phys 33(7):1569–1579
7. Nicolis G, Prigogine I (1989) Exploring complexity. Freeman, New York, W.H
8. Nottale L (1993) Fractal space-time and microphysics: towards a theory of scale relativity. World Scientific, Singapore
9. Reif F (1965) Fundamentals of statistical and thermal physics. McGraw-Hill Book Co., Singapore
10. Rosen N (1993) Int J Theor Phys 32(8):1435–1440
11. Ruffini R, Zang LZ (1983) Basic concepts in relativistic astrophysics. World Scientific, Singapore, p. 111ff.
12. Science, December 1998.
13. Science, December 2003.
14. Sidharth BG (1998a) Int J Mod Phys A 13(15):2599ff.
15. Sidharth BG (1998b) Int J Theor Phys 37(4):1307–1312
16. Sidharth BG (1998c) Frontiers of quantum physics (1997) Lim SC et al. (eds.) Springer Verlag, Singapore.
17. Sidharth BG (1999) Proceeding of the eighth marcell grossmann meeting on general relativity (1997) Piran T (ed.) World Scientific, Singapore, pp. 476–479.
18. Sidharth BG (2001) Chaotic universe: from the planck to the hubble scale. Nova Science, New York
19. Sidharth BG (2002) Found. Phys Lett 15(6):577–583
20. Sidharth BG (2003) Chaos Solitons Fractals 16(4):613–620
21. Sidharth BG (2007) Encounters: abdus salam in new advances. Physics 1(1):1–17
22. Sidharth BG (2008) The thermodynamic universe. World Scientific, Singapore
23. Sidharth BG (2010) Int J Theor Phys 49(10):2476–2485

24. Singh J (1961) Great ideas and theories of modern cosmology. Dover, New York, pp. 168ff.
25. Tryon EP (1973) Is the universe a vacuum fluctuation? Nature 246:396–397
26. Weinberg S (1972) Gravitation and cosmology. Wiley, New York, p. 61ff.
27. Weinberg S (1979) Phys Rev Lett 43:1566
28. Zee A (1982) Unity of forces in the universe (Vol. II). World Scientific, Singapore, p. 40ff.

Chapter 26
New Aspects of Collective Phenomena at Nanoscales in Quantum Plasmas

P. K. Shukla and B. Eliasson

Abstract We present two novel collective effects is quantum plasmas. First, we discuss novel attractive force between ions that are shielded by the degenerate electrons in quantum plasmas. Here we show that the electric potential around an isolated ion has a hard core negative part that resembles the Lennard-Jones (LJ)-type potential. Second, we present theory for stimulated scattering instabilities of electromagnetic waves off quantum plasma modes. Our studies are based on the quantum hydrodynamical description of degenerate electrons that are greatly influenced by electromagnetic and quantum forces. The relevance of our investigation to bringing ions closer for fusion in high-energy solid density plasmas at atomic dimensions, and for producing coherent short wavelength radiation in the x-ray regime at nanoscales are discussed.

26.1 Introduction

Recently, there has been a surge in the investigation of collective processes [1–4] in quantum plasmas. The latter are ubiquitous in a variety of physical environments, including the cores of Jupiter [5, 6] and white dwarf stars [7–11], magnetars [12, 13], warm dense matters [14–16], compressed plasmas produced by intense laser beams [17–24], as well as the processing devices for modern high-technology e.g. semiconductors [25], thin films and nano-metallic structures [26]. In quantum plasmas, the electrons are degenerate and obey the Fermi-Dirac distribution function, while non-degenerate strongly correlated ions are coupled with degenerate electrons

P. K. Shukla (✉)
International Centre for Advanced Studies in Physical Sciences and Institute for Theoretical Physics, Astronomie, Ruhr-University Bochum, 44780 Bochum, Germany

B. Eliasson
Department of Mechanical and Aerospace Engineering and Center for Energy Research, University of California San Diego, La Jolla, CA 92093, USA

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_26,

via electromagnetic fields. Quantum mechanical effects become important when inter-electron distance is of the order of the thermal de Broglie wavelength. Here overlapping of electron wave functions occurs due to the Heisenberg uncertainty and Pauli's exclusion principles [27]. The quantum statistical pressure law for degenerate electrons was obtained by Chandrasekhar in his classic paper [28]. He found for the electron pressure [28–30]

$$P_e \frac{\pi m_e^4 c^5}{3h^3} \left[R(2R^2 - 3)\sqrt{1+R^2} + 3\ln(R + \sqrt{1+R^2}) \right], \quad (26.1)$$

where m_e is the rest mass of the electrons, c the speed of light in vacuum, h the Planck constant, $R = (n_e/n_c)^{1/3}$, n_e the electron number density, and $n_c \simeq 5.9 \times 10^{29}\ \mathrm{cm}^{-3}$. The electron pressure, P_e can be represented as a polytropic form $P_e = P_0(n/n_0)^\gamma$, where we have $P_0 = \pi^{2/3}\hbar^2 n_0^{5/3}/5m_e)(3\hbar c n_0^{4/3}/4)$ with $\gamma = 5/3(4/3)$ for non-relativistic (ultra-relativistic) degenerate electrons, and $\hbar = h/2\pi$. The change in the equation of state (from 5/3 to 4/3) due to ultra-relativistic degeneracy of the electrons was responsible for the discovery of the Chandrasekhar critical mass [28] ($M_c \approx (\hbar c/G)^{3/2} m_n^{-2} \approx 1.4\, M_\odot$, where $M_\odot$ is the solar mass) of a white dwarf star, due to the balance between the gradient of ultra-relativistic electron pressure $(3\hbar c/4) n_e^{4/3}$ and the gravitational force $GM/R_s^2) m_e m_n$, where G is the gravitational constant, M and R_s are the mass and radius of a star, and m_n mass of the nuclei ($n_e m_n = M/R_c^3$).

Besides the quantum statistical electron pressure in dense plasmas, there is also a negative electron pressure PC due to Coulomb interactions [31], viz.

$$P_c = -\frac{8\pi^3 m_e^4 c^5 R^4}{h^3} \frac{\alpha_0 Z_i^{2/3}}{10\pi^2} \left(\frac{4}{9\pi}\right)^{1/3}, \quad (26.2)$$

where $\alpha_0 = 1/137$ is the fine structure constant and Z_i the atomic number of ions. Furthermore, there are novel quantum forces associated with (i) the quantum electron recoil/electron tunneling effect [32–37] caused by overlapping of electron wave functions due to the Heisenberg uncertainty and Pauli exclusion principles, and (ii) electron-exchange and electron correlations [38, 39] caused by electron spin effect [40]. These quantum forces (the quantum pressure gradients, the quantum Bohm force due to the quantum electron recoil effect [32, 33], and the gradient of the the potential associated with electron-exchange and electron-correlations effects [38, 39]), and electromagnetic forces control the dynamics of degenerate electrons in dense quantum plasmas. It turns out that the quantum forces produce new features to both electron and ion plasma oscillations [1–4] in an unmagnetized quantum plasma.

In this paper, we present two novel aspects of collective interactions in quantum plasmas. First, we report on the discovery of novel attractive force [41, 42] at atomic dimensions/nanoscales that can bring ions closer. Novel attractive force arises due to the interplay between the quantum Bohmian force [32, 33], as well as forces due to the quantum statistical pressure [28] of non-relativistic degenerate electrons and electron-exchange and electron correlations effects [38, 39]. Specifically, we shall use here the generalized quantum hydrodynamical (G-QHD) equations [1–4, 41, 42] for non-relativistic degenerate electron fluids, the generalized viscoelastic (GV) ion momentum equations [1, 2] including ion correlations and ion fluid viscosity effects, supplemented by Poisson's equation. Thus, our G-QHD-GV model captures the essential physics of electrostatic collective modes in quantum plasmas.

The manuscript is organized in the following fashion. In Sect. 26.2, we present the governing equations for electrostatic perturbations and derive the electron and ion susceptibilities in an unmagnetized quantum plasma. The linear response theory for quasi-stationary electrostatic perturbations involving immobile ions reveals that the electric potential around an isolated ion in a quantum plasma has a new profile [41, 42], which resembles the Lennard-Jones-type potential. Thus, our newly found electric potential embodies a short-range negative hard core electric potential. The latter would be responsible for ion clustering and oscillations of ion lattices under the new attractive force we have recently discovered [41, 42] in a strongly coupled quantum plasma. Section 26.3 contains an investigation of stimulated scattering instabilities of coherent high-frequency electromagnetic (EM) waves off quantum electron and ion plasma modes. Here we discuss stimulated Raman, stimulated Brillouin, and modulational instabilities of a constant amplitude EM pump wave. Explicit expressions for the growth rates are presented. The relevance of our investigation to the field of high-energy density quantum plasma physics is pointed out.

26.2 The Governing Equations

Let us consider an unmagnetized quantum plasma in the presence of non-relativistic degenerate electron fluids and mildly coupled ion fluids. In our quantum plasma, the electron and ion coupling parameters are $\Gamma_e = e^2/a_e k_B T_F$ and $\Gamma_i = Z_i^2 e^2/a_i k_B T_i$, respectively, where $a_e \sim a_i = (3/4\pi n_0)^{1/3}$ is the average interparticle distance, k_B the Boltzmann constant, $T_F = (\hbar^2/2m_e k_B)(3\pi^2 n_0)^{2/3}$ the Fermi electron temperature, and T_i the ion temperature. It turns out that $\Gamma_i/\Gamma_e = Z_i^2 T_F/T_i \gg 1$, since in quantum plasmas we usually have $T_F > T_i$. The dynamics of degenerate electron fluids is governed by the generalized quantum hydrodynamic (G-QHD) equations composed of the continuity equation

$$\frac{\partial n_e}{\partial t} + \nabla \cdot (n_e \mathbf{u}_e) = 0, \tag{26.3}$$

the momentum equation

$$m_e n_e \left(\frac{d\mathbf{u}_e}{dt} \right) = n_e e \nabla \phi - \nabla P_* + n_e \nabla V_{xc} + n_e \nabla V_B, \tag{26.4}$$

and Poisson's equation

$$\nabla^2\phi = 4\pi e(n_e - n_i) - 4\pi Q\delta(\mathbf{r}), \tag{26.5}$$

Here $d/dt = (\partial/\partial t) + \mathbf{u}_e \cdot \nabla$, $n_e(n_i)$ is the electron (ion) number density, and $\mathbf{u}_e$ the electron fluid velocity, and ϕ the electrostatic potential. Furthermore, we have denoted the quantum statistical pressure $P_* = (n_0 m_e v_*^2/5)(n/n_0)^{5/3}$, where $v_* = \hbar(3\pi^2)^{1/3}/m_e r_0$ is the electron Fermi speed and $r_0 = n_0^{-1/3}$ represents the Wigner-Seitz radius. The sum of the electron exchange and electron correlations potential is [38, 39] $V_{xc} = 0.985(e^2/\varepsilon)n^{1/3}\left[1 + (0.034/a_B n^{1/3})\ln(1 + 18.37 a_B n^{1/3})\right]$, where $a_B = \hbar^2/m_e e^2$ is the Bohr atomic radius. The quantum Bohm potential reads [1, 2, 4, 32, 33] $V_B = (\hbar^2/2m_e)(1/\sqrt{n_e})\nabla^2\sqrt{n_e}$. Equation (26.4) is valid if the plasmonic energy density $\hbar\omega_{pe}$ is smaller (or comparable) than (with) the Fermi electron kinetic energy $k_B T_F$, where $\omega_{pe} = (4\pi n_0 e^2/m_e)^{1/2}$ is the electron plasma frequency, and the electron-ion collision relaxation time is greater than the electron plasma period. The test charge is denoted by Q.

The ion number density perturbation n_{i1} is obtained from the ion continuity equation

$$\frac{\partial n_i}{\partial t} + \nabla \cdot (n_i \mathbf{u}_i) = 0, \tag{26.6}$$

where the ion fluid velocity $\mathbf{u}_i$ is determined from the generalized viscoelastic ion momentum equation

$$\left(1 + \tau_m \frac{D}{Dt}\right)\left[m_i n_i \frac{D\mathbf{u}_i}{Dt} + \nabla P_i + Z_i e n_i \nabla\phi\right] - \eta\nabla \cdot \nabla\mathbf{u}_i - \left(\xi + \frac{\eta}{3}\right)\nabla(\nabla \cdot \mathbf{u}_i) = 0, \tag{26.7}$$

where $D/Dt = (\partial/\partial t) + \mathbf{u}_i \cdot \nabla$, τ_m is the viscoelastic relaxation time for ion correlations, $P_i = \gamma_i \mu_i k_B T_i \nabla n_i$ the ion thermal pressure involving strong ion coupling effects, m_i the ion mass, γ_i the adiabatic index, $\mu_i \left[= (1/k_B T_i)(\partial P_i/\partial n_i)_{k_B T_i}\right] \equiv 1 + U_i(\Gamma_i)/3 + (\Gamma_i/9)\partial U_i(\Gamma_i)/\partial\Gamma_i$ the isothermal compressibility factor for non-degenerate ion fluids, the function $U_i(\Gamma_i)$ is the measure of the excess internal energy of the system, which is related with the correlation energy E_c by $U_i(\Gamma_i) = E_c/n_i k_B T_i$ [$\equiv \Gamma_i(0.9 + 1.5 r_i^2/a_i^2)$, where r_i is the ion core radius which depends on the degree of ion stripping [43]]. For one-component plasma model, we [44–48] usually adopt $U_i(\Gamma_i) \approx -0.9\Gamma_i$ for $\Gamma_i \gg 1$ Furthermore, the coefficients of the shear and bulk ion fluid viscosities are denoted by η and ξ, respectively. We note that Eq. (26.7) is similar to that used by Frenkel [49] and Ichimaru [44–48] in the context of ordinary fluids and one-component strongly coupled plasmas, respectively. Furthermore, Kaw and Sen [50] adopted a generalized viscoelastic dust momentum equation for studying the properties of dust acoustic waves [51] in a multi-component dusty plasma with highly- charged dust grains.

Letting $n_j = n_0 + n_{j1}$, where $n_{j1} \ll n_0$ and j equals e for electrons and i for ions, we linearize Eqs. (26.3) and (26.4) as well as (26.6) and (26.7) to obtain the electron and ion number density perturbations n_{j1}. The latter can be substituted into the Fourier transformed version of Eq. (26.5), obtaining, in the linear approximation, the electric potential [41, 42]

$$\phi(\mathbf{r}) = \frac{Q}{2\pi^2} \int \frac{\exp(i\mathbf{k}\cdot\mathbf{r})}{k^2 D(\omega, \mathbf{k})} d^3k, \tag{26.8}$$

where the dielectric constant $D(\omega, \mathbf{k})$ for an unmagnetized quantum plasma is

$$D(\omega, \mathbf{k}) = 1 + \chi_e(\omega, \mathbf{k}) + \chi_i(\omega, \mathbf{k}). \tag{26.9}$$

Here the electron and ion susceptibilities are, respectively,

$$\chi_e(\omega, \mathbf{k}) = -\frac{\omega_{pe}^2}{\omega^2 - k^2U_*^2 - \hbar^2k^4/4m_e^2}, \tag{26.10}$$

and

$$\chi_i(\omega, \mathbf{k}) = -\frac{\omega_{pi}^2}{\omega^2 - k^2V_{ti}^2 + i\omega\eta_* k^2/(1 - i\omega\tau_m)}, \tag{26.11}$$

where $U_* = (v_*^2/3 + v_{ex}^2)^{1/2}$, $v_{ex} = (0.328e^2/m_e r_0)^{1/2}[1 + 0.62/(1 + 18.36\, a_B n_0^{1/3})]^{1/2}$, $\omega_{pi} = (4\pi n_0 Z_i^2 e^2/m_i)^{1/2}$, $V_{ti} = (\gamma_i \mu_i k_B T_i/m_i)^{1/2}$ the ion thermal speed, and $\eta_* = (\xi + 4\eta/3)/m_i n_0$. Furthermore, the frequency and the wave vector are denoted by ω and $\mathbf{k}$, respectively.

26.2.1 Linear Quantum Plasma Waves

Two cases are of interest. First, for the high-frequency electron plasma waves with $\omega \gg \omega_{pi}$, we have from $1 + \chi_e(\omega, \mathbf{k}) = 0$, the frequency of the electron plasma oscillations (EPOs)

$$\omega(k) = \left(\omega_{pe}^2 + k^2U_*^2 + \frac{\hbar^2k^4}{4m_e^2}\right)^{1/2} \equiv \Omega_L(k). \tag{26.12}$$

Second for $\omega \ll k(U_*^2 + \hbar^2k^2/4m_e^2)^{1/2} \equiv kU_{q*}$, we have from $\chi_e(\mathbf{k}) = \omega_{pe}^2/k^2U_{q*}^2$, so that $1 + \chi_e + \chi_i = 0$ gives

$$\omega^2 + i\omega\frac{\omega_v}{(1 - i\omega\tau_m)} - k^2C_s^2 - \frac{\hbar^2k^4}{4m_e m_i} = 0, \tag{26.13}$$

for the ion plasma oscillations (IPOs). Here we have denoted $\omega_v = \eta_* k^2$ and $C_s = (m_e U_*^2/m_i + V_{ti}^2)^{1/2}$. In the hydrodynamic limit, viz. $\omega\tau_m \ll 1$, we have viscous damping of the quantum ion plasma mode. The real and imaginary parts of the frequencies ($\omega = \omega_r + i\omega_i$) are, respectively,

$$\omega_r(k) = \left[k^2 C_s^2 + \frac{\hbar^2 k^4}{4m_e m_i} - \omega_i^2\right]^{1/2} \equiv \Omega_s(k), \tag{26.14}$$

and

$$\omega_i(k) = -\frac{\omega_v}{2}. \tag{26.15}$$

Furthermore, in the kinetic regime characterized by $\omega\tau_m \gg 1$, we have from (26.13)

$$\omega(k) = \left(\frac{\omega_v}{\tau_m} + k^2 C_s^2 + \frac{\hbar^2 k^4}{4m_e m_i}\right)^{1/2} \equiv \Omega_I(k). \tag{26.16}$$

Generally, $\tau_m = \tau_0 Y_g(k)$, where $\tau_0 = 1/\omega_v(1 - \mu_i + 4U_i(\Gamma_i)/15)$ and $Y_g(k) = \exp(-k/k_g)$ for a Gaussian distribution, and $Y_l(k) = (1+k^2/k_L^2)^{-1}$ for a Lorentzian distribution. Here k_g and k_l are the scale factors [45–48].

26.2.2 Potential Distribution Around an Isolated Ion

Let us now present novel attractive force [41] in quantum plasmas, by using $1/D$ that is associated with quasi-stationary ($\omega \to 0$)) propagating electrostatic perturbations in a quantum plasma with immobile ions. Accordingly, we have [41]

$$\frac{1}{D} = \frac{(k^2/k_s^2) + \alpha k^4/k_s^4}{1 + (k^2/k_s^2) + \alpha k^4/k_s^4}, \tag{26.17}$$

where $k_s = \omega_{pe}/U_*$ is the inverse modified (due to the potential of the electron-exchange and electron-correlations) Thomas-Fermi screening length, and $\alpha = \hbar^2\omega_{pe}^2/4m_e^2 U_*^2$. We note that α depends only on r_0/a_B.

Inserting Eq. (26.17) into Eq. (26.8), we can express the latter as [41]

$$\phi(\mathbf{r}) = \frac{Q}{4\pi^2}\int\left[\frac{(1+b)}{k^2 + k_+^2} + \frac{(1-b)}{k^2 + k_-^2}\right]\exp(i\mathbf{k}\cdot\mathbf{r})d^3k, \tag{26.18}$$

where $b = 1/\sqrt{1-4\alpha}$, and $k_\pm^2 = k_s^2[1 \mp \sqrt{1-4\alpha}]/2\alpha$.

The integral in Eq. (26.18) can be evaluated using the general formula

$$\int \frac{\exp(i\mathbf{k}\cdot\mathbf{r})}{k^2 + k_\pm^2}d^3k = 2\pi^2\frac{\exp(-k_\pm r)}{r}, \tag{26.19}$$

where the branches of $k_\pm$ must be chosen with positive real parts so that the boundary condition $\phi \to 0$ at $r \to \infty$ is fulfilled. From Eqs. (26.18) and (26.19) we then have

$$\phi(\mathbf{r}) = \frac{Q}{r}\,[\cos(k_i r) - ib\sin(k_i r)]\exp(-k_r r). \tag{26.20}$$

Several comments are in order. First, for $\alpha > 1/4$, the solution of $k_\pm^2 = k_s^2(1 \mp i\sqrt{4\alpha - 1})/2\alpha$ yields $k_\pm = (k_s/\sqrt{4\alpha})[(\sqrt{4\alpha}+1)^{1/2} \mp i(\sqrt{4\alpha}-1)^{1/2}] \equiv k_r \mp ik_i$, and the corresponding electric potential reads [41]

$$\phi(\mathbf{r}) = \frac{Q}{r}\big[\cos(k_i r) + b_*\sin(k_i r)\big]\exp(-k_r r), \tag{26.21}$$

where $b_* = 1/\sqrt{4\alpha - 1}$. Second, for $\alpha \to 1/4$, we have $k_+ = k_- = \sqrt{2}k_s$ and

$$\phi(\mathbf{r}) = \frac{Q}{r}\left(1 + \frac{k_s r}{\sqrt{2}}\right)\exp(-\sqrt{2}\,k_s r). \tag{26.22}$$

Third, for $\alpha < 1/4$, the expression $\sqrt{1-4\alpha}$ is real, and we obtain $k_\pm = k_s(1 \mp \sqrt{1-4\alpha})^{1/2}/\sqrt{2\alpha}$. The resultant electric potential is

$$\phi(\mathbf{r}) = \frac{Q}{2r}\big[(1+b)\exp(-k_+ r) + (1-b)\exp(-k_- r)\big]. \tag{26.23}$$

We note that in the absence of the quantum recoil effect, viz. $\alpha \to 0$, we recover from Eq. (26.23) the modified Thomas-Fermi screened Coulomb potential $\phi(\mathbf{r}) = (Q/r)\exp(-k_s r)$. Finally, in the limit $\hbar^2 k^2 \gg m_e^2 U_*^2$, the quantum recoil force dominates over the quantum statistical pressure and the forces associated with electron-exchange and electron correlations effects. Here we have the exponential cosine-screened Coulomb potential [52]

$$\phi(\mathbf{r}) = \frac{Q}{r}\cos(k_q r)\exp(-k_q r), \tag{26.24}$$

where $k_q = (2m_e\omega_{pe}/\hbar)^{1/2}$ is the quantum wave number.

Figure 26.1 displays the variation of α against a_B/r_0, with a maximum α at $a_B/r_0 \approx 0.15$, corresponding to a number density $n_0 \approx 2 \times 10^{22}\,\text{cm}^{-3}$ (with $a_B = 5.3 \times 10^{-9}$ cm), a few times below solid densities. The value of α is above the critical value 1/4 only for a limited range of the plasma densities. In Fig. 26.2, we depict the profiles of the potential [given by Eqs. (26.21), (26.22) and (26.23) for $\alpha > 1/4$, $\alpha = 1/4$ and $\alpha < 1/4$, respectively] for different values of α. We clearly see the new short-range attractive electric potential that resembles the LJ-type potential for $0.25 < \alpha < 0.627$, while for smaller values of α, the attractive potential is absent. Figure 26.3a exhibits the distance $r = d$ from the test ion charge where $d\phi/dr = 0$ and the electric potential has its minimum, and Figs. 26.3b and 26.3c show the values of ϕ and $d^2\phi/dr^2$ at $r = d$. The value of $(d^2\phi/dr^2)$ determines the frequency and dispersion properties of the ion plasma oscillations, as shown below.

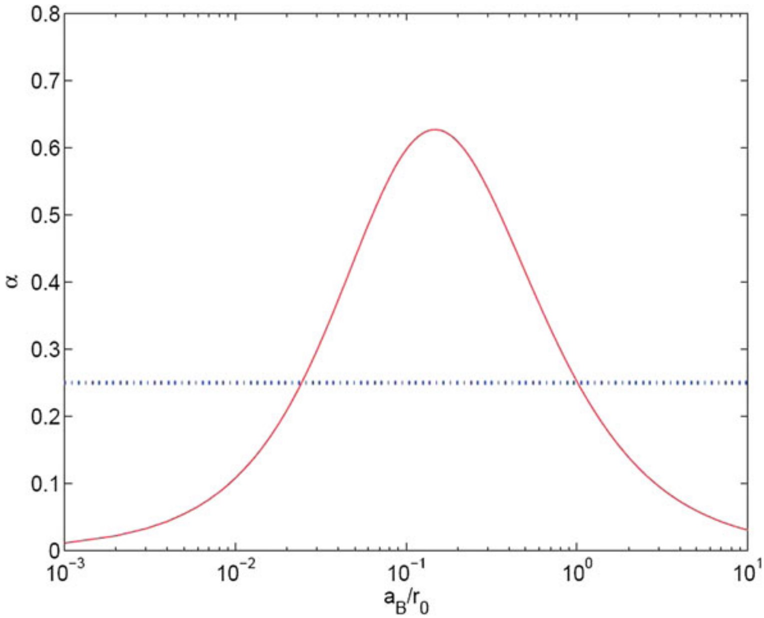


Fig. 26.1 The variation of α against a_B/r_0. The critical value $\alpha = 1/4$ is indicated with a *dotted line*

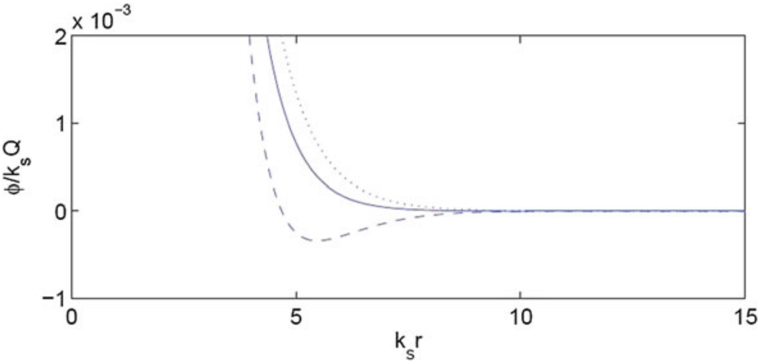


Fig. 26.2 The electric potential ϕ as a function of r for $\alpha = 0.625$ (*dashed curve*), $\alpha = 1/4$ (*solid curve*) and $\alpha = 0$ (*dotted curve*). The value 0.627 is the maximum possible value of α in our model, obtained for $a_B/r_0 \approx 0.15$. After P. K. Shukla and B. Eliasson, *Phys. Rev. Lett.* **108**, 219902 (2012)

For $\alpha \leq 0.25$, the electric potential and its second derivative vanish, and there is no attractive potential associated with the stationary test ion charge.

The interaction potential energy between two dressed ions with charges Q_i and Q_j at the positions $\mathbf{r}_i$ and $\mathbf{r}_j$ can be represented as $U_{i,j}(\mathbf{R}_{ij}) = Q_j\phi_i(\mathbf{R}_{ij})$, where ϕ_i is the potential around particle i, and $\mathbf{R}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. For $\alpha > 1/4$, it reads, using

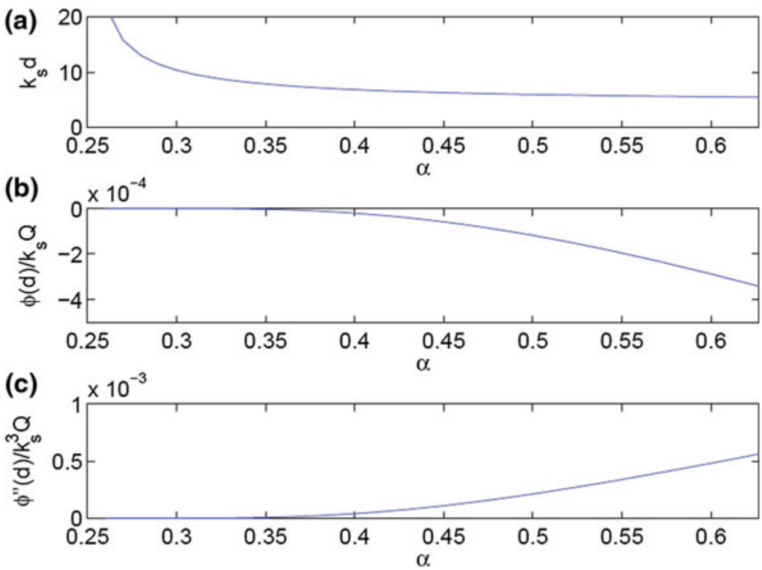


Fig. 26.3 **a** The distance $r = d$ from the test ion charge where $d\phi/dr = 0$ and the electric potential has its minimum, and **b** the values of the potential ϕ and c) its second derivative $d^2\phi/dr^2$ at $r = d$. After P. K. Shukla and B. Eliasson, *Phys. Rev. Lett.* **108**, 219902 (2012)

Eq.(26.21),

$$U_{i,j}(\mathbf{R}_{ij}) = \frac{Q_i Q_j}{|\mathbf{R}_{ij}|} \exp(-k_r|\mathbf{R}_{ij}|) \times \left[\cos(k_i|\mathbf{R}_{ij}|) + b_* \sin(k_i|\mathbf{R}_{ij}|)\right]. \quad (26.25)$$

On account of the interaction potential energy, ions would suffer vertical oscillations around their equilibrium position. The vertical vibrations of ions in a crystallized ion string in quantum plasmas are governed by

$$m_i \frac{d^2\delta z_j(t)}{dt^2} = -\sum_{i\neq j} \frac{\partial U_{ij}(\mathbf{r}_i, \mathbf{r}_j)}{\partial z_j}, \quad (26.26)$$

where $\delta z_j(t)[= z_j(t) - z_{j0}]$ is the vertical displacement of the jth ion from its equilibrium position z_{j0}. Assuming that $\delta z_j(t)$ is proportional to $\exp[-i(\omega t - jka)]$, where ω and k are the frequency and wave number of the ion lattice oscillations, respectively, and that $Q_i = Q_j = Q$, Eq. (26.26) for the nearest-neighbor ion interactions gives

$$\omega^2 = \frac{4Q^2}{m_i d^3} S \exp(-k_r d) \sin^2\left(\frac{kd}{2}\right), \quad (26.27)$$

where $S = [2(1 + k_r d) + (k_r^2 - k_i^2)d^2](\cos\xi + b_* \sin\xi) + 2k_i d(1 + k_r d)(\sin\xi - b_* \cos\xi)$. $\xi = k_i d$, and d is the separation between two consecutive ions. We note that Eq. (26.27) can also be obtained from the formula [53]

$$\omega^2 = \frac{4}{m_i}\left[\frac{d^2 W(r)}{dr^2}\right]_{r=d} \sin^2\left(\frac{kd}{2}\right), \tag{26.28}$$

where the inter-ion potential energy is represented as $W(r) = (Q^2/r)\exp(-k_r r)$ $[\cos(k_i r) + b_* \sin(k_i r)]$ for $\alpha > 1/4$. Hence, the lattice wave frequency is proportional to $[d^2\phi(r)/dr^2]^{1/2}_{r=d}$ [cf. Fig. 26.3c], and decreases rapidly for $\alpha < 0.25$.

26.3 Stimulated Scattering of Electromagnetic Waves

Large amplitude high-frequency (HF) electromagnetic (EM) waves are used for heating inertial confined fusion plasmas [54, 55], as well as for diagnostic purposes [14–16] in solid density plasmas that are created by intense laser and charged particle beams. The HF-EM pulses also appear as localized bursts of x-rays and γ-rays from compact astrophysical objects. Furthermore, the generation of coherent HF-EM waves is of great importance in the context of free-electron lasers (FELs) involving EM wigglers [56–59]. Therefore, studies of nonlinear phenomena (e.g parametric instabilities [60] and HF-EM wave localizations [61]) associated with a large amplitude HF-EM waves in dense quantum plasmas are of practical interest.

In the following, we present an investigation of stimulated scattering instabilities of coherent circular polarized electromagnetic (CPEM) waves carrying orbital angular momentum (OAM) in an unmagnetized dense quantum plasma. The nonlinear interactions between the CPEM waves and electrostatic plasma oscillations in quantum plasmas is governed by the EM wave equation [60, 62]

$$\left(\frac{\partial^2}{\partial t^2} - c^2\nabla^2 + \omega_p^2\right)\mathbf{A} + \omega_{pe}^2 N\mathbf{A} = 0, \tag{26.29}$$

which is derived from the Maxwell equation and the electron equation of motion with the electromagnetic fields $\mathbf{E} = -c^{-1}\partial\mathbf{A}/\partial t$ and $\mathbf{B} = \nabla \times \mathbf{A}$, using the Coulomb gauge $\nabla \cdot \mathbf{A} = 0$. Here $\mathbf{A}$ is the vector potential, $N = n_{e1}/n_0 \ll 1$, and n_{e1} the electron number density perturbation associated with low-frequency electrostatic plasma oscillations (EPOs) that are reinforced by the ponderomotive force of the CPEM waves. In the absence of nonlinear couplings between the latter and the EPOs, the paraxial EM wave solution, $\mathbf{A}(r, z)\exp(-i\omega_e t + ik_e z)$, where $\omega_e = (k_e^2 c^2 + \omega_{pe}^2)^{1/2}$ and k_e are the frequency and the wave number of the CPEM wave, respectively, of Eq. (26.28) is [63, 64]

$$\mathbf{A}(r, z) = \mathbf{A}F_{p,l}(r, z)\exp(il\varphi). \tag{26.30}$$

Here we have denoted

$$F_{p,l}(r,z) = \frac{1}{2\sqrt{\pi}}\left[\frac{(l+p)!}{p!}\right]^{1/2} X^{|l|}L_p^{|l|}(X)\exp(-X/2), \tag{26.31}$$

where $X = r^2/w^2(z)$, $w(z)$ the beam waist, and the associated Laguerre polynomial $L_p^{|l|}(X)$ are defined by the Rodriguez formula, $L_p^{|l|}(X) = (X^l p!)^{-1}\exp(X)d^p[X^{(l+p)}\exp(-X)]/dX^p$, p and l are the radial and angular mode numbers of the EM orbital angular momentum states, respectively, φ the azimuthal angle, $r = (x^2+y^2)^{1/2}$ the radial of the cylindrical coordinates $((r,\varphi,z)$, so that $\nabla^2 = \nabla_\perp^2 + \partial_z^2$, where $\nabla_\perp^2 = (1/r)(\partial/\partial r)(r\partial/\partial r) + (1/r^2)\partial^2/\partial\varphi^2$. The Laguerre-Gauss (LG) solutions (26.30) describe CPEM waves with a finite OAM.

The dynamics of the low-frequency (in comparison with the CPEM wave frequency ω_e) plasma oscillations involving degenerate electrons and non-degenerate ion fluids is governed by the G-QHD equations composed of the linearized electron continuity equation

$$\frac{\partial n_{e1}}{\partial t} + n_0\nabla\cdot\mathbf{u}_e = 0, \tag{26.32}$$

the electron momentum equation

$$\frac{\partial \mathbf{u}_e}{\partial t} - \frac{e}{m_e}\nabla(\phi-\phi_p) + U_*^2\nabla n_{e1} + \frac{\hbar^2}{4m_e^2}\nabla\nabla^2 n_{e1} = 0, \tag{26.33}$$

and Poisson's equation

$$\nabla^2\phi = 4\pi e(n_{e1} - ni1), \tag{26.34}$$

together with the linearized Eqs. (26.6) and (26.7). Here $\phi_p = e|\mathbf{A}|^2/m_e c^2$ is the ponderomotive potential of the CPEM waves. The EM ponderomotive force [54, 55] $-e\nabla\phi_p$ comes from the averaging (over the CPEM wave period) of the advection and nonlinear Lorentz forces involving the electron quiver velocity and the laser wave magnetic field.

Let us now derive the governing equations for the electron and ion plasma oscillations in the presence of the ponderomotive force of the CPEM waves. First, we consider the driven electron plasma oscillations on the time scale of the electron plasma period, so that the ions do not have time to respond and the ion density perturbation is zero. We combine Eqs. (26.32) and (26.33) and substitute the resultant equation into (26.34) with $n_{i1} = 0$ to obtain the electron plasma wave equation [62]

$$\left(\frac{\partial^2}{\partial t^2} + \omega_{pe}^2 - U_*^2\nabla^2 + \frac{\hbar^2}{4m_e^2}\nabla^4\right)N = \frac{e^2}{2m_e^2c^2}\nabla^2|\mathbf{A}|^2. \tag{26.35}$$

Second, we consider driven ion oscillations by supposing that $|\partial^2 N/\partial t^2| \ll |U_*^2\nabla^2 N + (\hbar^2/4m_e^2)\nabla^4 N|$. Here, one can neglect the right-hand side in Eq. (26.33), and use the resultant equation to eliminate the electric field $-\nabla\phi$ from Eq. (26.7) to obtain, after linearization of the resultant equation under the quasi-neutral approximation $n_{e1} = n_{i1}$, the driven ion oscillation equation [62]

$$\left(1 + \tau_m \frac{\partial}{\partial t}\right)\left(\frac{\partial^2}{\partial t^2} - C_s^2\nabla^2 + \frac{\hbar^2}{4m_e m_i}\nabla^4\right) N$$
$$- \frac{\eta}{m_i n_0}\nabla\cdot\nabla\frac{\partial N}{\partial t} - \frac{\left(\xi + \frac{\eta}{3}\right)}{m_i n_0}\nabla^2\frac{\partial N}{\partial t} = \left(1 + \tau_m \frac{\partial}{\partial t}\right)\frac{Z_i^2 e^2}{2m_0 m_i c^2}\nabla^2|\mathbf{A}|^2, \quad (26.36)$$

where we have used the linearized Eq. (26.6) to eliminate $\nabla\cdot\mathbf{u}_i$ from the ion momentum equation (26.7).

Equations (26.29), (26.35) and (26.36) are the desired equations [62] for studying nonlinear effects (viz. parametric instabilities [54] and localization of light pulses [55] associated with L-G CPEM beams in quantum plasmas at nanoscales.

26.3.1 Nonlinear Dispersion Relations

In the following, we present an investigation of stimulated Raman, stimulated Brillouin, and modulational instabilities [65–67] of L-G CPEM waves. Accordingly, we decompose the vector potential as

$$\mathbf{A} = \mathbf{A}_{0+}\exp(-i\omega_0 t + i\mathbf{k}_0\cdot\mathbf{r}) + \mathbf{A}_{0-}\exp(i\omega_0 t - i\mathbf{k}_0\cdot\mathbf{r})$$
$$+ \sum_{+,-}\mathbf{A}_{\pm}\exp(-i\omega_{\pm} t + i\mathbf{k}_{\pm}\cdot\mathbf{r}), \quad (26.37)$$

where the subscripts 0 and $\pm$ denote the CPEM pump and CPEM sidebands, respectively, and $\omega_{\pm} = \Omega \pm \omega_0$ and $\mathbf{k}_{\pm} = \mathbf{K} \pm \mathbf{k}_0$ are the frequency and wave vectors of the CPEM sidebands that are created by the beating of the pump $(\omega_0, \mathbf{k}_0)$ and electrostatic oscillations $(\Omega, \mathbf{K})$.

Inserting (26.37) into (26.29), (26.35), and (26.36), and supposing that N is proportional to $\exp(-i\Omega t + i\mathbf{K}\cdot\mathbf{r})$, we Fourier decompose the resultant equations to obtain the nonlinear dispersion relations [62]

$$S_R = \frac{\omega_{pe}^2 e^2 K^2}{2m_e^2 c^2}\sum_{+,-}\frac{|\mathbf{A}_0 F_{p,l}|^2}{D_{\pm}}, \quad (26.38)$$

and

$$S_B = \frac{\omega_{pe}^2 Z_i^2 e^2 K^2}{2m_e m_i c^2}\sum_{+,-}\frac{|\mathbf{A}_0 F_{p,l}|^2}{D_{\pm}}. \quad (26.39)$$

Here we have denoted $S_R = \Omega^2 - \omega_{pe}^2 - K^2U_*^2 - \hbar^2K^4/4m_e^2$ and $S_B = \Omega^2 - K^2C_s^2 - (\hbar^2K^4/4m_em_i) + i\Omega\eta_*K^2/m_in_0(1-i\Omega\tau_m)$, and $D_\pm = \omega_\pm^2 - k_\pm^2c^2 - \omega_{pe}^2 \approx \pm 2\omega_0(\Omega - \mathbf{K}\cdot\mathbf{V}_g \mp \delta)$, where $\mathbf{V}_g = c^2\mathbf{k}_0/\omega_0$ is the group velocity of the CPEM pump, $\omega_0 = (\omega_{pe}^2 + k_0^2c^2)^{1/2}$ the pump wave frequency, and $\delta = K^2c^2/2\omega_0$ the small frequency shift arising from the nonlinear interaction between the CPEM pump and the electrostatic plasma oscillations in our quantum plasma.

26.3.2 Growth Rates

We now present a summary of formulas for the growth rates of stimulated raman (SR) and stimulated brillouin (SB) scattering instabilities, as well as of modulational instabilities of a constant amplitude pump that is scattered off a quantum electron plasma wave, a quantum ion mode, and a spectrum of non-resonant electron and ion density perturbations. For three-wave decay interactions, one assumes that $D_- = 0$ and $D_+ \neq 0$. Thus, one ignores D_+ from Eqs. (26.38) and (26.39). Letting $\Omega = \mathbf{K}\cdot\mathbf{V}_g - \delta + i\gamma_{R,B}$, and $\Omega = \Omega_L + i\gamma_R$ and $\Omega = \Omega_s\ (\Omega_I) + i\gamma_B$ in the resultant equations, we obtain the growth rates for Raman and Brillouin backscattering ($|\mathbf{K}| = 2k_0$) instabilities, respectively,

$$\gamma_R = \frac{\omega_{pe}k_0e|\mathbf{A}_0F_{p,l}|}{\sqrt{2\omega_e\Omega_R}m_ec}, \tag{26.40}$$

and

$$\gamma_B = \frac{\omega_{pe}k_0Z_ie|\mathbf{A}_0F_{p,l}|}{\sqrt{2\omega_0\Omega_Bm_em_i}c}, \tag{26.41}$$

where $\Omega_R = \Omega_L(K = 2k_0)$, $\Omega_B = \Omega_{s,I}(K = 2k_0)$, and $|2k_0\mathbf{V_g} - \delta| \sim \Omega_R, \Omega_B$. Since the growth rates of SR and SB scattering instabilities, given by Eqs. (26.40) and (26.41), respectively, are proportional to Ω_R and Ω_B, one notices that quantum and strong ion correlation effects significantly affect the e-folding time of the instabilities. Furthermore, the growth rates, which are proportional to $F_{p,l}$, are minimum at the center of the vortex pump wave with OAM.

Next, for the modulational instabilities, we have $D_\pm \neq 0$ and $S_{R,B} \neq 0$. Here, we have to retain both upper and lower CPEM sidebands on an equal footing in (26.38) and (26.39), and write them as

$$S_R\left[(\Omega - \mathbf{K}\cdot\mathbf{V}_g)^2 - \delta^2\right] = \frac{\delta\omega_{pe}^2e^2K^2|\mathbf{A}_0F_{p,l}|^2}{2\omega_0m_e^2c^2}, \tag{26.42}$$

and

$$S_B\left[(\Omega - \mathbf{K}\cdot\mathbf{V}_g)^2 - \delta^2\right] = \frac{\delta\omega_{pe}^2Z_i^2e^2K^2|\mathbf{A}_0F_{p,l}|^2}{2\omega_0m_em_ic^2}. \tag{26.43}$$

Equations (26.42) and (26.43) can be analyzed numerically to obtain the growth rates of the modulational instabilities. However, some analytical results follow for $\mathbf{K} \cdot \mathbf{V}_g = 0$, in which case we have from (26.42) and (26.43), respectively,

$$\Omega^2 = \frac{1}{2}(\Omega_L^2 + \delta^2) \pm \frac{1}{2}\left[(\Omega_L^2 - \delta^2)^2 + \frac{2\delta\omega_{pe}^2 e^2 K^2 |\mathbf{A}_0 F_{p,l}|^2}{\omega_0 m_e^2 c^2}\right]^{1/2}, \quad (26.44)$$

and

$$\Omega^2 = \frac{1}{2}(\Omega_{s,I}^2 + \delta^2) \pm \frac{1}{2}\left[(\Omega_{s,I}^2 - \delta^2)^2 + \frac{2\delta\omega_{pe}^2 Z_i^2 e^2 K^2 |\mathbf{A}_0 F_{p,l}|^2}{\omega_0 m_e m_i c^2}\right]^{1/2}, \quad (26.45)$$

which exhibit oscillatory modulational instabilities.

26.4 Summary and Conclusions

In this paper, we have discussed two new aspects of collective interactions in an unmagnetized quantum plasma. First, we have reported the existence of novel attractive force that can bring ions closer. The appearance of the novel attractive force is attributed to tunneling of degenerate electrons through the Bohmian potential on account of overlapping electron wave functions at atomic dimensions. There are several consequences of our newly found short-range attractive force at quantum scales. For example, due to the trapping of ions in the negative part of the exponential oscillating-screened Coulomb potential, there will arise ordered ion structures/ion clustering depending on the electron density concentration, which in fact controls the Wigner-Seitz radius r_0. The formation of ion clusters/ion atoms will emerge as new features in a dense quantum plasma. Furthermore, we have shown that both the electron and ion plasma oscillations in quantum plasmas can be excited by a large amplitude CPEM waves due to stimulated Raman and Brillouin scattering instabilities. We also have the possibility of the modulational instabilities of the CPEM waves, via which non-resonant electron density perturbations are created. Hence, there are enhanced electrostatic fluctuations at nanoscales in dense quantum plasmas. In conclusion, we stress that the results of the present investigation are useful for understanding the novel phenomena of ion clustering and the excitation of density fluctuations by the CPEM waves at nanoscales in high-energy density plasmas that are of interest for inertial confinement fusion schemes. Specifically, stimulated scattering instabilities of a coherent high-frequency short-wavelength radiation can offer a possible mechanism for the quantum free-electron lasers in the x-ray and gamma-ray regimes, in addition to serving a diagnostic tool for determining the plasma parameters (viz, the plasma number density), once the spectra of enhanced density fluctuations are experimentally measured by using x-ray spectroscopic techniques. Finally, the present investigation, which has revealed the new physics of collective interactions between an ensemble of electrons at nanoscales, will open a new window for research in one of the modern areas of physics dealing with strongly correlated degenerate electrons and non-degenerate mildly coupled ions in dense quantum plasmas that share knowledge with cooperative phenomena (e.g. the formation of ion lattices) in condensed matter physics and in astrophysics.

Acknowledgments This work was supported by the Deutsche Forschungsgemeinschaft through the project SH21/3-2 of the Research Unit 1048. The authors thank Lennart Stenflo and Massoud Akbari-Moghanjoughi for useful discussions.

References

1. Shukla PK, Eliasson B (2010) Phys Usp 53:51
2. Shukla PK, Eliasson B (2011) Rev Mod Phys 83:885
3. Vladimirov SV, Tyshetskiy YuO (2011) Phys Usp 54:1243.
4. Haas F (2011) Quantum plasmas: an hydrodynamical approach. Springer, New York
5. van Horn HM (1991) Science 252:384
6. Fortov VE (2009) Phys Usp 52:615
7. Schatzman E (1958) White dwarfs. Interscience Publishers, New York
8. Liebert J (1980) Ann Rev Astron Astrophys 18:363
9. Shapiro SL, Teukolsky SL (1983) Black holes, White Dwarfs and neutron stars: the physics of compact objects. John Wiley & Sons, New York
10. Abrahams AM, Shapiro SL (1991) Astrophys J 374:652
11. Balberg S, Shapiro SL (2000) The Properties of Condensed Matter in White Dwarfs and Neutron. In: Levy M (ed) Stars. Academic, London
12. Lai D (2001) Rev Mod Phys 73:629
13. Harding AK, Lai D (2006) Rep Prog Phys 69:2631
14. Glenzer SH, Landen OL, Neumayer P et al (2007) Phys Rev Lett 98:065002
15. Murillo MS (2010) Phys Rev E 81:036403
16. Chapman DA, Gericke DO (2011) Phys Rev Lett 107:165004
17. Kremp D et al (1999) Phys Rev E 60:4725
18. Eliezer S, Norreys P, Mendonça JT, Lancaster K (2005) Phys Plasmas 12:052115
19. Malkin VM, Fisch NJ, Wurtele JS (2007) Phys Rev E 75:026404
20. Azechi H et al (1991) Laser Part Beams 9:193
21. Kodama R et al. (2001) Nature 412:798 (London).
22. Glenzer S, Redmer R (2009) Rev Mod Phys 81:1625
23. Redmer R, Röpke G (2010) Contrib Plasma Phys 50:970
24. Kritcher AL, Neumayer P, Castor J et al (2008) Science 322:69
25. Markowich PA et al (1990) Semiconductor equations. Springer, Berlin
26. Crouseilles N et al (2008) Phys Rev B 78:155412
27. Landau LD, Lifshitz EM (1980) Statistical physics. Butterworth-Heinemann, Oxford
28. Chandrasekhar S (1935) Mon Not R Astron Soc 113:667
29. Chandrasekhar S (1939) An introduction to the study of stellar structure. Chicago University Press, Chicago
30. Chandrasekhar S (1984) Science 226:4674
31. Salpeter EE (1961) Astrophys J 134:669
32. Wilhelm HL (1971) Z. Physik 241:1
33. Gardner CL, Ringhofer C (1996) Phys Rev E 53:157
34. Manfredi G, Haas F (2001) Phys Rev B 64:075316

35. Manfredi G (2005) Fields Inst Commun 46:263
36. Tsintsadze NL, Tsintsadze LN (2009) EPL 88:35001
37. Mendonça JT (2011) Phys Plasmas 18:062101
38. Brey L et al (1990) Phys Rev B 42:1240
39. Hedin L, Lundqvist BI (1971) J Phys C Solid State Phys 4:2064
40. Shukla PK (2009) Nature Phys 5:92
41. Shukla PK, Eliasson B (2012) Phys Rev Lett 108:219902
42. Shukla PK, Eliasson B (2012) Phys Rev Lett 108:165007
43. Atrazhev VM, Iakubov IT (1995) Phys Plasmas 2:2624
44. Frenkel YI (1946) Kinetic theory of liquids. Clarendon, Oxford
45. Ichimaru S, Tanaka S (1986) Phys Rev Lett 56:2815
46. Tanaka S, Ichimaru S (1987) Phys Rev A 35:4743
47. Ichimaru S, Iyetomi H, Tanaka S (1987) Phys Rep 149:91
48. Berkovsky MA (1992) Phys Lett A 166:365
49. Ichimaru S (1994) Statistical plasma physics: condensed plasmas. Addison Wesley, New York
50. Kaw PK, Sen A (1998) Phys Plasmas 10:3552
51. Rao NN, Shukla PK, Yu MY (1990) Planet Space Sci 38:543
52. Shukla PK, Eliasson B (2008) Phys Lett A 372:2897
53. Kittel C (1986) Introduction to solid state physics. John Wiley & Sons Inc, New York p 83
54. Kruer WL (1973) The physics of laser plasma interactions. Addison-Wesley Publishing Company, Redwood City
55. Shukla PK, Rao NN, Yu MY, Tsintsadze NL (1986) Phys Rep 138:1
56. Stenflo L (1976) Phys Scr 14:320
57. Yan YT, Dawson JM (1986) Phys Rev Lett 57:1599
58. Huang Z, Kim K-J (2007) Phys Rev ST Accel Beams 10:034801
59. Eliasson B, Shukla PK (2012) Phys Rev E 85:065401(R)
60. Shukla PK, Stenflo L (2006) Phys Plasmas 13:044505
61. Shukla PK, Eliasson B (2007) Phys Rev Lett 99:096401
62. Shukla PK, Eliasson B, Stenflo L (2012) Phys Plasmas 19 (submitted).
63. Allen L, Beijersbergen MW, Spereeuw RJC, Woerdman JP (1992) Phys Rev A 45:8185
64. Mendonça JT, Thide B, Then H (2009) Phys Rev Lett 102:185005
65. Yu MY, Spatschek KH, Shukla PK, Naturforsch Z (1974) A 29:1736
66. Shukla PK, Yu MY, Spatschek KH (1975) Phys Fluids 18:265
67. Murtaza G, Shukla PK (1984) J Plasma Phys 31:423

Chapter 27
String Theory and Regularisation of Space–Time Singularities

Martin O'Loughlin

Abstract After a discussion of the general status of space-time singularities in string theory we will concentrate on the case of a singular space-time as seen by an observer moving at the speed of light and we will show that the resulting metric has a very simple universal structure provided that the original space-time had a stress-energy source that satisfies the dominant energy condition. String theory in such singular homogeneous plane wave backgrounds is exactly solvable and the essential details can be described by matrix quantum mechanics with a time-dependent potential. We discuss in some detail the regularisation of this matrix quantum mechanics.

27.1 Introduction

String theory as a quantum theory of gravity has had a considerable amount of success. It is to date the only known consistent quantum theory of gravity that also includes a more or less standard particle physics picture. In the field of quantum gravity it has also had considerable success in elucidating the interesting phenomenon of black hole thermodynamics. There are precise quantitative results identifying the stringy/gravitational degrees of freedom that give rise to black hole entropy. The study in non-perturbative string theory of scattering from black holes has also led to the extremely powerful conjecture of AdS/CFT holography.

Singularities in General Relativity and the stringy modification thereof, are intuitively signalled by a divergent curvature as one approaches a point, or surface in a space-time. When we say divergent curvature we really mean that some curvature invariant is divergent. However singularities may also arise in other ways, often as an "incompleteness" or conical singularity of the space-time manifold. The "invariant" way to identify many singularities of a space-time manifold is to show that there

M. O'Loughlin (✉)
University of Nova Gorica, Gorica, Slovenia

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_27,

exists a geodesic which runs to a boundary of the space-time manifold at finite affine parameter.

Singularities have been shown to arise generically for solutions to Einsteins equations - this is the substance of the Hawking-Penrose singularity theorems. The most well studied examples arise in black hole and cosmological space-times and can be broadly classified by their orientation and strength.

- Space-like singularities, the most well-known examples being the big bang of Friedmann Robertson Walker cosmology and that of the Schwarzschild Black Hole.
- Time-like, the classic examples are those inside the inner horizon of Reissner-N\"ordstrom and Kerr(-Newman) black holes. In string theory time-like singularities also arise in the compactified part of the space-time: orbifolds and conifolds are the simplest examples.
- Null–conjectured to arise under generic perturbations of inner horizons. In the context of string theory there has been much recent investigation into the properties of Singular (Homogeneous) Plane Waves which possess a null singularity of particular simplicity.

There are various approaches to studying the physics of singularities.

- Singularities may be resolved in a geometric sense, for instance replacing the region close to and including the singularity by a smooth geometry. Conifold singularities are resolved in this way.
- The metric remains singular, however it may turn out that extra degrees of freedom concentrated at the location of the singularity mean that physical processes in the presence of the singularity remain well-defined.
- There is some other quantum gravity related resolution: Fuzzballs that hide horizons and consequently also singularities; Loop Quantum Gravity has a minimum distance element and thus there is encoded in this theory an upper bound on the space-time curvature.
- Gravitation and space-time enter a non-geometrical phase: a gas of black holes; a new phase of quantum geometry; or an alternative Yang-Mills description of the physics.

There is much that one could say about these different approaches, of which one could also make a complete catalogue. We will concentrate on null singularities and their possible non-geometric resolution via non-perturbative physics in String Theory. For more details refer to [9].

27.2 String Theory and Singularities

As opposed to the essentially point particles of quantum field theory strings are extended fundamental objects and consequently they may actually "see" singularities differently. It is in fact well known that a variety of time-like singularities can actually be resolved when they are probed by strings rather than point particles. For example, in the case of orbifold singularities (essentially conical singularities) the string itself couples not just to the metric but also to an additional anti-symmetric tensor field. There are additional stringy states that are always attached to the singularities the net effect of which is the absence of a true singularity thus giving rise to the smooth propogation of a string across an orbifold.

In the case of space-like singularities the situation is considerably more complex as one is forced to deal with time-dependent backgrounds and out of equilibrium systems. The time-like singularity of a conifold for example is resolved due to the condensation of non-perturbative states that become massless at the geometrical singularity, in a time-dependent and non-adiabatic system we have very few tools that enable us to consider such phase transitions. In the following we will illustrate how one may neverthless use string theory to study null singularities (as a limit of space-like ones).

Consider a time-dependent metric of the form

$$ds^2 = -f(t)\,dt^2 + g(t)\,dr^2 + r^2 d\Omega_d^2$$

and assume that there is a singularity at $t = 0$. All common space-like singularities have such a form where $f(t)$ and $g(t)$ are to leading order simply powers of t. Now to simplify our problem (without hopefully removing all the possibly interesting physics related to singualrities) we will take a particular limit, the Penrose limit, of this space-time.

The Penrose Limit corresponds to zooming into the space-time in a tubular neighbourhood of a null geodesic. The profile of the plane-wave that results upon taking the Penrose limit of a metric with respect to the null geodesic $g(u)$ is

$$ds^2 = -2du\,dv + A_{ab}(u)\,z^a z^b du^2 + d\vec{z}^2$$

with

$$A_{ab}(u) = -R_{aubu}(\gamma(u))$$

where on the right hand side we have frame components of the curvature tensor of the original metric evaluated along the null geodesic.

Thus, the Penrose limit is actually encoding some physical information about the original metric giving an exact description of the space-time along the null geodesic. The geometrical significance of the wave-profile $A_{ab}(u)$ is that it is the transverse null geodesic deviation matrix along $g(u)$ of the original metric

$$\frac{d^2}{du^2} Z^a = A_{ab}(u)\,Z^b$$

where Z is the transverse geodesic deviation vector. More precisely the Penrose limit encodes the information about tidal forces along the corresponding null geodesic in the original geometry. For a singular homogeneous plane wave we have

$$A_{ab}(u) = \frac{1}{u^2} A_{ab}$$

and the SHPW's have a divergent tidal force thus retaining this important feature of the singular gravitational field in the original geometry as $u \to 0$.

In [4] we demonstrated that:

Penrose Limits of spherically symmetric space-like or time-like singularities of power-law type satisfying (but not saturating) the Dominant Energy Condition (DEC) are singular homogeneous plane waves with profile

$$A_{ab}(u) = -\omega_a^2 \delta_{ab} u^{-2}$$

27.3 Yang-Mills from Discrete Light Cone Quantization

The Discrete Light Cone Quantization construction applied to string theory (Sen, Seiberg) is naturally adapted to SHPW's due to their symmetries and geometrical structure. Applied to the metric

$$ds^2 = -2du dv + A_{ab} z^a z^b \frac{du^2}{u^2} + d\vec{z}^2$$

and expanding the resulting Dirac-Born-Infeld D-string action around a classical solution, one finds that the fluctuations around this trajectory are described by the action

$$S = \frac{1}{2} \int d^2\sigma \left(-\eta^{\alpha\beta} \partial_\alpha z^a \partial_\beta z_b + A_{ab}(t) z^a z^b + \frac{1}{2} g_{YM}^2 \left[z^a, z^b \right]^2 \right)$$

and the Yang-Mills coupling is related to the original dilaton

$$g_{YM} \sim \frac{1}{g_s l_s} e^{-\phi}$$

In the following we concentrate on negative frequency squared for which case we have a strong string coupling singularity corresponding to weak Yang-Mills coupling. Thus there is hope that strongly coupled string theory may have an alternative description in terms of a weakly coupled Yang-Mills theory. An analysis of classical and quantum mechanics with this lagrangian [8] shows that near the singularity the typical quartic interaction is not important compared to the time dependent mass-term.

An attractive picture suggested in the original paper by CSV, was that near the singularity the weak coupling should give rise to a highly non-geometric picture of the space-time, where the D-string position is no longer described by a set of commuting coordinates but rather by a set of non-commuting matrices. Furthermore there was a significant hope that this new picture would lead to a tractable picture of the near singularity physics in terms of these non-singular and non-geometric fields. However suggestive this picture, it has been notoriously difficult to make this quantitative.

Nevertheless the general analysis of string theory in the singular plane-wave background, following classical and semi-classical reasoning [8] for dynamics in the time-dependent potential leads one to consider the considerably simpler lagrangian ($t = u$)

$$S_{bc} = -\frac{1}{2}\int d\sigma^2 \left(\eta^{\alpha\beta}\partial_\alpha z^a \partial_\beta z^b - \omega_a^2(t)\, z^a z^a\right)$$

To analyse this we will restrict to $1+1$ spacetime dimensions and decompose into Fourier modes in the spatial s worldsheet direction, leading to an infinite number of quantum mechanical systems labelled by the level n

$$S_{bc} = -\frac{1}{2}\int dt \left(\dot{z}^2 - \omega_n(t)^2 z^2\right)$$

where

$$\omega_n(t)^2 = \frac{a(1-a)}{t^2} + kn^2$$

and a appears also in the corresponding dilaton field

$$\varphi(t) = -4a(a-1)\log(|t|)$$

One can easily show by basic quantum mechanics that this system is singular at $t \to 0$. Is there an (essentially) unique way to propagate this system through the singularity?

27.4 Regularisation of the Singularity

A geometrical regularisation of the singularity consists in replacing the singular frequency by a smooth function with a parameter ε such that when the parameter is taken to zero this function returns to the original singular one. The simplest requirement that such a regularisation should satisfy is to provide a finite propagator between a time preceeding the singularity to a time in the future of the singularity. It turns out that a good regularisation (not the only possibility) is

$$\omega_\epsilon(t)^2 = a(1-a)\frac{\left(t^2 - \beta\epsilon^2\right)}{\left(t^2+\epsilon^2\right)^2}$$

and the corresponding propagator is

$$\lim_{\epsilon\to 0} F\left(t;s\right) = \frac{1}{1-2\mathrm{a}}\left(q\left(t^{a}\right)\left(s^{1-a}\right) + q^{-1}\left(s^{a}\right)\left(t^{1-a}\right)\right)$$

where $q = \pm 1$. This result is interesting in that it does not correspond to a naïve analytic continuation of the original propagator, however it is also unsatisfactory as it does not provide any further intuition about the possible physical mechanism that allows propagation through the origin. To further investigate this question we also looked at the propagator between any finite t and the origin finding that it is always singular. More elaborate regularisations can also make this propagator *F(t;0)* finite but at the expense of introducing additional free parameters into the final result. One of the main hopes of this study was that one could regularise in such a way that there are no (or at most just one) free parameter so we deem this situation to be unsatisfactory. We can still read off some general lessons about the behaviour of strings near null singularities and we will discuss these in the following section.

27.5 Discussion

For the case of most interest, with strong string coupling divergent at $t = 0$, we find that there is a good regularisation but it is still not completely clear how this regularises the physics as the string grazes the singularity and runs off to infinity and thus out of the Penrose Limit of the original space-time. This seems to indicate that either the Penrose limit, or the DLCQ procedure is removing some degrees of freedom that may be important at large *z,* causing the resulting physics to remain singular. In addition, in taking these limits one needs to confront various problems with the order of limits and these need to be analysed more carefully.

The Penrose limit is indeed the leading term in an expansion of the space-time metric around a null geodesic, the Penrose-Fermi expansion, and it would be a useful exercise to study the next terms in this expansion to investigate further the possibility that the singularity may be removed be some additional degrees of freedom.

As mentioned above the original conjecture of CSV [5] involved a highly non-Abelian phase of the 1+1 dimensional Yang-Mills theory, a phase that is very weakly coupled. At this point one should also attempt to construct alternative models that enable an analytic treatment of this highly non-geometric phase given that our current and simple minded approach does not lead to a resolution.

Acknowledgments I would like to thank my collaborators, Matthias Blau and Lorenzo Seri, for all of the time and effort that they have put into this project.

References

1. Penrose R (1976) Any spacetime has a plane wave as a limit. Reidel, Dordrecht 271–275.
2. Seiberg N (1997) Why is the matrix model correct? Phys Rev Lett 79:3577
3. Sen A (1997) D0 branes on Tn and matrix theory. Adv Theor Math Phys 2:51

4. Blau M, Borunda M, O'Loughlin M, Papadopoulos G (2004) Penrose Limits and Spacetime singularities. Class Quantum Gravity 21:L43–L49
5. Craps B, Sethi S, Verlinde EP (2005) A matrix big bang. JHEP 0510:005
6. Blau M, O'Loughlin M (2008) DLCQ and plane wave matrix big bang models. JHEP 0809:097
7. Craps B, De Roo F, Evnin O (2008) Quantum evolution across singularities: The case of geometrical resolutions. JHEP 0804:036
8. O'Loughlin M, Seri L (2010) The non-abelian gauge theory of matrix big bangs. JHEP 1007:036
9. Blau M, O'Loughlin M, Seri L (2011) Aspects of plane wave (matrix) string dynamics. arXiv:1112.3182.

Chapter 28
Fuzzy Space–Time, Quantization and Gauge Invariance

S. N. Mayburov

Abstract Quantum space-time with Dodson-Zeeman topological structure is studied. In its framework the states of massive particle m correspond to elements of fuzzy set called fuzzy points. Due to their weak ordering, m space coordinate x acquires principal uncertainty σ_x. Quantization formalism is derived from consideration of m evolution in fuzzy phase space with minimal number of additional assumptions. Particle's interactions on fuzzy manifold are studied and shown to be gauge invariant.

Structure of space-time at microscopic (Plank) scale and its relation to axiomatic of Quantum Mechanics (QM) is actively discussed now [1, 2]. In particular, it was proposed that such fundamental properties of space-time manifold M_{ST} as its metrics and topology can differ significantly at Planck scale from standard Riemanian formalism [2, 3]. In particular, Posets and the fuzzy ordered sets (Fosets) were used for the construction of different variants of the novel fuzzy topology (FT) [4, 5]. Hence it's instructive to study what kind of physical theory such topology induces [1, 3]. In our previous works it was shown that in its framework the quantization procedure by itself can be defined as the transition from ordered phase space to fuzzy one. Therefore, the quantum properties of particles and fields are induced directly by FT of their phase space and don't need to be postulated separately of it [1, 3]. As simple example of such transition the quantization of nonrelativistic particle was studied; it was argued that FT induces the particle's dynamics which is equivalent to QM evolution [1, 3]. Yet in its derivation some phenomenological assumptions were used, here new and rather simple formalism will be described.It will be shown also that the interactions on such fuzzy manifold are gauge invariant and under simple assumptions correspond to Yang-Mills fields [3]. It is worth to mention here the extensive studies of noncommutative fuzzy spaces, both finite (sphere, tori) and infinite ones, these are, in fact, alternative variant of such approach based on operator algebras [6].

S. N. Mayburov (✉)
Lebedev Institute of Physics, Leninsky Prospect 53, Moscow RU-117924, Russia,
e-mail: mayburov@sci.lebedev.ru

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_28,

Here we consider only the most important steps in construction of mechanics on fuzzy manifold called fuzzy mechanics (FM), the details can be found in [1, 3] . In 1-dimensional Euclidean Geometry, the elements of its manifold X are the points x_a which constitute the ordered set. For the elements of partially ordered set (Poset) $\{d_i\}$, beside standard ordering relation between its elements $d_k \le d_l$ (or vice versa), the incomparability relation $d_k \sim d_l$ is also permitted; if it's true, then both $d_k \le d_l$ and $d_l \le d_k$ propositions are false. To illustrate it, consider Poset $D^T = A \cup B$, which includes the subset of 'incomparable' elements $B = \{b_j\}$, and ordered subset $A = \{a_i\}$. In A the element's indexes grow correspondingly to their ordering, so that $\forall\, i,\, a_i \le a_{i+1}$. As the example, consider some A interval $\{a_l, a_n\}$ and suppose that $b_j \in \{a_l, a_n\}$, i.e. $a_l \le b_j;\ b_j \le a_n$ and $b_j \sim a_i;$ iff $l \le i \le n$. In this case, b_j in some sense is 'smeared' over $\{a_l, a_{l+n}\}$ interval, which is analogue of b_j coordinate uncertainty relative to A 'axe'. To introduce the fuzzy relations, let's put in correspondence to each b_j, a_i pair the weight $w_i^j \ge 0$ with the norm $\sum_i w_i^j = 1$. Then D^T is fuzzy set, b_j called fuzzy point. Continuous 1-dimensional fuzzy set C^F is defined analogously; $C^F = B \cup X$ where B is the same as above, X is the continuous ordered subset, which is equivalent to R^1 axis of real numbers. Correspondingly, fuzzy relation between b_j, x_a are described by $w^j(x_a) \ge 0$ with norm $\int w^j dx_a = 1, b_j$ uncertainty of X coordinate described by w^j dispersion. Note that in FT $w^j(x)$ doesn't have any probabilistic meaning but only the geometric one.

In these terms the particle's state in 1-dimensional classical mechanics corresponds to ordered point $x(t)$ in X. Analogously to it, in 1-dimensional fuzzy mechanics (FM) the particle m corresponds to fuzzy point $b(t)$ in C^F; it characterized by normalized positive density $w(x,t)$. Yet m fuzzy state $|g\}$ can depend also on other m degrees of freedom (DF). The obvious one is $\frac{\partial w}{\partial t}$, yet it's more convenient to replace it by related DF, which describes w flow velocity $v(x,t)$, it permit to decompose formally w flow $j(x)$, as $j = wv$. Assuming FM to be local theory, flow continuity equation should hold :

$$\frac{\partial w}{\partial t}(x) = -\frac{\partial j}{\partial x} = -v\frac{\partial w}{\partial x} - \frac{\partial v}{\partial x}w \tag{28.1}$$

Alternatively, locality violation induces nonlocal w correlations incompatible with FT and causality. We replace $v(x)$ by equivalent DF:

$$\gamma(x) = r\int_{-\infty}^{x} v(\xi)d\xi \tag{28.2}$$

where r is theory constant. If $|g\}$ doesn't depend on any other DFs, then the calculations based on theorem of [7] show that in FM m pure state unambiguously expressed as:

$$g(x) = \sqrt{w(x)}e^{i\gamma(x)}. \tag{28.3}$$

analogously to QM dirac vector in $X-$representation. Evolution equation for m evolution supposedly should be of first order in time, i.e.:

$$i\frac{\partial g}{\partial t} = \hat{H}g. \tag{28.4}$$

In general $\hat{H}$ is nonlinear operator, for simplicity we shall consider first linear case and turn to nonlinear one afterwards. Free m evolution is invariant relative to X shifts performed by the operator $\hat{W}(a) = \exp(a\frac{\partial}{\partial x})$. Because of it, $\hat{H}$ should commute with $\hat{W}(a)$ for arbitrary a, i.e. $[\hat{H}, \frac{\partial}{\partial x}] = 0$. It holds only if $\hat{H}$ is differential polinom which can be written as:

$$\hat{H} = \hat{H}_0 + \Delta\hat{H} = -c_1\frac{\partial}{\partial x} - c_2\frac{\partial^2}{\partial x^2} - \sum_{l=3}^{n} c_l\frac{\partial^l}{\partial x^l} \tag{28.5}$$

where $\Delta\hat{H}$ denotes the sum over l; $c_{1,2}, c_l$ are arbitrary real constants, $n \geq 3$. From X-reflection invariance it follows that $c_l = 0$ for noneven l. If to substitute $v(x)$ by $\gamma(x)$ in Eq. (28.1) and transform it to $\sqrt{w}$ time derivative, then left part of (28.4) is equal to:

$$i\frac{\partial g}{\partial t}(x) = -(\frac{i}{r}\frac{\partial\sqrt{w}}{\partial x}\frac{\partial\gamma}{\partial x} + \frac{i}{2r}\sqrt{w}\frac{\partial^2\gamma}{\partial x^2} + \sqrt{w}\frac{\partial\gamma}{\partial t})e^{i\gamma} \tag{28.6}$$

Imaginary terms in brackets of (28.6) and analogous terms of $e^{-i\gamma}\hat{H}_0 g$ coincide up to c_2/r ratio, from that n can be obtained. Really, imaginary part of $e^{i\gamma}\Delta\hat{H}g$ should include the term proportional to $c_n\frac{\partial^n\gamma}{\partial x^n}$, yet Eq. (28.6) includes the highest term corresponding to $n = 2$ only. Hence $\Delta H = 0$ and $c_2 = \frac{1}{2r}$, as the result m free evolution described by Schroedinger equation mass $m_0 = r$. Note that in FM it's equivalent to the system of two equations, one for w and other for γ time derivatives, their meaning will be discussed below. Plainly, $\gamma(x)$ corresponds to quantum phase, m states described by $g(x)$ are dirac vectors (rays) of Hilbert space $\mathcal{H}$ [8].

Concerning with nonlinear case, the conditions of dynamics linearity obtained by Jordan recenlty are rather weak [7]. In particular, it was shown that if evolution unambiguously maps the set of pure states onto itself, then such evolution is linear. Yet for FM such condition is generic, no mixed state can appear in free evolution of fuzzy state, and so FM evolution should be linear [1]. In FM $\hat{p}_x = -i\frac{\partial}{\partial x}$ describes m momentum and all self-adjoint operator functions $\hat{F}_Q(x, p)$ are also m observables. Note that in such formalism the commutation relations of the kind $[x, p_x] = i$ are obtained from topological premises which constitute FM basis. Planck constant $\hbar = 1$ in our FM ansatz, but the same value ascribed to it in Relativistic unit system together with $c = 1$; in FM framework $\hbar$ only connects x, p scales and doesn't have any other meaning. Generalization of FM formalism on three dimensions is straightforward and doesn't demand any serious modification of described ansatz. For relativistic free evolution the linearity of state evolution becomes the important criterion for the choice of consistent ansatz. From these premises it was found that for massive particle m the minimal solution is 4-spinor $g_i(\overrightarrow{r}, t)$; $i = 1, 4$, its evolution is described by Dirac equation for spin-$\frac{1}{2}$, i. e. such particle is fermion.

Now we shall consider briefly the particle interactions on such fuzzy manifold. Note first that by derivation FM free Hamiltonian H_0 induces $\mathcal{H}$ dynamical asymmetry between $|\overrightarrow{r}\rangle$ and $|\overrightarrow{p}\rangle$ 'axes' which is absent in standard QM formalism. As was shown, in FM m free dynamics is described by the system of two equations which define $\frac{\partial\sqrt{w}}{\partial t}$ and $\frac{\partial\gamma}{\partial t}$. Yet the first of them is equivalent to Eq. (28.1) which describes $w(x)$ balance and so is, in fact, kinematical one. Thus any m interactions can be accounted only via second equation, which for 1-dimensional case is written as:

$$\frac{\partial\gamma}{\partial t} = -\frac{1}{2m_0}[(\frac{\partial\gamma}{\partial x})^2 - \frac{1}{\sqrt{w}}\frac{\partial^2\sqrt{w}}{\partial\vec{x}^2}] + H_{int} \tag{28.7}$$

where H_{int} is interaction term. Since γ corresponds to quantum phase, it supposes that in FM all m interactions should be gauge invariant. Despite that fermion state is described by several phases the same invariance if fulfilled for them and can be extended also on relativistic case. Basing on it, QED formalism was derived by us with minimum of additional assumptions. Preliminary results for interactions of fermion multiplets show that in such theory their interactions also possess $SU(n)$ gauge invariance and transferred by corresponding Yang-Mills fields [3].

In conclusion, we have shown that the quantization of elementary systems can be derived directly from axiomatic of Set Theory and Topology together with the natural assumptions about system evolution. It allows to suppose that the quantization phenomenon has its roots in foundations of mathematics and logics [8]. The main aim of FM, as well as other studies of fuzzy spaces, is the construction of nonlocal QFT (or other more general theory). In this vein, FM provides the interesting opportunities, being generically nonlocal theory which, in the same time, is Lorentz covariant and manifests the gauge invariance.

References

1. Mayburov S (2008) Fuzzy geometry of phase space and quantization of massive fields. J Phys A 41:164071–164080 hep-th 1205.3019.
2. Isham C (1994) Introduction into canonical gravity quantization, in canonical gravity: from classical to quantum Lecture Notes in Phys. Lect Notes Phys 433:11–28
3. Mayburov S (2010) Fuzzy space-time geometry and particle's dynamics. Int J Theor Phys 4:3192–3198
4. Zeeman C (1961) Topology of brain and visual perception in topology of 3-manifolds. Prentice-Hall, New Jersey
5. Dodson C (1975) Tangent structures for hazy space. J London Math Soc 2:465–474
6. Madore J (1992) Fuzzy sphere. Class Quantum Gravity 9:69–83
7. Jordan T (2006) Assumptions of linear quantum dynamics. Phys Rev A73:022101–022108
8. Jauch J (1968) Foundations of quantum mechanics. Addison-Wesly, MA Reading

Chapter 29
On Fluid Maxwell Equations

Tsutomu Kambe

Abstract *Fluid mechanics* is a field theory of Newtonian mechanics of Galilean symmetry, concerned with fluid flows represented by the velocity field such as $\boldsymbol{v}(\boldsymbol{x}, t)$ in space-time. A fluid is a medium of continuous mass. Its mechanics is formulated by extending discrete system of point masses. Associated with two symmetries (translation and space-rotation), there are two gauge fields: $\boldsymbol{E} \equiv (\boldsymbol{v} \cdot \nabla)\boldsymbol{v}$ and $H \equiv \nabla \times \boldsymbol{v}$, which do not exist in the system of discrete masses. One can show that those are analogous to the electric field and magnetic field in the electromagnetism, and *fluid Maxwell equations* can be formulated for $\boldsymbol{E}$ and $\boldsymbol{H}$. Sound waves within the fluid is analogous to the electromagnetic waves in the sense that phase speeds of both waves are independent of wave lengths, i.e. non-dispersive.

29.1 Introduction

Fluid mechanics is a field theory of Galilean symmetry. Two symmetries are known as subgroups of the Galilean group: translation (space and time) -and space-rotation. From this point of view, a gauge-theoretic study has been developed by Kambe [1] for flows of an ideal compressible fluid with respect to both global and local invariances of the flow fields in the space–time $(\boldsymbol{x}, t)$, where $\boldsymbol{x} = (x^1, x^2, x^3)$ is the three-dimensional space coordinates.

Suppose that we have a velocity field $\boldsymbol{v}(t, \boldsymbol{x})$ depending on $\boldsymbol{x}$ and time t. Then one can define the convective derivative D_t (i.e. the Lagrange derivative in the fluid mechanics) by

$$\mathrm{D}_t \equiv \partial_t + \boldsymbol{v} \cdot \nabla, \qquad \text{where} \quad \partial_t = \partial/\partial t, \qquad \nabla = (\partial/\partial x^i). \tag{29.1}$$

T. Kambe (✉)
Department of Physics, University of Tokyo, Tokyo, Japan

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_29,

It is remarkable that D_t is *gauge-invariant*, namely D_t is invariant with respect to local gauge transformations [1]. This implies that the gauge theory can be applied to the fluid mechanics, since the covariant derivative is an essential building block of the gauge theory [2].

The fluid velocity $\boldsymbol{v}$ and acceleration $\boldsymbol{A}$ are defined by $\boldsymbol{v} = D_t\boldsymbol{x}$, and $\boldsymbol{A} = D_t\boldsymbol{v} = \partial_t\boldsymbol{v} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v}$, respectively. The second term on the right of $\boldsymbol{A}$ can be transformed by the following vector identity:

$$(\boldsymbol{v}\cdot\nabla)\boldsymbol{v} = -\boldsymbol{v}\times(\nabla\times\boldsymbol{v}) + \nabla(\frac{1}{2}|\boldsymbol{v}|^2). \tag{29.2}$$

It is verified by Kambe ([1, 2] that the vorticity defined by

$$\boldsymbol{\omega} = \nabla\times\boldsymbol{v} = \mathrm{curl}\boldsymbol{v} \tag{29.3}$$

is a *gauge field* with respect to the rotation symmetry, and the acceleration $\boldsymbol{A}$ is given by the following gauge-invariant expression: $D_t\boldsymbol{v} = \partial_t\boldsymbol{v} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v} = \partial_t\boldsymbol{v} + \boldsymbol{\omega}\times\boldsymbol{v} + \nabla(\frac{1}{2}|\boldsymbol{v}|^2)$.

In a dynamical system of n discrete point masses with their positions denoted by $\boldsymbol{X}_k(t)\,(k = 1,\ldots,n)$, the equations of motion are described by the form, $(\mathrm{d}/\mathrm{d}t)^2\boldsymbol{X}_k(t) = \boldsymbol{F}_k(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_n)$, where $\boldsymbol{F}_k$ is the force acting on the kth particle. For a fluid of continuous mass distribution, the discrete positions $\boldsymbol{X}_k(t)$ are replaced by a continuous field representation $\boldsymbol{X}(\boldsymbol{a},t)$, where $\boldsymbol{a}$ is the Lagrange parameters $\boldsymbol{a} = (a_1, a_2, a_3)$, or labels identifying fluid particles. Associated with the field representation, the time derivative $\mathrm{d}/\mathrm{d}t$ should be replaced by D_t, and the acceleration is given by $\partial_t\boldsymbol{v} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v}$. Thus, we have a *connection* term $(\boldsymbol{v}\cdot\nabla)\boldsymbol{v}$, which is the second *gauge field*.

Clearly, the two gauge fields $(\boldsymbol{v}\cdot\nabla)\boldsymbol{v}$ and $\nabla\times\boldsymbol{v}$ do not exist in discrete systems of point masses. They are the fields defined only in continuous differentiable fields such as the fluid flow. In fact, it is remarkable that they are analogous to the electric field and magnetic field in the electromagnetism, and a system of fluid Maxwell equations can be formulated for $\boldsymbol{E}\equiv(\boldsymbol{v}\cdot\nabla)\boldsymbol{v}$ and $\boldsymbol{H}\equiv\nabla\times v$ [3].

It is shown in Sect. 29.2 that there exists a similarity between the wave equations of electromagnetism and fluid mechanics. This suggests existence of correspondence between the variables of electromagnetism and fluid-mechanics. The Sect. 29.3 describes that the correspondence permits formulation of a system of fluid Maxwell equations. In Sect. 29.4 a wave equation is derived from the system for a vector field. However in the fluid case, the vector form reduces to a scalar wave equation for sound waves, since all the terms of the equation are expressed by gradients of scalar fields. In Sect. 29.5, another analogy between fluid mechanics and electromagnetism is presented for the equations of motion of a test particle in flow field and in electromagnetic field.

29.2 Equations of Fluid Mechanics, Compared with Electromagnetism

29.2.1 Equations of Fluid Mechanics

Equation of motion of an ideal fluid is given by the Euler's equation of motion:

$$\partial_t \boldsymbol{v} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = -\frac{1}{\rho}\nabla_p \tag{29.4}$$

This equation is supplemented by the followings:

$$\begin{aligned} \partial_t \rho + \boldsymbol{v} \cdot \nabla\rho + \rho\nabla \cdot \boldsymbol{v} &= 0, & (29.5)\\ \partial_t s + \boldsymbol{v} \cdot \nabla s &= 0, & (29.6)\\ \partial_t \boldsymbol{\omega} + \nabla \times (\boldsymbol{\omega} \times \boldsymbol{v}) &= 0, & (29.7) \end{aligned}$$

where ρ is the fluid density, s the entropy per unit mass, p the pressure. The Eq. (29.5) is the continuity equation, and (29.7) is the vorticity equation, while (29.6) is the entropy equation, stating that each fluid particle keeps its initial entropy (i.e. adiabatic).

If initial entropy field is uniform with a constant value s_0, the fluid keeps the *isentropic* state $s = s_0$ at any later time and everywhere. In this case, we have $(1/\rho)\nabla p = \nabla h$ by the thermodynamics where h is the enthalpy per unit mass.[1] In isentropic flows, an enthalpy variation Δh and a density variation $\Delta\rho$ are related by

$$\Delta h = \frac{1}{\rho}\Delta p = \frac{a^2}{\rho}\Delta\rho, \qquad \text{where} \quad \Delta p = a^2\Delta\rho, \quad a^2 = (\partial p/\partial\rho)_s. \tag{29.8}$$

The notation $(\partial p/\partial\rho)_s$ denotes partial differentiation with s fixed, and $a = \sqrt{(\partial p/\partial\rho)_s}$ is the *sound speed*, as becomes clear below. From the above, we have $\partial_t\rho = (\rho/a^2)\partial_t h$ and $\nabla\rho = (\rho/a^2)\nabla h$. Therefore, the Eq. (29.5) is transformed to $(\rho/a^2)(\partial_t h + \boldsymbol{v} \cdot \nabla h + a^2\nabla \cdot \boldsymbol{v}) = 0$. Thus, the fluid equations (29.4)–(29.7) reduce to the following three equations:

$$\begin{aligned} \partial_t \boldsymbol{v} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} + \nabla h &= 0, & (29.9)\\ \partial_t h + \boldsymbol{v} \cdot \nabla h + a^2\nabla \cdot \boldsymbol{v} &= 0, & (29.10)\\ \partial t\boldsymbol{\omega} + \nabla \times (\boldsymbol{\omega} \times \boldsymbol{v}) &= 0, & (29.11) \end{aligned}$$

[1] From the thermodynamics, $\mathrm{d}h = (1/\rho)\mathrm{d}p + T\mathrm{d}s$ where T is the temperature. If $\mathrm{d}s = 0$, we have $\mathrm{d}h = (1/\rho)\mathrm{d}p$.

Using (29.2), the equation of motion (29.9) is rewritten as[2]

$$\partial_t \boldsymbol{v} + \boldsymbol{\omega} \times \boldsymbol{v} + \nabla(\frac{1}{2}|\boldsymbol{v}|^2 + h) = 0. \tag{29.12}$$

29.2.2 Equations of Electromagnetism

In electromagnetism, Maxwell's equations for electric field $\boldsymbol{E}^{\mathrm{em}}$ and magnetic field $\boldsymbol{H}^{\mathrm{em}}$ are

$$\begin{aligned} \nabla \times \boldsymbol{E}^{\mathrm{em}} + c^{-1}\partial_t \boldsymbol{H}^{\mathrm{em}} &= 0, \qquad \nabla \cdot \boldsymbol{H}^{\mathrm{em}} = 0, \\ \nabla \times \boldsymbol{H}^{\mathrm{em}} - c^{-1}\partial_t \boldsymbol{E}^{\mathrm{em}} &= \boldsymbol{J}^{\mathrm{e}}, \qquad \nabla \cdot \boldsymbol{E}^{\mathrm{em}} = q^{\mathrm{e}}. \end{aligned} \tag{29.13}$$

where $q^{\mathrm{e}} = 4\pi \quad \rho^{\mathrm{e}}$ and $\boldsymbol{J}^{\mathrm{e}} = (4\pi/c)\boldsymbol{j}^{\mathrm{e}}$ with ρ^{e} and $\boldsymbol{j}^{\mathrm{e}}$ being the charge density and current density vector respectively, and c the light velocity. The vector fields $\boldsymbol{E}^{\mathrm{em}}$ and H^{em} can be defined in terms of a vector potential $\boldsymbol{A}$ and a scalar potential $\phi^{(\mathrm{e})}$ by

$$E^{\mathrm{em}} = -c^{-1}\partial_t \boldsymbol{A} - \nabla\phi^{(e)}, \qquad \mathbf{H}^{\mathrm{em}} = \nabla \times \boldsymbol{A}. \tag{29.14}$$

Using these definitions, the above Maxwell equations require that the two fields $\boldsymbol{A}$ and $\phi^{(\mathrm{e})}$ satisfy the following equations (Landau and Lifshitz [4], Chap. 8):

$$\partial_t \phi^{(\mathrm{e})} + c\nabla \cdot \boldsymbol{A} = 0 \qquad \text{(Lorentz condition)}, \tag{29.15}$$

$$(\partial_t^2 - c^2\nabla^2)\boldsymbol{A} = c^2\boldsymbol{j}^{\mathrm{e}}, \qquad (\partial_t^2 - c^2\nabla^2)\phi^{(\mathrm{e})} = c^2 q^{\mathrm{e}}. \tag{29.16}$$

29.2.3 Analogy in Wave Property

Linearizing (29.9) by neglecting $(\boldsymbol{v} \cdot \nabla)\boldsymbol{v}$, and linearizing (29.10) by neglecting $\boldsymbol{v} \cdot \nabla h$ and replacing a with a constant value a^0, we have

$$\partial_t \boldsymbol{v} + \nabla h = 0, \qquad \partial_t h + a_0^2 \nabla \cdot \boldsymbol{v} = 0. \tag{29.17}$$

Eliminating $\boldsymbol{v}$ from the two equations, we obtain the wave equation $(\partial_t^2 - a_0^2\nabla^2)h = 0$ for sound waves. Using it, we obtain the wave equation for $\boldsymbol{v}$ as well. Thus, we have

$$(\partial_t^2 - a^2\nabla^2)h = 0, \qquad (\partial_t^2 - a^2\nabla^2)\boldsymbol{v} = 0 \tag{29.18}$$

[2] The vorticity equation (29.11) is also obtained by taking curl of (29.12).

It is remarkable that we have a close analogy between the two systems of fluid and electromagnetism. In vacuum space where $q^{\rm e} = 0, \boldsymbol{J}^{\rm e} = 0$, the wave equations (29.16) reduce to

$$(\partial_t^2 - c^2\,\nabla^2)\phi^{(\rm e)} = 0, \qquad (\partial_t^2 - c^2\,\nabla^2)\boldsymbol{A} = 0. \tag{29.19}$$

It is seen that c (light speed) $\Leftrightarrow a_0$ (sound speed). Notable feature of the sound wave equations (29.18) is *non-dispersive*. Namely, the dispersion relation for waves of wave number k and frequency ω is given by $\omega^2 = a_0^2\,k^2$, and the phase speed ω/k is equal to a_0 independent of the wave length $2\pi/k$. The same is true for the equation (29.19) of the electromagnetic wave. Note that this is closely related to the system of Maxwell equations.

In addition, the second equation of (29.17), obtained from the continuity equation by linearization, is analogous to the *Lorentz condition* (29.15) by the correspondence,

$$(c, \boldsymbol{A}, \phi^{(\rm e)}) \quad \Leftrightarrow \quad (a_0, a_0\boldsymbol{v}, h).$$

This implies possibility of formulation of *fluid* Maxwell equations for which the vector potential is played by $\boldsymbol{v}$ (except for the coefficient a_0, abbreviated in the present formulation for simplicity) and the scalar potential played by h. According to this finding, let us introduce two fields defined by

$$\boldsymbol{E} = -\partial_t\boldsymbol{v} - \nabla h, \qquad \boldsymbol{H} = \nabla \times \boldsymbol{v}.$$

Consider the following transformations: $\boldsymbol{v}' = \boldsymbol{v} + \nabla f$, $h' = h - \partial_t f$. The vector fields $\boldsymbol{E}'$ and $\boldsymbol{H}'$ defined by $\boldsymbol{v}'$ and h' are unchanged, namely we have $\boldsymbol{E}' = \boldsymbol{E}$ and $\boldsymbol{H}' = \boldsymbol{H}$. Thus in fluid flows, there exists the same gauge invariance as that in the electromagnetism.

29.3 Fluid Maxwell Equations

Based on the following definition of the two vector fields $\boldsymbol{E}$ and $\boldsymbol{H}$ (analogous to (29.14)),

$$\boldsymbol{E} = -\partial_t\boldsymbol{v} - \nabla h, \qquad \boldsymbol{H} = \nabla \times \boldsymbol{v}. \tag{29.20}$$

Fluid Maxwell Equations can be derived from the fluid equations (29.9)–(29.11) as follows:

$$\text{(A)} \quad \nabla\cdot\boldsymbol{H} = 0, \qquad \text{(B)} \quad \nabla\times\boldsymbol{E} + \partial_t\boldsymbol{H} = 0, \tag{29.21}$$

$$\text{(C)} \quad \nabla\cdot\boldsymbol{E} = q, \qquad \text{(D)} \quad a_0^2\nabla\times\boldsymbol{H} - \partial_t\boldsymbol{E} = \boldsymbol{J}, \tag{29.22}$$

$$\text{Where} \qquad q \equiv \nabla\cdot[(v\cdot\nabla)v], \qquad \boldsymbol{J} \equiv \partial_t^2 v + \nabla\partial_0^2\nabla\times(\nabla\times v), \tag{29.23}$$

and a_0 is a constant (the sound speed in undisturbed state). From the calculus ∂_t(C)+ div(D) operated on the two equations of (29.22), we have the *charge conservation*: $\partial_t q + \mathrm{div}\boldsymbol{J} = 0$. Using (29.9) and the definition $\boldsymbol{E} = -\partial_t \boldsymbol{v} - \nabla h$, the *fluid-electric* field $\boldsymbol{E}$ is given by

$$\boldsymbol{E} = (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} \quad = \boldsymbol{\omega} \times \boldsymbol{v} + \nabla(\frac{1}{2}|\boldsymbol{v}|^2). \tag{29.24}$$

Hence, the *charge density* q is given by

$$q = \nabla \cdot \boldsymbol{E} = \mathrm{div}[(\boldsymbol{v} \cdot \nabla)\boldsymbol{v}]. \tag{29.25}$$

29.3.1 Derivation

Derivation of the fluid Maxwell equations (29.21) and (29.22) is carried out as follows.

(a) Equation (A) is deduced immediately from the definition of $\boldsymbol{H} = \nabla \times \boldsymbol{v} = \boldsymbol{\omega}$.
(b) Equation (B) is an identity obtained from the definition (29.20). Moreover, if the expression (29.24) is substituted to $\boldsymbol{E}$ and $\boldsymbol{\omega}$ to $\boldsymbol{H}$, then the equation (B) reduces to the vorticity equation (29.11).
(c) Equation (C) is just div [Eq. (29.24)] with q defined by (29.23).
(d) Equation (D) can be derived in the following way. Applying ∂_t to $\boldsymbol{E} = -\partial_t \boldsymbol{v} - \nabla h$, we obtain

$$-\partial_t \boldsymbol{E} - \partial_t^2 \boldsymbol{v} = \nabla \partial_t h,$$

Adding the term $a_0^2 \nabla \times \boldsymbol{H} = a_0^2 \nabla \times (\nabla \times \boldsymbol{v})$ on both sides, this can be rearranged as follows:

$$a_0^2 \nabla \times \boldsymbol{H} - \partial_t \boldsymbol{E} = \boldsymbol{J}, \qquad \boldsymbol{J} = \partial_t^2 v + \nabla \partial_t h + a_0^2 \nabla \times (\nabla \times v).$$

which is nothing but the equation (D). The vector $\boldsymbol{J}$ can be given another expression by using $\partial_t h = -(\boldsymbol{v} \cdot \nabla)h - a^2 \nabla \cdot \boldsymbol{v}$ from the continuity Eq (29.10):

$$J = \partial_t^2 v + a_0^2 \nabla \times (\nabla \times v) - \nabla(a^2 \nabla \cdot v) - \nabla((v \cdot \nabla)h)$$

This can be rewritten as $\boldsymbol{J} = (\partial_t^2 - a_0^2 \nabla^2)\boldsymbol{v} + \boldsymbol{J}^*$, and $\boldsymbol{J}^* = a_0^2 \nabla(\nabla \cdot \boldsymbol{v}) - \nabla a^2 \nabla \cdot \boldsymbol{v} - \nabla(\boldsymbol{v} \cdot \nabla h)$, where the following identity is used:

$$\nabla(\nabla \cdot \boldsymbol{v}) = \nabla \times (\nabla \times \boldsymbol{v}) + \nabla^2 \boldsymbol{v}. \tag{29.26}$$

29.4 Equation of Sound Wave

Suppose that a localized flow is generated at an initial instant in otherwise uniform state at rest, where undisturbed values of the pressure, density and enthalpy are respectively p_0, ρ_0 and h_0. Equation of sound wave is derived from the system of fluid Maxwell equations (A)–(D) as follows.

Differentiating Eq.(D) with respect to t, and eliminating $\partial_t \boldsymbol{H}$ by using (B), we obtain

$$\partial_t^2 \boldsymbol{E} + a_0^2 \nabla \times (\nabla \times \boldsymbol{E}) = -\partial_t \boldsymbol{J}. \tag{29.27}$$

The second term on the left can be rewritten by using the identity (29.26), with $\boldsymbol{v}$ replaced by $\boldsymbol{E}$. Then the Eq. (29.27) reduces to

$$(\partial_t^2 - a_0^2 \nabla^2)(\boldsymbol{E} + \partial_t v) = -a_0^2 \nabla(\nabla \cdot \boldsymbol{E}) - \partial_t \boldsymbol{J}^*, \tag{29.28}$$

$$J^* = \nabla((a_0^2 - a^2)\nabla \cdot v) - \nabla(v \cdot \nabla h) \equiv a_0^2 \nabla \hat{Q}, \quad \hat{Q} = (1 - \hat{a}^2)\nabla \cdot v - a_0^{-2}(v \cdot \nabla)h, \tag{29.29}$$

where $\hat{a} = a/a0$. We have $\boldsymbol{E} + \partial_t \boldsymbol{v} = -\nabla h$ from (29.20), and also $\partial_t \boldsymbol{J}^* = a^2 \nabla \partial_{0t} \hat{Q}$ from (29.29). Therefore, we can integrate (29.28) spatially, since all the terms are of the form of gradient of scalar fields. Dividing (29.28) with $-a_0^2$ and integrating it, we obtain the following wave equation:

$$(a_0^{-2}\partial_t^2 - \nabla^2)\tilde{h} = s(x,t), \qquad s(x,t) \equiv \nabla \cdot \boldsymbol{E} + \partial_t \hat{Q}, \tag{29.30}$$

where $\tilde{h} \equiv h - h_0 = (p - p_0)/\rho$. Thus, the vectorial form of wave equation (29.28) has been reduced to the equation for a scalar field $\tilde{h}$ (see the first of (29.18)). The term $S(\boldsymbol{x}, t)$ is a source of the wave. Using (29.24) and (29.25), we obtain an explicit form of the first $\nabla \cdot \boldsymbol{E}$ of the source S as $\nabla \cdot \boldsymbol{E} = \mathrm{div}(\boldsymbol{\omega} \times \boldsymbol{v}) + \nabla^2 \frac{1}{2} v^2$. The first term $\mathrm{div}(\boldsymbol{\omega} \times \boldsymbol{v})$ implies that the motion of $\boldsymbol{\omega}$ generates sound waves. This is the source term of the *Vortex sound* [5], and contribution from the second term $\nabla^2 \frac{1}{2} v^2$ vanishes in an ideal fluid in which total kinetic energy $\int \frac{1}{2} v^2 d^3 x$ is conserved. Mach number of the source flow is defined by $M = |\boldsymbol{v}|/a_0$, then the second term $\partial_t \hat{Q}$ of S is $O(M^2)$, namely, higher order if M is small enough.

29.5 Equation of Motion of a Test Particle in a Flow Field

Analogy between fluid mechanics and electromagnetism is also found in the equation of motion of a test particle in a flow field as well. Suppose that a *test particle* of mass m is moving in a flow field $\boldsymbol{v}(\boldsymbol{x}, t)$, which is *unsteady, rotational and compressible*. The size of the particle and its velocity are assumed to be so small that its influence on the background velocity field $\boldsymbol{v}(\boldsymbol{x}, t)$ is negligible, namely the velocity field $\boldsymbol{v}$ is regarded as independent of the position and velocity of the particle.

The particle velocity is defined by $\boldsymbol{u}(t)$ relative to the fluid velocity $\boldsymbol{v}$. Then, the total particle velocity is $\boldsymbol{u} + \boldsymbol{v}$. In this circumstance, the ith component of total momentum P_i associated with the test particle moving in the flow field is expressed by the sum: $P_i = mu_i + m_{ik}u_k$,[3] and the equation of motion of the particle is given by

$$\frac{d}{dt}\boldsymbol{P} = m\boldsymbol{E} + mu \times \boldsymbol{H} = m\nabla\phi_g, \tag{29.31}$$

[2], where $\boldsymbol{E}$ and $\boldsymbol{H}$ take the same expressions as those of (29.20), and $\boldsymbol{P} = (P_i)$, $P_i = mu_i + m_{ik}u_k$, and $\phi_g = gz$. Obviously, the Eq. (29.31) is analogous to the equation of motion of a *charged* particle in an electromagnetism of electric field $\boldsymbol{E}^{\text{em}}$ and magnetic field $\boldsymbol{H}^{\text{em}}$:

$$\frac{\mathrm{d}}{\mathrm{d}t}(mv_e) = eE^{\text{em}} + (e/c)v_e \times H^{\text{em}} - m\nabla\Phi_g, \tag{29.32}$$

where $\boldsymbol{v}_e$ is the velocity of a charged particle, and c the light velocity. Rewriting the second term of (29.31) as $(m/a_0)\boldsymbol{u}\times(a_0\boldsymbol{H})$, and comparing the first two terms on the right of (29.31) and (29.32), it is found that there is correspondence: $e \Leftrightarrow m$, $\boldsymbol{E}^{\text{em}} \Leftrightarrow \boldsymbol{E}$, and $\boldsymbol{H}^{\text{em}} \Leftrightarrow a_0\boldsymbol{H}$.

29.6 Summary

It is shown that there exist similarities between electromagnetism and fluid mechanics. The correspondence between them permits formulation of a system of fluid Maxwell equations. It is found that the sound wave in the fluid is analogous to the electromagnetic wave, in the sense that phase speeds of both waves are independent of wave lengths, i.e. non-dispersive. Another analogy between fluid mechanics and electromagnetism is presented for the equations of motion of a test particle in flow field and in electromagnetic field.

[3] According to the hydrodynamic theory (e.g. Landau and Lifshitz [6]) when a solid particle moves through the fluid (at rest), the fluid energy induced by the relative particle motion of velocity $\boldsymbol{u} = (u_i)$ is expressed in the form $m_{ik}u_iu_k$

by using the *mass tensor* m_{ik}. Additional fluid momentum induced by the particle motion is given by $m_{ik}u_k$.

References

1. Kambe T (2008) Variational formulation of ideal fluid flows according to gauge principle. Fluid Dyn Res 40:399–426
2. Kambe T (2010a) Geometrical Theory of Dynamical Systems and Fluid Flows (Rev. ed), Chap. 7. World Scientific, Singapore.
3. Kambe T (2010b) A new formulation of equations of compressible fluids by analogy with Maxwell's equations. Fluid Dyn Res 42:055502
4. Landau LD, Lifshitz EM (1975) The classical theory of fields, 4th edn. Pergamon Press, Oxford
5. Kambe T (2010c) Vortex sound with special reference to vortex rings: theory, computer simulations, and experiments. Aeroacoustics 9(1–2):51–89
6. Landau LD, Lifshitz EM (1987) Fluid mechanics, 2nd edn. Pergamon Press, Oxford

Chapter 30
Dark Energy Condensate and Vacuum Energy

Houri Ziaeepour

Abstract Many candidate models for dark energy are based on the existence of a classical scalar field. In the context of Quantum Field Theory (QFT), we briefly discuss the condensation of such a field from a light quantum scalar field produced by gradual decay of a heavy particle during cosmological time. We obtain the necessary conditions for survival of the condensate in an expanding universe, and show that this process is directly related to quantum nature of the field which preserves the coherence of the condensate at cosmological distances. We also suggest a new interpretation of *vacuum energy* in QFT in curved spacetimes which can potentially solve the puzzle of huge deviation of what is considered to be the vacuum energy from observations of dark energy.

30.1 Introduction

Many alternatives to a cosmological constant have been proposed to explain the accelerating expansion of the Universe. They can be divided to two main groups: modified gravity models and models in which a field—usually a scalar but in some models a vector field—is responsible for what is called dark energy. This limited contribution does not allow to go to the details of each category, therefore we only concentrate on the models based on a scalar field, generally called *quintessence* [1, 2]. In simplest version of such models the quintessence field ϕ is a very light scalar with a self-interaction potential and no interaction with other components of the Universe. Under special conditions [3] the dynamics of the field at late times become independent of its initial value and the field approaches to what is called a tracking solution. In this case it varies very slowly with time and its equation of state—defined as:

H. Ziaeepour (✉)
Max Planck Institut für Extraterrestrische Physik (MPE), Giessenbachstraße 1,
85748 Garching, Germany

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_30,

$$w = \frac{P}{\rho} = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi}^2 + V(\phi)} \quad (30.1)$$

approaches $w \gtrsim -1$. Two types of potentials have tracking solutions: $V(\phi) = e^{-\alpha\phi}$, ϕ^{-n}, and polynomials including such terms. In the context of QFT these potentials are non-renormalizable, thus are assumed to be effective potentials.

A simple quintessence model suffers from various short comings. One of these problems is the very small mass of ϕ which must be $m_\phi \sim 10^{-32}$ eV $\sim H_0$. More importantly, this model cannot explain what is called *the coincidence problem*, i.e. why dark energy becomes dominant only after galaxy formation. This means that the density fraction of dark energy at the time of matter formation—presumably after inflation and during reheating—had to be $\sim 10^{42}$ times smaller than matter density. The only natural way to explain such an extreme fine tuning is to consider an interaction between dark energy and other components, notable with dark matter [4–6]. Moreover, a simple quintessence model is limited to $w > -1$, but for the time being many observations prefer $w \lesssim -1$, although due to measurement errors one cannot yet have a definitive conclusion about the sign of $w + 1$. Interacting dark energy models can explain $w < -1$ without violation of null energy principle because it has been shown [5, 7] that when the interaction is ignored, the effective equation of state $w_{eff} < -1$ when the real $w \geq -1$. As for the particle physics view, no particle is isolated and every species has some non-gravitational interaction with other particles.

The study of quintessence models is usually concentrated on the evolution of a classical scalar field without any concern about how such a field can be formed from a quantum field, specially in an expanding universe. In fact considering very small mass of a quintessence field and its very weak interaction to itself and to other particles, one expects that at their production—during inflation, reheating and/or later in the history of the Universe—they simply behave as relativistic particles and have an equation of state $w \sim 1/3$ which is very different from dark energy $w \approx -1$. Nonetheless, we also know that bosonic particles/fields can condensate and form a classical scalar field. We know this process from condense matter where Cooper pairs creates a non-zero expectation value—a condensate—at macroscopic scales, breaks the $U(1)$ symmetry, generates an effective mass for photons, and leads to phenomena such as superconductivity and super-fluidity. The Higgs field—if it exists—has a similar property but at microscopic scales. The formation of Higgs condensate at electroweak energy scale breaks $SU(2) \times U(1)$ and generates mass for leptons and quarks. The formation of a condensate has been studied, see e.g. [8] for some Higgs models as well as for inflaton [9]. These studies show that the problem of condensate formation and evolution is quite involved. In the case of dark energy it is even more complicated because one has to take into account the geometry of the expanding Universe and the evolution of other species, specially in the context of interacting quintessence models. The condensation issue is also more important because in contrast to Higgs and inflation, dark energy condensate must be very

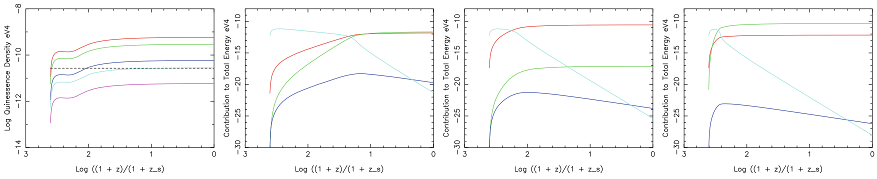


Fig. 30.1 From *left* to *right*: *1* Density of dark energy for various branching ration to the quintessence field $\Gamma_0 \equiv \Gamma_\phi/\Gamma = 10^{-16}$ (*magenta*), $5\Gamma_0$ (*cyan*), $10\Gamma_0$ (*blue*), $50\Gamma_0$ (*green*), $100\Gamma_0$ (*red*). *Dash line* is the observed value of the dark energy. $m_\phi = 10^{-6}$ eV, self-coupling $\lambda = 10^{-20}$. *2*, *3*, *4* Evolution of the contribution to the total energy density of ϕ for $\Gamma_0 = 10^{-16}$ and *2* $m_\phi = 10^{-8}$ eV and $\lambda = 10^{-20}$; *3* $m_\phi = 10^{-6}$ eV and $\lambda = 10^{-20}$; *4* $m_\phi = 10^{-6}$ eV and $\lambda = 10^{-10}$. Curves are: mass (*red*), self-interaction (*green*), kinetic energy (*cyan*) and interaction with DM (*blue*)

uniform and homogeneous both spatially and during cosmic time. These properties cannot be obtained trivially and should strongly constrain quintessence models.

In this proceeding we briefly review the technique, issues, and results obtained recently for a simple and generic interacting quintessence model [10]. We also describe an idea about a modified definition of vacuum energy which can solve the enormous deviation of the value obtained from usual definition in QFT.

30.2 Dark Energy from Decay of Dark Matter

In the context of inflation-reheating models, all constituents of the Universe were produced either during reheating from the decay of inflaton or a curvaton field, or later on from the decay of other species. In [6] we have studied the decay of a massive long life metastable dark matter with a small branching ratio to a light scalar field. A classical treatment of such a model show that the energy density of the light scalar field from very early times is roughly constant despite the expansion of the Universe, i.e. it behaves very similar to a cosmological constant, see Fig. 30.1. In contrast to many quintessence models, in this model the self-interaction potential is a simple ϕ^4 polynomial. The scalar field has a $w \sim -1$ for a large range of parameters and is not very sensitive to self-interaction potential because it is mainly the interaction/decay term that control its evolution during cosmic time. Indeed, such a setup has an internal feedback—if the density of dark energy increases, expansion rate increases, reduces the density of dark matter and thereby the rate of production of scalar field from decay of dark matter, thus the density of dark energy decreases. For a metastable dark matter, this stability can last for very long time. In addition, there would be no big ripe in the future because when a large fraction of dark matter decays, the stability of the system breaks, the energy density of dark energy decreases and the Universe becomes matter or radiation dominated, thus the accelerating expansion rate slow downs. Such a model can easily explain the observed slightly negative value of $w + 1$ [5, 7]. Therefore, for studying the condensation of quintessence field we consider this model.

30.3 Condensation of an Interacting Dark Energy

In condense matter a condensate is defined as a system in which the majority of particles are in their ground state. In quantum field theory when there is no conserved quantum number, such as in the case of a single scalar field without internal symmetry and with self interaction, a system is not usually in an eigen state of the number operator. Therefore a condensate which behaves classically is defined as a state in which number operator has a large expectation value—large occupation number—equivalent to a classical system with a large number of particles—in the minimum of their potential energy. Mathematically a condensate state $|\Psi\rangle$ is defined as:

$$\langle\Psi|\phi|\Psi\rangle \equiv \varphi \neq 0 \tag{30.2}$$

It is easy to verify that in contrast to Bose–Einstein condensate in quantum mechanics, for a quantum system containing only free or weakly interacting—perturbative—fields with finite number of particles φ is zero. It is possible to construct states which satisfies Eq. (30.2) using an expansion to coherent states. A special case is suggested by [11] and a more general state that we call *multi-condensate* is obtained in [10]. For a real scalar field it has the following expression:

$$|\Psi_{GC}\rangle \equiv \sum_k A_k e^{C_k a_k^\dagger}|0\rangle = \sum_k A_k \sum_{i=0}^{N\to\infty} \frac{C_k^i}{i!}(a_k^\dagger)^i|0\rangle \tag{30.3}$$

$$\begin{aligned}\chi(x,\eta) &\equiv a(\eta)\langle\Psi_{GC}|\Phi|\Psi_{GC}\rangle \\ &= \sum_k C_k\mathcal{U}_k(x) + C_k^*\mathcal{U}_k^*(x) \quad \Longrightarrow \quad C_k = \frac{\mathcal{U}_k(x) + \mathcal{U}_k^*(x)}{\chi(x)}\end{aligned} \tag{30.4}$$

where operators a_k and $a_k^\dagger$ are respectively annihilation and creation operators with $[a_k, a_k^\dagger] = 1$. Coefficients A_k and C_k^i are arbitrary, but can depend on the spacetime coordinate because creation and annihilation operators in a curved space depend on the coordinates. $\mathcal{U}_k$ is a solution of the free Green's function of ϕ.

To study the formation and evolution of such a state we consider a toy model for the Universe after reheating. Inspired by the classical model explained in the previous section, we assume a heavy particle—for the sake of simplicity a scalar X—presumably the dark matter, that decays to a light scalar field ϕ and some other particles that we collectively call A. In general ϕ can have a self-interaction considered here to be a simple power-law with positive exponent $V(\phi) = \phi^n,\ n > 0$. For X we consider only a mass term and no self-interaction. The collective field A

can have a self-interaction too. The field ϕ can be decomposed to $\phi = \Phi + I\varphi$ with $\langle\Psi|\Phi|\Psi\rangle = 0$, I the unit operator, and $\varphi(x)$ a classical scalar field (a $\mathcal{C}$-number). According to this definition $\langle\Psi|\phi|\Psi\rangle = \varphi(x)$. After inserting this decomposition to the Lagrangian, the dynamic equation of the classical field can be obtained from variation principle:

$$\frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu\varphi) + m_\Phi^2\varphi + \frac{\lambda}{n}\sum_{i=0}^{n-1}(i+1)\binom{n}{i+1}\varphi^i\langle\phi^{n-i-1}\rangle - \mathsf{g}\langle XA\rangle = 0 \tag{30.5}$$

A proper solution of this apparently simple equation needs a complete solution of Boltzmann or Kadanov-Baym equations—if we want to consider full nonequilibrium quantum field. This is a very complex problem specially in an expanding universe with a curved spacetime, and needs numerical solution of all the coupled equations. Therefore, rather than considering the full formulation, we assume that the evolution of other components affects φ only through the expansion rate of the Universe $a(t)$ in a FLRW cosmology with metric:

$$ds^2 = dt^2 - a^2(t)\delta_{ij}dx^i dx^j = a^2(\eta)(d\eta^2 - \delta_{ij}dx^i dx^j), \quad dt \equiv a d\eta \tag{30.6}$$

This approximation is applicable to radiation domination and matter domination epochs, but not to redshifts $z \lesssim 1$ when the dark energy becomes dominant. Although the decoupling of $a(t)$ simplifies the problem, we have yet to consider a state, or in classical limit a distribution, for other fields to determine the expectation values in Eq. (30.5). Fowllowinf cosmological observations, we assume a thermal distribution for other components. Evidently this is a valid assumption only at redshifts much smaller than reheating epoch, but due to the long lifetime of X particles only a negligible fraction of them decay earlier and their impact on the later state of matter should be small.

We solve Eq. (30.5) separately for radiation and matter domination epochs because they have very different evolution equations. To determine expectation values, we use Schwinger closed time path integral method, but we only consider tree level diagrams and only need to determine free propagators. Considering the very weak coupling of ϕ, this is a good approximation, up to the precision we need here. The Green's function of ϕ is coupled to the classical field φ, but considering the smallness of the self-coupling λ, we first determine the solution without taking into account the coupling term, then we use WKB approximation to obtain a more precise solution. Finally, we use α-vacuum as the initial condition for the Green's function.

During radiation domination epoch the Green's function equation without φ term has an exact solution. The evolution equation for φ also has the same form when interactions, including expectation values, are neglected. Their effect can be added through a WKB approximation. We consider an initial value $\varphi(t_0) = 0$ for the condensate. Finally, we obtain two independent solutions of the evolution equation which are plotted in Fig. 30.2. As this figure shows, the amplitude of the condensate has an exponential growth, similar to what happens during preheating and resonant

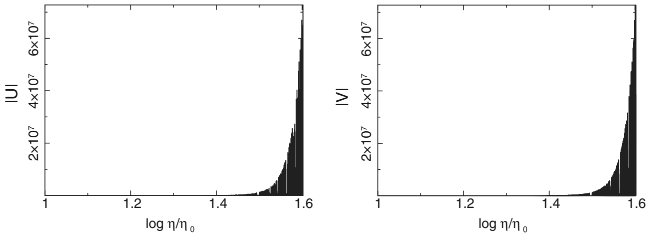


Fig. 30.2 An example of absolute value of independent solutions of evolution function of φ Note that although there are resonant jumps in the solution, due to the complexity of the interaction terms they are not regular like in preheating case

decay of inflaton to other fields. This is not a surprise because X and ϕ have a relation analogue to inflaton and matter fields, and have very similar evolution equations. Evidently, the exponential growth of the amplitude cannot continue forever, and backreactions due to nonlinearities in the evolution equation stop the growth rate. In particular, the interaction between the condensate component and *free* ϕ particles through nonlinear self-interaction terms in (30.5) has the tendency to free particles from condensate, in another word when the density of condensate grows, it begins to *evaporate*. Due to their tiny mass, free particles are relativistic and with the expansion of the Universe they become diluted very quickly. On the other hand, if a large number of them *evaporate*, their energy loss during the expansion of the Universe increases the probability of joining the condensate again. Therefore, as long as the expansion of the Universe is not very quick, this process is self-regulatory. It can be shown that the amplitude of modes decreases very rapidly with increasing $|k|$, i.e. for small distance scales. This is consistent with the lack of significant spatial fluctuation in dark energy density.

In the same way one can solve the Green's function and evolution equations during matter domination epoch. However, even when the interactions are ignored, these equations have a known analytical solution only if $m_\phi = 0$ or $k = 0$. Because we are specially interested in the modes with $|k| \to 0$ (Similar to radiation domination epoch, it is possible to show that the amplitude of modes for large $|k|$'s decreases quickly), we use the analytical solution for $k = 0$ as zero-order approximation, and apply WKB to obtain a better solution. Finally, we find a solution for the linearized evolution equation of φ which is proportional to $1/\eta$, thus decreases with time. This could be a disastrous for this model, because this leads to a dark energy with $w > -2/3$ which is already ruled out by observations. Nonetheless, when the full nonlinear equation is considered, although we cannot solve it, there is evidence that under special conditions a roughly constant amplitude—a tracking solution—similar to observed dark energy can be obtained. In fact, expectation values of type $\langle\phi^i\rangle$ induce negative power of φ into the evolution equation of φ because C_k is proportional to φ^{-1}, see Eq. (30.4). As we mentioned in the Introduction, polynomials with negative power are proved to have a tracking solution. A counting of power of self-interaction terms after replacing expectation values with their approximate solution show that for $n \leq 3$ there exits a tracking solution. For $n = 4$ the decrease rate of the condensate density can be enough slow to be consistent with present observations. In 4D spacetimes these potentials are the only renormalizable self-interactions for a quantum scalar field model. Giving the fact that we did not impose any constraint on renormalizability of the model, these results are very interesting and encourage more work on quantum description and origin of dark energy.

If these conclusions are confirmed by a more precise numerical calculations, they would be a proof of the reign of quantum mechanics at largest scales in the Universe because it is the quantum coherence of dark energy that saves it from being diluted by the expansion. The dominance of dark energy at late times in one hand proves that in contrast of general believes, the Universe is dominantly in a coherent quantum state, and in the other hand dark energy provides a natural envionment for decoherence of other constituents.

30.4 Vacuum Energy

In QFT energy is calculated as the expectation value of classical expression for the energy momentum tensor $T^{\mu\nu}$ in which classical field is replaced by its quantized counterpart:

$$
\begin{aligned}
E =& \langle\psi| \int d^4k \delta(k_0^2 - m^2) T^{00} |\psi\rangle = \frac{1}{2} \langle\psi| \int d^3\omega_k (a_k^\dagger a_k + a_k a_k^\dagger) |\psi\rangle \\
=& \langle\psi| \int d^3\omega_k (\hat{N}_k + \frac{1}{2}) |\psi\rangle, \quad \omega_k \equiv \mathbf{k}^2 + m^2 \qquad (30.7)
\end{aligned}
$$

Vacuum energy is defined as $|\psi\rangle = |0\rangle$, thus $E_{vac} = \frac{1}{2} \int d^3\omega_k \to \infty$. To regularize this integral usually a UV cutoff is imposed that leads to a finite but very large value for the energy density of vacuum. In QFT in Minkovsky space without gravity ordering operator is imposed to the above definition i.e. $\langle\psi| : T^{00} : |\psi\rangle$ is used. This simple operation removes the constant (infinite) term and makes a strictly zero vacuum energy. When gravity is present, it is usually supposed that ordering operator cannot be applied because it shifts the energy, an unauthorized operation in the context of general relativity and gravity that define an absolute reference for energy.

The constant term in $\langle\psi|T^{00}|\psi\rangle$ is due to noncommutative creation and annihilation operators in field theory. The presence of the constant term looks like the memory of spacetime or the ghost of a particle, i.e. when a particle is created, its annihilation does not completely restore the initial state. This can be interpreted as a manifestation of energy conservation. In fact, considering two operators $a^\dagger a$ and $aa^\dagger$, their physical interpretations are very different. The former counts the number of particles in a state. For an asymptotically free system, it behaves as a detector of particles without changing their state. By contrast, operator $aa^\dagger$ first creates a particle i.e. changes the energy of the system by an amount equal to the energy of the particle. Because in general relativity energy and momentum are locally conserved, this operatation necessarily violates the closeness of the system and must actually play the role of a bridge between the system—state—under consideration and another system that provides the energy. Moreover, energy and momentum are eigen values of translation operator. Therefore, the application of right part of this operator changes the translation (symmetry) state of the system. Because symmetry is related to information, the return of energy to its initial reservoir i.e. the annihilation of the particle restore the energy state, but quantum mechanics tells us that it does not restore the information. This is another manifestation of nonlocality or state collapse in quantum mechanics. Nonetheless, if we are only interested in energy conservation, annihilation restores energy-momentum state and in this regard the system should be considered as unchanged. Base on this argument we suggest that operator ordering must be applied to $\langle\psi|T^{00}|\psi\rangle$ even in the context of general relativity.

References

1. Ratra B, Peebles PJE (1988) Phys Rev D 37:3407
2. Wetterich C (1988) Nucl Phys B 302:668
3. Steinhardt PJ, Wang L, Zlatev I (1999) Phys Rev D 59:123504 [astro-ph/9812313]
4. Amendola L (2000) Phys Rev D 62:043511 [astro-ph/9908023]
5. Ziaeepour H (2000) Cosmic equation of state, quintessence and decaying dark matter [astro-ph/0002400].
6. Ziaeepour H (2004) Phys Rev D 69:063512 [astro-ph/0308515]
7. Das S, Corasaniti PS, Khoury J (2006) Phys Rev D 73:083509 [astro-ph/0510628]
8. Inagaki T (1994) Nucl Phys Proc Suppl 37A:197 [hep-ph/9402235]
9. Parker L, Zhang Y (2000) Phys Rev D 47:416
10. Ziaeepour H (2010) Phys Rev D 81:103526 [arXiv:1003.2996]
11. Matsumoto S, Moroi T (2008) Phys Rev D 77:045014 [arXiv:0709.4338]
12. Ziaeepour H (2005) Progress in dark matter research. Nova Science Inc, New York, p. 175 [astro-ph/0406079].

Chapter 31
General Relativistic Quantum Theories: Foundations, the Leptons Masses

Claudio Parmeggiani

Abstract The space-time is represented by a usual, four-dimensional differential manifold, X. Then it is assumed that at every point x of X, we have a Hilbert space H(x) and a quantum description (states, observables, probabilities, expectations) based on H(x). The Riemannian structure of X induces a connection on the fiber bundle H and this assumption has many relevant consequences: the theory is regularized; the interaction energy is a well defined self-adjoint operator; finally, applying the theory to electro-weak interactions, we can obtain a "specter" of leptons masses (electron, muon, tau).

31.1 Introduction: Summary

We shall discuss (the foundations of) a General Relativistic Quantum Theory: now the Hilbert space of the quantum description is replaced by a *fiber bundle* (here called *quantum bundle*) based on a four-dimensional differential manifold (the space-time); the typical fiber is a complex, infinite-dimensional Hilbert space; on every fiber there are states (vectors or rays), observables (operators), probabilities and expectations (Sect. 31.1).

The ordinary derivative (in Schrödinger equations) is replaced by a sort of *covariant derivative* on the quantum bundle. This derivative, at a space-time point, depends on the value of the space-time metric tensor (the gravitational field) at the same point: this is the *main hypothesis*. The standard quantum field theories are formally recovered, for a "constant" metric tensor (that is when the gravitational field goes to zero); but, in fact, even for an isolated elementary particle we have to consider the gravitational field of the particle itself (Sect. 31.2).

C. Parmeggiani (✉)
University of Milan, Milan, Italy
e-mail: clparm@infinito.it

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_31,

The quantum fields and the energy operator (the "Hamiltonian") are now well defined (eventually self-adjoint) operators; the Pauli-Jordan distributions are now true, differentiable functions (gravitation dependent, obviously); therefore the Theory is regularized: there are no more *divergences* (Sect. 31.3).

The presence of our "covariant" derivative not only regularizes the theory but, as a consequence, the (ratios of the) masses of the leptons are, by some supplementary assumptions, theoretically predictable. In fact, imposing the constancy of the sum of the space-time dependent *self mass* (logarithmically diverging, with zero gravity) and of a space-time dependent *counter term*, we arrive at a second order partial differential equation; the self values of this equation are (proportional to) the leptons masses (Sects. 31.4, 31.5).

Finally note that we have only to postulate the existence of three stable elementary leptons (electrons, positrons and electron neutrinos); the other ones (the μ and the τ particles, their neutrinos) and their masses are derived, so to speak, as excitations.

31.2 Fiber Bundles and Quantum Bundles

A *fiber bundle* (see, for example, [2] or [3]) is a tuple (F, X, π, $\mathbf{F}$) where F and X (the *base space*) are manifolds; π is a continuous surjection from F to X; the F(x) := $\pi^{-1}(x)$ (the fibers at x, a point of X) are all isomorphic to the typical fiber $\mathbf{F}$. A (cross) *section* is a function Ψ from X to F such that:

$$\pi(\Psi(x)) = x \tag{31.1}$$

Locally (in an open subset U of X) F(x) can be identified to $\mathbf{F}$; hence, locally, the sections can be expressed as (ordinary) maps from U to $\mathbf{F}$; in the following we shall generally employ *local expressions.*

A *quantum bundle* is defined as a tuple (H, X, π, $\mathbf{H}$) where the base space X is a four-dimensional Riemannian manifold ($\mathbf{g}$ is its (+*1*, –*1*, –*1*, –*1*) metric tensor); the fibers at x, H(x), and the typical fiber, $\mathbf{H}$, are infinite-dimensional, complex, isomorphic Hilbert spaces. The Schrödinger function is now a section of the quantum bundle, locally expressed by a map from X to $\mathbf{H}$; the observables are expressed by operator-valued maps, assigned on X. For example, the average value of the observable ξ, the "system" being in the state Ψ, can be (locally) written as:

$$(\Psi(x)|\xi(x)|\Psi(x))_{\mathrm{H}}/(\Psi(x)|\Psi(x))_{\mathrm{H}} \tag{31.2}$$

where the inner product is taken on $\mathbf{H}$. Obviously in a Schrödinger picture the section ξ will be "constant" (in a sense make precise in Sect. 2.2).

31.3 Derivatives Dynamics

To relate quantum objects defined at different space-time points we need a way to connect them, at least when they are "infinitely near"; thence we shall introduce a notion of "covariant" derivative.

31.3.1 Covariant Derivatives

If Ψ is the local expression of a *section* of the Quantum Bundle and u is a 4-dimensional *vector* (space-like or time-like), the *covariant derivative* of Ψ, in the direction u, is defined as:

$$\mathsf{D}_u\Psi(x) := \partial_u\Psi(x) + \mathsf{C}_u(x).\Psi(x) \tag{31.3}$$

where ∂_u is the ordinary derivative (along u) and $\mathsf{C}_u(x)$ is a space-time dependent operator on the Hilbert space $\mathbf{H}$. And for an observable (an operator) ξ:

$$\mathsf{D}_u\xi(x) = \partial_u\xi(x) + [\mathsf{C}_u(x), \xi(x)] \tag{31.4}$$

If γ is a smooth path on X, parameterized by the real variable s, a section Ψ along γ is said *parallel* if:

$$\mathsf{D}_u\Psi(\gamma(s)) = 0(u = \gamma'(s)) \tag{31.5}$$

31.3.2 Schrödinger Equation

The self-adjoint operator on $\mathbf{H}$ (u is again a 4-dimensional vector)

$$\mathsf{P}_u(x) = \mathsf{P}_0(x)u^0 + \mathsf{P}_1(x)u^1 + \mathsf{P}_2(x)u^2 + \mathsf{P}_3(x)u^3 \tag{31.6}$$

is the (local expression of the) e*nergy-momentum* observable, at space-time point x; it is an energy operator if u is time-like, a momentum operator when u is space-like. Typically $\mathsf{P}_u(x)$ can be decomposed as:

$$\mathsf{P}_u(x) = \mathsf{P}_u^{(\mathrm{FREE})}(x) + \mathsf{P}_u^{(\mathrm{INT})}(x) \tag{31.7}$$

where $\mathsf{P}_u^{(\mathrm{INT})}(x) = 0$, for a space-like u.

A section Ψ satisfies a *Schrödinger equation* (in a Schrödinger picture) if

$$\mathsf{D}_u\Psi(x) = \partial_u\Psi(x) + \mathsf{C}_u(x).\Psi(x) = -iP_u(x).\Psi(x) \tag{31.8a}$$

while, for *all* observables ξ:

$$\mathsf{D}_u\xi(x) = \partial_u\xi(x) + [\mathsf{C}_u(x), \xi(x)] = 0 \tag{31.8b}$$

Observe that, because $\mathsf{P}_u(x)$ and $\mathsf{P}_v(x)$ commute and $\mathsf{D}_u\mathsf{P}_v(x) = 0$ (for every couple of vectors, u and v), we have $\mathsf{D}_u\mathsf{D}_v\Psi(x) = \mathsf{D}_v\mathsf{D}_u\Psi(x)$.

In an *interaction picture*

$$\mathsf{D}_u\Psi^{(\mathrm{INT})}(x) = -i\mathsf{P}_u^{(\mathrm{INT})}(x).\Psi^{(\mathrm{INT})}(x) \tag{31.9a}$$

$$\mathsf{D}_u\xi^{(\mathrm{INT})}(x) = i[\mathsf{P}_u^{(\mathrm{FREE})}(x), \xi^{(\mathrm{INT})}(x)] \tag{31.9b}$$

In the following we shall always use the interaction picture.

31.3.3 Explicit Expression of the Covariant Derivative

Having defined a *vacuum state* (at space-time point x), $\Phi_0(x)$, others states (and observables) can be generated by means of the *creation* and *annihilation* operators, $\mathbf{a}^*_{n,p}(x)$ and $\mathbf{a}_{n,p}(x)$ (see, for example, [1] or [5]). Here n is an index = *1, 2, …, N* and

$$p = (\varepsilon, \mathbf{p}), \qquad \varepsilon = \sqrt{(\mathbf{p}^2 + M^2)} \tag{31.10}$$

ε, $\mathbf{p}$ and M are the particles energy, momentum and (bare, unobservable) mass. Evidently we must have, for the vacuum, $\mathsf{D}_u\Phi_0(x) = \partial_u\Phi_0(x) + \mathsf{C}_u(x).\Phi_0(x) = 0$ (and also $\partial_u\Phi_0(x) = 0$).

In our interaction picture

$$[P_u^{(\mathrm{FREE})}(x), \mathbf{a}_{n,p}(x)] = -(p.u)\mathbf{a}_{n,p}(x) \tag{31.11}$$

so: $\mathsf{D}_u\mathbf{a}_{n,p}(x) = -i\,(p.u)\,\mathbf{a}_{n,p}(x)$. Afterward we shall assume (this is a *fundamental postulate*) that:

$$[\mathsf{C}_u(x), \mathbf{a}_{n,p}(x)] = i\kappa c_u(x, p)\mathbf{a}_{n,p}(x) \tag{31.12}$$

where c_u is an *ordinary numerical function* dependent on x and p (but *not* on n) and, at least in a first approximation, that:

$$c_u(x, p) = -\sum_{\alpha\beta} \Gamma_u^{\alpha\beta}(x).p_\alpha p_\beta = 1/2\sum_{\alpha\beta} \partial_u g^{\alpha\beta}(x).p_\alpha p_\beta (13) \tag{31.13}$$

Here the $g^{\alpha\beta}$ are the "contravariant" components of the metric tensor $\mathbf{g}$; the Γ are its Christoffel symbols; κ is a numerical, universal parameter. We are using absolute unit of measure, that is $h_{PLANK} = 2\pi$ and $M_{PLANK} = 1$.

Hence we are lead to the differential equation ($c(\boldsymbol{x},p) := 1/2 \sum_{\alpha\beta} g^{\alpha\beta}(\boldsymbol{x}).p_\alpha p_\beta$)

$$i\partial_u \mathbf{a}_{n,p}(\boldsymbol{x}) = ((p.u) + \kappa\partial_u c(\boldsymbol{x}, p))\mathbf{a}_{n,p}(\boldsymbol{x}) \tag{31.14}$$

which can be immediately integrated:

$$\mathbf{a}_{n,p}(\boldsymbol{x}) = \exp(-ip(\boldsymbol{x} - \boldsymbol{x}_0)).\exp(-i\kappa(c(\boldsymbol{x}, p) - c(\boldsymbol{x}_0, p))).\mathbf{a}_{n,p}(\boldsymbol{x}_0) \tag{31.15}$$

$\boldsymbol{x}_0$ is a *fixed* space-time point. Evidently for a constant, "special relativistic" $\mathbf{g}$, we shall recover the usual definitions, relatively to the free creation and annihilation operators (see again [1] or [5]).

For a "near constant" $\mathbf{g}$ (that is in a weak gravitational field approximation, see [4])

$$c(\boldsymbol{x}, p) = 1/2M^2 + (M^{(\mathrm{GRAV})}/r(\boldsymbol{x}))(\varepsilon^2 + \mathbf{p}^2) \tag{31.16}$$

$M = \sqrt{}(\varepsilon^2 - \mathbf{p}^2)$ is the bare mass; $M^{(\mathrm{GRAV})}$ is the gravitating mass; $r(\boldsymbol{x})$ is the distance from $\boldsymbol{x}$ to the source of the gravitational field.

31.3.4 Commutations and Anti-commutations Relations

At the same space-time point $\boldsymbol{x}_0$

$$[\mathbf{a}_{m,p}(\boldsymbol{x}_0), \mathbf{a}_{n,q} * (\boldsymbol{x}_0)]_\pm = N_\mathrm{p}\delta_{m,n}\delta_{\mathrm{p,q}} \tag{31.17}$$

hence, for a *real* κ (positive or negative):

$$[\mathbf{a}_{m,p}(\boldsymbol{x}), \mathbf{a}_{n,q} * (y)]_\pm = N_\mathrm{p} \exp(ip(y - \boldsymbol{x})).\exp(i\kappa(c(y, p) - c(\boldsymbol{x}, p)))\delta_{m,n}\delta_{\mathrm{p,q}} \tag{31.18}$$

where $\boldsymbol{x}$ and y are different space-time points. The bracket $[\ ,\ ]_\pm$ refer to Fermi or Bose statistics; N_p is a normalization factor, usually set equal to *1* (as in [5]) or to $\sqrt{}(\mathbf{p}^2 + M^2)/M$.

Observe that, if κ is *imaginary*, $\kappa = -i\kappa'$, κ' positive, we get a term like $\exp(-\kappa'(c(\boldsymbol{x},p) + c(y,p)))$.

31.4 Quantum Fields

It is possible to define neutral and charged *quantum fields* ϕ and ψ (now well defined, operator-valued functions assigned on X) as:

$$\phi(x) = \int (\mathbf{a}_p(x) + \mathbf{a}_p * (x)) d\mu(\mathbf{p}) \tag{31.19}$$

$$\psi(x) = \int \mathbf{a}_p(x) d\mu(\mathbf{p}) + \int \mathbf{b}_p * (x) dv(\mathbf{p}) \tag{31.20}$$

$d\mu$ and dv are measure on $\mathbf{R}^3$, eventually matrix-valued. Obviously, for $\kappa = 0$, we shall obtain the usual, singular quantum fields. These are *free fields* (the $\mathsf{P}_u^{(\mathrm{INT})}(x)$ term is not considered), but the gravitational effects are fully included in the field operators.

Afterward we can build the Pauli-Jordan functions $\Delta(x, y)$: they depend on $\kappa(c(y,p)–c(x,p))$ and are, generally, not translation invariant. But they are now true, differentiable functions (assigned on X^2) and, for $\kappa = 0$, we shall obtain the usual distributions (singular functions), translation invariant. We have, for example, for a real κ:

$$\Delta(x, y) = [\phi(x), \phi(y)] = 2\int \sin(p(y - x) + \kappa(c(y, p) - c(x, p))) d\mu(\mathbf{p}) \tag{31.21}$$

but for an imaginary $\kappa = -i\kappa'$, κ' positive:

$$\Delta(x, y) = [\phi(x), \phi(y)] = 2\int \sin(p(y - x)).\exp(-\kappa'(c(y, p) + c(x, p))) d\mu(\mathbf{p}) \tag{31.22}$$

that leads to an exact, canonical *equal time* commutation relation.

31.5 Interactions Self-Energies

31.5.1 Interaction-Energy Operator

Let us consider a tri-linear *interaction-energy operator* (the Hamiltonian, always in an interaction picture):

$$H^{(\mathrm{INT})}(x) = \lambda \int\int (\xi_p * (x) \mathbf{A}_{p,q}(x) \xi_q(x)) d\mu(\mathbf{p}, \mathbf{q}) \tag{31.23}$$

here $d\mu$ is a measure on $\mathbf{R}^6$; $\mathbf{A}_{p,q}(x)$ is a self-adjoint operator (Bose statistics) related to the gauge fields which mediate the interaction; the operators $\boldsymbol{\xi}_q(x)$ describe the interacting particles (Fermi statistics) and they are linear combinations of the $\mathbf{a}_p(x)$, $\mathbf{a}_p{*}(x)$, $\mathbf{b}_p(x)$, $\mathbf{b}_p{*}(x)$; λ is a coupling constant. Note that we are only considering the particles interactions and not the interactions between the gauge fields.

When $\kappa = 0$ the *interaction-energy operator* (formally) reduces to the standard expression ($x = (t, \mathbf{r})$):

$$H^{(\mathrm{INT})}(t) = \lambda \int \int (\xi_p * (t) \mathbf{A}_{p,q}(t) \xi_q(t)) d\mu(\mathbf{p}, \mathbf{q}) \tag{31.24}$$

time dependent and needing a regularization.

31.5.2 Particles Self-Masses

If we try to calculate the self-energies (or the *self-masses*) of our particles considering only the one-loop Feynman graphs, we arrive at a $\mathbf{r}$ (and $\mathbf{g}$) dependent expression ([6], see also for a more “modern”, but essentially equivalent, presentation [1] or [5]):

$$M^{(\mathrm{SELF})}(\mathbf{r}) = -M.\Lambda.\log(\mathrm{F}_g(\mathbf{r})) + const. + \cdots \tag{31.25}$$

logarithmic diverging when $\kappa = 0$ (that is ignoring gravitation). Here M is again the bare mass and Λ a numerical, positive parameter, proportional to the squared coupling constant λ. In the weak gravitational field approximation (Sect. 2.3):

$$F_g(\mathbf{r}) = M^{(\mathrm{GRAV})}/r \tag{31.26}$$

where $M^{(\mathrm{GRAV})}$ is the *effective* (inertial, gravitational) particle mass.

31.5.3 Counter-Terms

Now we shall assume that the interaction-energy operator contains a mass-like counter-term

$$\int (\boldsymbol{\xi}_p * (x) M^{(\mathrm{CTERM})}(x) \boldsymbol{\xi}_p(x)) d\mu(\mathbf{p}) \tag{31.27}$$

and that the real function $\mathrm{M}^{(\mathrm{CTERM})}(\mathbf{r})$ (> 0) is proportional to

$$-M.(\Delta k(\mathbf{r})/k(\mathbf{r})) \tag{31.28}$$

where M is the bare mass, Δ the Laplace, second order, differential operator and k is a scalar, classical field. Alternatively we can include the counter-term in the free part of the energy operator so the "bare mass" will be a function of $\mathbf{x}$. Presumably the counter-term is of *geometrical origin*, for example it is the Ricci scalar of a (modified) metric tensor **g**. In any case, imposing the constancy of the sum

$$\mathrm{M}^{(\mathrm{SELF})}(\mathbf{r}) + \mathrm{M}^{(\mathrm{CTERM})}(\mathbf{r}) \tag{31.29}$$

(the effective mass cannot depends on $\mathbf{r}$), we arrive at a $\mathbf{R}^3$ differential equation (note that here the bare mass do not appears):

$$\Delta k(\mathbf{r})/k(\mathbf{r}) = -\Lambda' \log(\mathrm{F_g}(\mathbf{r})) + const. \tag{31.30}$$

For a radial symmetric k, rescaling the dependent variable and using the weak field expression of $\mathrm{F}_g(\mathbf{r})$, $M^{(\mathrm{GRAV})}$ / r,

$$K''(r) + (2/r)K'(r) = (log(r) - w)K(r) \tag{31.31}$$

We are looking for a $K(r)$ going to zero when r diverge, hence we are lead to a countable family of solutions of the above differential equation; the first self-values of w are:

$$1.044;\ 1.847;\ 2.290;\ 2.596;\ 2.830;\ 3.020;\ 3.179 \tag{31.32}$$

Obviously we can only obtain in this way the *masses ratios* of the first, the second, the third, ... *variety* (or generations) of particles.

31.6 Leptons Masses

Consider now the case of the electro-weak interactions between *leptons*, mediate by the gauge fields *A, Z, W, W**; M_Z and M_W are the *Z*-particle and *W*-particle masses. To calculate the self-masses we need the three (one loop) Feynman graph (*lept* stands for a *charged* lepton):

$$\begin{aligned} lept(lept - photon)lept + lept(lept - Zparticle)lept \\ + lept(neutrino - Wparticle)lept \end{aligned} \tag{31.33}$$

then in the differential equation of Sect. 31.5.3 appears a sum of three mass-dependent terms:

$$\Lambda_A.\log(M_{lept})+\Lambda_Z.\log(M_{lept}+M_Z)+\Lambda_W.\log(M_{neutrino}+M_W) \tag{31.34}$$

The Λ_A, Λ_Z and Λ_W are three positive parameters proportional to the squared electro-weak coupling constants; $\Lambda_A + \Lambda_Z + \Lambda_W = 1$, as a consequence of the rescaling the dependent variable (Sect. 31.5.3). Apparently the experimental data lead to $\Lambda_A/\Lambda_W = \mathit{0.231}$ and to $\Lambda_A/\Lambda_Z = \mathit{0.769}$ so:

$$\Lambda_A = \mathit{0.151},\ \Lambda_Z = \mathit{0.196}, \Lambda_W = \mathit{0.653} \tag{31.35}$$

Finally, assuming that $M_{neutrino} << M_W$, we arrive at:

$$\log(M_\mu/M_{electron}) = 5.32, \log(M_\tau/M_{electron}) = 8.22 \tag{31.36}$$

in reasonable accord to the experimental values: *5.332* and *8.154* (we have introduced many approximations). But if, for example, we add to the self-masses a $(\log(r))^2$ term (coming from some two loops Feynman graphs), we obtain a quite better accord. It would be also of some interest to consider the not radial symmetric solutions of the k-field differential equation.

Obviously we also obtain other, bigger mass values (corresponding, presumably, to highly instable particles):

$$M_4 \approx 11\ Gev, M_5 \approx 34\ Gev, M_6 \approx 72\ Gev, \ldots \tag{31.37}$$

If we try to calculate the neutrinos masses starting from the interaction-energy operator defined at the Sect. 37.1, we arrive at a quite problematic result: the neutrinos predicted masses are very large, greater than the W-particle and Z-particle masses. To cure this problem it is apparently necessary to modify the structure of the energy operator.

References

1. Bogoliubov NN, Shirkov DV (1980) Introduction to the theory of quantized fields. Wiley, New York.
2. Bourbaki N (2007) Eléments de Mathématique. Fascicule de résultats, Springer-Verlag, Variétés différentielles et analytiques
3. Lang S (1985) Differential manifolds. Springer-Verlag, Berlin
4. Weinberg S (1972) Gravitation and cosmology: principles and applications of the general theory of relativity. Wiley, New York
5. Weinberg S (1995) The quantum theory of fields. Volume I, Cambridge UP.
6. Weisskopf VF (1939) On the self-energy and the electromagnetic field of the electron. Phys Rev 56:72

Chapter 32
Direction of Time from the Violation of Time Reversal Invariance

Joan A. Vaccaro

Abstract We show that the violation of time reversal invariance (T) observed in meson decay may have a profound effect on the direction of time. Our starting point is a universe without a presumed direction of time. Time evolution is modeled by taking an equal superposition of steps in both directions of time. This gives rise to multiple paths through time. The presence of T violating processes is shown to give rise to destructive interference between all paths except for the two that comprise continuously forwards and continuously backwards steps. We show that this leads to a new kind of irreversibility that is quite unlike that found in thermodynamics.

32.1 Introduction

The direction of time is typically studied in terms of various arrows of time [5]. Each arrow is associated with some time-asymmetric phenomenon and is oriented to point to the future. Perhaps the best known is the thermodynamic arrow which points in the direction of an increase in the entropy of an isolated system [2]. Essentially, an increase in entropy is the *effect* that results from a fixed direction of time and an incomplete knowledge of a system. The thermodynamic arrow requires the universe to be initially in a low-entropy non-equilibrium state and evolve towards equilibrium. This implies an asymmetry between the past and future boundary conditions. The time asymmetry of other arrows of time can also be attributed to asymmetric temporal boundary conditions. For example, the cosmological arrow points in the direction of the expansion of the universe and so relies on a small universe in the past and a large universe in the future. All such arrows are based on time-symmetric dynamical laws. The one notable exception is the matter-antimatter arrow of time (also called the weak arrow of time) which partially explains the dominance of matter over antimatter in the universe. It arises from the observed violation of charge conjugation and parity

J. A. Vaccaro (✉)
Centre for Quantum Dynamics, Griffith University, Brisbane, Australia

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_32,

inversion invariance (CP) in meson decay due to the weak interaction [1]. Under the CPT theorem, CP violation is equivalent to violation of time reversal invariance (T). This means that meson decay is described by a time-asymmetric dynamical law and so the matter-antimatter arrow originates from the time asymmetry of a dynamical law rather than boundary conditions. Of all the arrows of time, it is the only one to have the potential of being the *cause* of the direction of time, rather than being simply an *effect*. Indeed T violation in meson decay has recently been shown to have potentially large scale physical effects in a model of the universe in which the direction of time is not predefined [6]. Here we re-examine the implications of the model for the direction of time.

In Sect. 32.2 we briefly review the model that was introduced in [6]. Next in Sect. 32.3 we show how a new kind of irreversibility emerges from the model when sufficient T violation processes are present. In Sect. 32.4 we explore the repercussions of baryogenesis and end with a discussion of the implications for the direction of time in Sect. 32.5.

32.2 Effect of T Violation in a Universe Without a Presumed Temporal Direction

We want to explore how T violation can give rise to a preferred direction of time. For this we need to model the universe in such a way that it has no predefined direction of time if T violation processes are absent. We have two choices: the universe does not evolve at all, in which case there is nothing more to be said, or both directions of time have an equal footing. To incorporate the second option, rather than take time steps in a particular direction, time evolution must somehow be in both directions of time. In other words we must allow for the possibility of multiple paths through time. Feynman's sum over paths [3] provides the mathematical framework for dealing with such situations. We let the evolution in one direction of time, which we shall call "forwards" and associate with the positive time direction, be described by

$$|\psi_F\rangle = U_F(\tau)\,|\psi_0\rangle\,, \tag{32.1}$$

and in the other direction of time, which we shall call "backwards" and associate with the negative time direction, by

$$|\psi_B\rangle = U_B(\tau)\,|\psi_0\rangle\,, \tag{32.2}$$

where τ represents a time interval and

$$\begin{aligned} U_F(\tau) &= \exp(-iH_F\tau) \\ U_B(\tau) &= TU_F(\tau)T^{-1} = \exp(iTH_FT^{-1}\tau) = \exp(iH_B\tau). \end{aligned}$$

Here, H_F and $H_B \equiv T H_F T^{-1}$ are the Hamiltonians for the universe with respect to the forward and backward directions of time, respectively, T is Wigner's time reversal operator [7] and we use units where $\hbar = 1$. The Hamiltonian violates time reversal invariance if

$$H_B \neq H_F.$$

The expressions $\langle\phi| U_F(\tau) |\psi_0\rangle$ and $\langle\phi| U_B(\tau) |\psi_0\rangle$ represent the probability amplitudes for the universe to evolve from the state $|\psi_0\rangle$ to the state $|\phi\rangle$ via two different paths, each of which corresponds to a different direction of time. We have no reason to favor one path over the other so, according to Feynman's sum over paths method, we take the total probability amplitude to evolve from $|\psi_0\rangle$ to $|\phi\rangle$ as the sum $\langle\phi| U_F(\tau) |\psi_0\rangle + \langle\phi| U_B(\tau) |\psi_0\rangle = \langle\phi| U_F(\tau) + U_B(\tau) |\psi_0\rangle$. As this result is true for all states $|\phi\rangle$ we can write the unbiased evolution of the universe as

$$|\Psi(\tau)\rangle = [U_F(\tau) + U_B(\tau)] |\psi_0\rangle . \tag{32.3}$$

We call the process described by Eq. (32.3) *symmetric time evolution*. This analysis can be repeated for an additional step of symmetric time evolution. After N such steps the state of the universe is given by

$$|\Psi(N\tau)\rangle = [U_F(\tau) + U_B(\tau)]^N |\psi_0\rangle . \tag{32.4}$$

To simplify the description we have ignored the normalization of the state. Eq. (32.4) is represented as a binary tree in Fig. 32.1a.

Expanding the N-fold product on the right side of Eq. (32.4) gives a series of 2^N terms each of which represents a different path through time. For example the terms $U_F^N(\tau) |\psi_0\rangle$ and $U_F^{N-1}(\tau)U_B(\tau) |\psi_0\rangle$ represent paths of N steps containing zero and one backward step, respectively. The series can be separated into subseries whose terms involve the same number of steps in the forward and backward directions, but with steps in different orders. For example, the small circle in Fig. 32.1a represents a subseries characterized by two forward steps and one backward step:

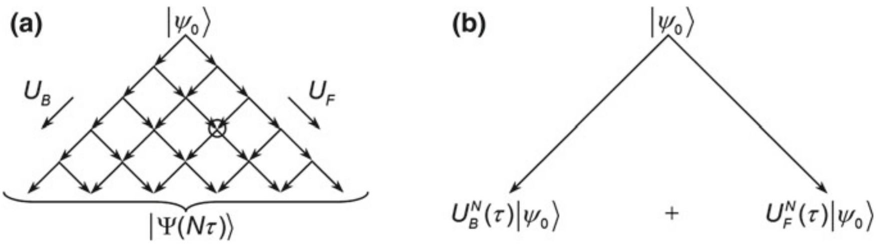


Fig. 32.1 **a** and **b** are representations of Eqs. (32.4) and (32.5), respectively, as binary trees. *Leftward arrows* (↙) represent the application of U_B and *rightward arrows* (↘) represent the application of U_F

$$U_F(\tau)U_F(\tau)U_B(\tau)\,|\psi_0\rangle + U_F(\tau)U_B(\tau)U_F(\tau)\,|\psi_0\rangle + U_B(\tau)U_F(\tau)U_F(\tau)\,|\psi_0\rangle\,.$$

Note that the operators $U_F(\tau)$ and $U_B(\tau)$ do not commute if the Hamiltonian is not T invariant. In that case the terms of the subseries have been shown to interfere destructively [6].

In order to model the interference effect we use the conventional phenomenological model of neutral kaon decay [4, 8] as the prototypical T violation process. We quantify the number of kaons as $f \times 10^{80}$ where f is the fraction of the estimated 10^{80} total number of particles in the universe and set the size of the time step to the Planck time, i.e. $\tau \approx 10^{-44}$ s. With these conditions, a relatively lengthy calculation [6] reveals that for $N\tau \gg f^{-1/2}10^{-13}$ s Eq. (32.4) becomes

$$|\Psi(N\tau)\rangle \approx [U_F^N(\tau) + U_B^N(\tau)]\,|\psi_0\rangle\,. \qquad (32.5)$$

The right side comprises only two parts, one representing consistent evolution in the forward direction, $U_F^N(\tau)\,|\psi_0\rangle$, and the other representing consistent evolution in the backward direction, $U_B^N(\tau)\,|\psi_0\rangle$, as illustrated in Fig. 32.1b. Destructive interference has eliminated all other paths. This should be contrasted with the situation in Eq. (32.4) for which the direction of the evolution can change from step to step as illustrated in Fig. 32.1a.

32.3 New Kind of Irreversibility

We now turn to new results. The derivation of Eq. (32.4) relies on the construction of paths through time and it is this feature we wish to explore in more detail here. If the direction of time is known to be the forwards direction, say, then we can determine whether a universe that is in state $|A\rangle$ will be found in a different state $|B\rangle$ after a time interval t from the probability

$$|\langle B|\,U_F(t)\,|A\rangle|^2\,. \qquad (32.6)$$

A nonzero probability value means that there is a *temporal path* from $|A\rangle$ to $|B\rangle$. Expression (32.6) is essentially the probability of transitioning from $|A\rangle$ to $|B\rangle$ after a time delay of t.

In the case where the direction of time is not predetermined, we have seen in the previous section that the construction of temporal paths of N steps each of duration τ is given by

$$[U_F(\tau) + U_B(\tau)]^N\,.$$

With this operator we can check whether a universe that is in state $|A\rangle$ can evolve to state $|B\rangle$ in N time steps of unspecified direction by seeing if the (un-normalized) probability

$$\begin{aligned} P_N(B|A) &= \left| \langle B| \left[U_F(\tau) + U_B(\tau)\right]^N |A\rangle \right|^2 \\ &= \langle A| \left[U_F^\dagger(\tau) + U_B^\dagger(\tau)\right]^N |B\rangle \langle B| \left[U_F(\tau) + U_B(\tau)\right]^N |A\rangle \end{aligned} \tag{32.7}$$

is nonzero. The converse situation where the universe can evolve from $|B\rangle$ to $|A\rangle$ is given by checking for a nonzero value of

$$\begin{aligned} P_N(A|B) &= \left| \langle A| \left[U_F(\tau) + U_B(\tau)\right]^N |B\rangle \right|^2 \\ &= \langle B| \left[U_F^\dagger(\tau) + U_B^\dagger(\tau)\right]^N |A\rangle \langle A| \left[U_F(\tau) + U_B(\tau)\right]^N |B\rangle \,. \end{aligned} \tag{32.8}$$

For a non T-violating universe $H_B = H_F$ and so $U_F^\dagger(\tau) = \exp(iH_F\tau) = U_B(\tau)$ and $U_B^\dagger(\tau) = \exp(-iH_B\tau) \quad = U_F(\tau)$, and thus $P_N(B|A) = P_N(A|B)$. In other words, if a non T-violating universe can evolve from $|A\rangle$ to $|B\rangle$ it can also evolve from $|B\rangle$ to $|A\rangle$, as one would expect if both directions of time are allowed.

However, the situation is very different for a T violating universe, i.e. one for which for $H_B \neq H_F$. In this case the right sides of Eqs. (32.7) and (32.8) are not equal in general. Indeed, in the extreme case, one may be zero while the other is not. To see this let $|B\rangle$ be given by

$$|B\rangle = \left[U_F(\tau) + U_B(\tau)\right]^M |A\rangle \tag{32.9}$$

where M is an integer, for which we find

$$\begin{aligned} P_N(B|A) &= \left| \langle B| \left[U_F(\tau) + U_B(\tau)\right]^N |A\rangle \right|^2 \\ &= \left| \langle A| \left[U_F^\dagger(\tau) + U_B^\dagger(\tau)\right]^M \left[U_F(\tau) + U_B(\tau)\right]^N |A\rangle \right|^2 \,. \end{aligned}$$

In the special case where $N = M$

$$P_M(B|A) = |\langle B|B\rangle|^2$$

which is the square of the norm of $|B\rangle$. Normalizing the probability results in a value of unity. This implies that the evolution from $|A\rangle$ to $|B\rangle$ is *guaranteed* as would be expected from Eq. (32.9). However for the converse case,

$$\begin{aligned} P_N(A|B) &= \left| \langle A| \left[U_F(\tau) + U_B(\tau)\right]^N |B\rangle \right|^2 \\ &= \left| \langle A| \left[U_F(\tau) + U_B(\tau)\right]^{N+M} |A\rangle \right|^2 \,. \end{aligned} \tag{32.10}$$

If the value of M is sufficiently large that destructive interference eliminates all paths but two, as illustrated by Eq. (32.5), then the state $|B\rangle$ can be written as

$$|B\rangle \approx [U_F^M(\tau) + U_B^M(\tau)]\,|A\rangle$$

and it follows that Eq. (32.10) becomes

$$\begin{aligned} P_N(A|B) &\approx \left|\langle A|\,U_F^{N+M}(\tau) + U_B^{N+M}(\tau)\,|A\rangle\right|^2 \\ &= \left|\langle A|\,U_F^{N+M}(\tau)\,|A\rangle + \langle A|\,U_B^{N+M}(\tau)\,|A\rangle\right|^2. \end{aligned} \tag{32.11}$$

Moreover, let the value of M be such that

$$\langle A|B\rangle = 0$$

and so $P_0(A|B) = 0$. It would be reasonable to assume that the observable universe is not cyclical and never returns to the same state in a fixed direction of time. For our model this means that $\langle A|\,U_F^{N+M}(\tau)\,|A\rangle = 0$ and $\langle A|\,U_B^{N+M}(\tau)\,|A\rangle = 0$, and so from Eq. (32.11)

$$P_N(A|B) \approx 0$$

for any positive value of N. We have succeeded in showing, therefore, that *the universe does not evolve from* $|B\rangle$ *to* $|A\rangle$ *despite the fact that it is guaranteed to evolve from* $|A\rangle$ *to* $|B\rangle$. This is a new kind of irreversibility. But to fully appreciate what it means we first need to consider the relationship between the two terms on the right side of Eq. (32.5).

32.4 Superposition of Matter and Antimatter

If the state $|\psi_0\rangle$ is T invariant, i.e. if $|\psi_0\rangle = T\,|\psi_0\rangle$, the two components of the right side of Eq.(32.5) are related by the time reversal operation

$$U_B(N\tau)\,|\psi_0\rangle = T U_F(N\tau)\,|\psi_0\rangle\,.$$

Multiplying both sides by CP, where C and P are the charge conjugation and parity inversion operators, and invoking the CPT theorem shows that $U_F(N\tau)\,|\psi_0\rangle$ and $U_B(N\tau)\,|\psi_0\rangle$ are related by charge-parity conjugation. It is interesting to consider the case where the evolution $U_F(N\tau)\,|\psi_0\rangle$ symbolizes baryogenesis. In this case, if $U_F(N\tau)\,|\psi_0\rangle$ is dominated by matter then $U_B(N\tau)\,|\psi_0\rangle$ is dominated by antimatter. Every particle in $U_F(N\tau)\,|\psi_0\rangle$ has a corresponding partner antiparticle in $U_B(N\tau)\,|\psi_0\rangle$ and vise versa, and so a planet composed of matter in $U_F(N\tau)\,|\psi_0\rangle$ is correspondingly composed of antimatter in $U_B(N\tau)\,|\psi_0\rangle$. The two versions of the planet evolve independently of each other according to Eq. (32.5) and so particle-antiparticle annihilation between the two versions of the planet would not occur. Moreover the evolution of the matter-dominated universe in $U_F(N\tau)\,|\psi_0\rangle$ is

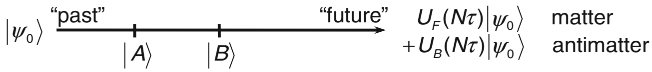


Fig. 32.2 A reinterpretation of Eq. (32.5) treating the two directions, forward and backward, as equivalent. The states $|A\rangle$*and* $|B\rangle$ are related by the relationship given in Eq. (32.9)

equivalent, by the CPT theorem, to the evolution of the antimatter-dominated universe in $U_B(N\tau)\,|\psi_0\rangle$. This symmetry between $U_F(N\tau)\,|\psi_0\rangle$ and $U_B(N\tau)\,|\psi_0\rangle$ implies that evolution in the forward direction is equivalent to the evolution in the backward direction. The directions forward and backward are redundant in this sense. The two arrows in Fig. 32.1b can therefore be considered to be one as illustrated in Fig. 32.2.

32.5 Discussion

We began with a model of the universe that has no presumed direction of time. Time evolution is described by a superposition of all possible paths that zigzag forwards and backwards through time as illustrated in Fig. 32.1a. We found that by incorporating sufficient T-violating processes only two paths representing always forwards and always backwards evolution survive destructive interference, as illustrated in Fig. 32.1b. We also showed that the two directions of time are essentially equivalent, as illustrated in Fig. 32.2.

Moreover we found that the universe can evolve from one state $|A\rangle$ to another $|B\rangle$, as illustrated in Fig. 32.2, but not vice versa. This endows the universe with a temporal direction pointing from a "past" to a "future". We have therefore described a physical mechanism which, if it exists in nature, could account for the directedness of time that we observe in our universe. This mechanism has the potential of being the *cause* of the direction of time. It is quite different to the asymmetry of the thermodynamic arrow which is an *effect* of a fixed direction of time and an incomplete knowledge of a system.

References

1. Christenson JH, Cronin JW, Fitch VL, Turlay R (1964) Evidence for the 2π decay of the K_2^0 meson. Phys Rev Lett 13:138
2. Eddington AS (1928) The nature of the physical world. Macmillan, New York
3. Feynman RP (1948) Space-time approach to non-relativistic quantum mechanics. Rev Mod Phys 20:367–387
4. Lee TD, Wolfenstein L (1965) Analysis of CP-noninvariant interactions and the K_1^0, K_2^0 system. Phys Rev 138:B1490–B1496
5. Price H (1996) Time's arrow and Archimedes' point. Oxford University Press, New York
6. Vaccaro JA (2011) T violation and the unidirectionality of time. Found Phys 41:1569–1596
7. Wigner EP (1959) Group theory and its application to the quantum mechanics of atomic spectra. Academic Press, New York
8. Yao W-M et al (2006) Review of particle physics. J Phys G: Nucl Part Phys 33:666–684

Chapter 33
Study of Stability Matter Problem in Micropolar Generalised Thermoelastic

Arminder Singh and Gurpinder Singh

Abstract The theory of micropolar thermoelasticity has many applications. One form of the recent years concerning the problem of propagation of thermal waves at finite speed and the possibility of "second sound" effects established a new thermo mechanical theory of deformable media that uses a general entropy balance as postulated and the theory is illustrated in detail in the context of flow of heat in a rigid solid, with particular reference to the propagation of thermal waves at finite speed. Then theory of thermoelasticity for non-polar bodies, based on the new procedures, was discussed and employed the eigen value approach to study the effect of rotation and relaxation time in two dimensional problem of generalized thermoelasticity. Recently investigation shows the dynamic response of a homogeneous, isotropic, generalized thermoelastic half-space with voids subjected to normal, tangential force and thermal stress. In this paper we introduce the eigen value approach, following Laplace and Fourier transformation has been employed to find the general solution of the field equation in a micropolar generalized thermoelastic medium for plane strain problem. An application of an infinite space with an impulsive mechanical source has been taken to illustrate the utility of the approach. The integral transformation has been inverted by using a numerical inversion technique to get result in physical domain. The result in the form of normal displacement, normal force stress, tangential force stress, tangential couple stress and temperature field components have been obtained numerically and illustrated graphically. Special case of a thermoelastic solid has also been deduced.

A. Singh (✉)
Kounsil C.T. Institute of Technology, Jalandhar, India

G. Singh
Samra, Lyallpur Khalsa College, Jalandhar, India

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_33,

33.1 Introduction

The theory of micropolarthermoelasticity has been a subject of intensive study. A comprehensive review of works on the subject was given by [4] and [19]. There has been very much written in recent years concerning the problem of propagation of thermal waves at finite speed. A generalized theory of linear micropolarthermoelasticity that admits the possibility of “second sound” effects was established by [1]. Recently, [9] established a new thermomechanical theory of deformable media that uses a general entropy balance as postulated by [8]. The theory is illustrated in detail in the context of flow of heat in a rigid solid, with particular reference to the propagation of thermal waves at finite speed. A theory of thermoelasticity for non-polar bodies, based on the new procedures, was discussed by [10]. Bahshi et al. [2] employed the eigen value approach to study the effect of rotation and relaxation time in two dimensional problem of generalized thermoelasticity. Kumar and Rani [15] studied the deformation due to mechanical and thermal sources in generalized orthorhomtic-thermoelastic material. Kumar and Rani [16] investigated the dynamic response of a homogeneous, isotropic, generalized thermoelastic half-space with voids subjected to normal, tangential force and thermal stress. The micropolar theory was extended to include thermal effects by [4] and [19]. Kumar and Chadha [13] derived the expressions for displacements, microrotation, force stress, couple stress and first moment for a half - space subjected to an arbitrary temperature field and a particular case of line heat source has been discussed in detail. The uniqueness of the solution of some boundary value problems of the linear micropolarthermoelasticity was investigated by [3]. Passarella [21] solved the initial-boundary value problem for micropolarthermoelasticity and proved a uniqueness theorem for the problem. Mahalanabis and Manna [17] discussed eigen value approach to linear micropolarthermoelasticity by arranging basic equations of elasticity in the form of matrix deferential equation in the Hankel transform and extended the approach to linear thermoelasticity. Marin and Lupu [18] investigated harmonic vibrations in thermoelasticity of micropolar bodies. Kumar and Deswal [14] discussed the disturbance due to mechanical and thermal sources in homogeneous isotropic micropolar generalized thermoelastic half-space.

33.2 Formulation and Solution of the Problem

We consider a homogeneous, isotropic, micropolar generalized thermoelastic solid in an undisturbed state and initially at uniform temperature. We take a cartesian system (x, y, z) and z-axis pointing vertically into the medium.

Following [6], [12] and [11], the field equations and the constitutive relations in micropolar generalized thermoelastic solid without body forces, body couples and heat sources can be written as

$$(\lambda + 2\mu + K)\,\nabla(\nabla.\mathbf{u}) - (\mu + K)\,\nabla \times \nabla \times \mathbf{u} + K\nabla \times \boldsymbol{\phi} - \nu\left(1 + \tau_1 \frac{\partial}{\partial t}\right)\nabla \mathrm{T} = \rho \frac{\partial^2 \mathbf{u}}{\partial t^2}$$

$$(\alpha + \beta + \gamma)\,\nabla(\nabla.\boldsymbol{\phi}) - \gamma\nabla \times \nabla \times \boldsymbol{\phi} + K\nabla \times \mathbf{u} - 2\mathrm{K}\boldsymbol{\phi} = \rho\, j \frac{\partial^2 \boldsymbol{\phi}}{\partial t^2}$$

$$K^*\nabla^2 T = \rho C^*\left(\frac{\partial T}{\partial t} + \tau_0 \frac{\partial^2 \mathrm{T}}{\partial t^2}\right) + \nu T_0\left(\frac{\partial}{\partial t} + \Xi\, \tau_0 \frac{\partial^2}{\partial t^2}\right)\nabla \cdot \mathbf{u}$$

$$m_{ij} = \alpha\varphi_{r,r}\delta_{ij} + \beta\varphi_{i,j} + \gamma\varphi_{j,i},$$

$$t_{ij} = \lambda u_{r,r}\delta_{ij} + \mu\left(u_{i,j} + u_{j,i}\right) + K\left(u_{j,i} - \varepsilon_{ijr}\varphi_r\right) - \nu\left(\mathrm{T} + \tau_1 \frac{\partial T}{\partial t}\right)\delta_{\mathrm{ij}},$$

For the L-S (Lord Shulman) theory $\tau_1 = 0$, $\Xi = 1$ and for G - L (Green Lindsay) theory $\tau_1 = 0$, $\Xi = 0$,

The thermal relaxations τ_0 and τ_1 satisfy the inequality $\tau_1 \geq \tau_0 > 0$ for the G-L theory only. However, it has been proved by [22] that the inequalities are not mandatory for τ_0 and τ_1 to follow.

For two dimensional plane strain problem parallel to *xz*-plane, we assume

$$\mathbf{u} = (u_1, 0, u_3)\,, \quad \boldsymbol{\phi} = (0, \phi_2, 0)$$

The displacement components $u_1, u_3$and microrotation component depend upon x, z and t and are independent of co-ordinate y, so that $\frac{\partial}{\partial y} \equiv 0$. With these considerations and using (2.6) and introducing the non-dimensional quantities as

$$x' = \frac{\omega^* x}{C_1}, \qquad z' = \frac{\omega^* z}{C_1},$$
$$T' = \frac{T}{T_0}, \qquad u_1' = \frac{\rho\omega^* C_1 u_1}{\nu T_0},$$
$$m_{32}' = \frac{\omega^*}{C_1 \nu T_0} m_{32},$$

where

$$\omega^* = \frac{C^*(\lambda + 2\mu)}{K^*}, \qquad C_1^2 = \frac{\lambda + 2\mu}{\rho}.$$

Now applying Laplace and Fourier transform defined by

$$\overline{f}(x, z, p) = \int_0^\infty f(x, z, t) exp(-pt) dt,$$
$$\tilde{f}(\xi, z, p) = \int_{-\infty}^\infty \overline{f}(x, z, p) \mathrm{e}(-\iota\xi x) dx,$$

on the set of Eq. (2.1)–(2.3), after suppressing primes, we get

$$\frac{d^2\tilde{u}_1}{dz^2} = \frac{1}{\mathrm{m}_3}\left[\left(\mathrm{m}_1\xi^2 + \mathrm{p}^2\right)\tilde{u}_1 - \iota\xi\mathrm{m}_2\frac{d\tilde{u}_3}{dz} + \mathrm{m}_4\frac{d\tilde{\phi}_2}{dz} + \iota\xi\left(1+\tau_1 p\right)\tilde{T}\right]$$
$$\frac{d^2\tilde{u}_3}{dz^2} = \frac{1}{\mathrm{m}_1}\left[-\iota\mathrm{m}_2\xi\frac{d\tilde{u}_1}{dz} + \left(\mathrm{m}_3\xi^2 + \mathrm{p}^2\right)\tilde{u}_3 - \mathrm{m}_4\iota\xi\tilde{\phi}_2 + \frac{d\tilde{T}_1}{dz}\right]$$
$$\frac{d^2\tilde{\phi}_2}{dz^2} = -\mathrm{m}_5\frac{d\tilde{u}_1}{dz} + \iota\xi\mathrm{m}_5\tilde{u}_3 + \left(2\mathrm{m}_5 + \xi^2 + \mathrm{m}_6\mathrm{p}^2\right)\tilde{\phi}_2$$
$$\frac{d^2\tilde{T}}{dz^2} = \varepsilon p\left(1+\tau_0 p\Xi\right)\left\{\iota\xi\tilde{u}_1 + \frac{d\tilde{u}_3}{dz}\right\} + \left\{\xi^2 + p\left(1+\tau_0 p\right)\tilde{T}\right\}$$

where

$$m_1 = \frac{\lambda + 2\mu + K}{\rho C_1^2}, \qquad m_2 = \frac{\lambda + \mu}{\rho C_1^2},$$
$$m_4 = \frac{K}{\rho C_1^2}, \qquad m_5 = \frac{K C_1^2}{\rho \omega^{*2}},$$
$$m_7 = \frac{\mu}{\rho C_1^2}, \qquad m_8 = \frac{\lambda}{\rho C_1^2},$$
$$\varepsilon = \frac{\mathrm{T}_0\beta_1^2}{\rho K^*\omega^*}.$$

Equations (2.9)– (2.12) can be written in the vector matrix differential equation form as

$$\frac{d}{dz}W\left(\xi, z, p\right) = \mathrm{A}\left(\xi, p\right)\mathrm{W}\left(\xi, z, p\right)$$

where

$$W = \begin{bmatrix} U \\ DU \end{bmatrix}, \qquad A = \begin{bmatrix} O & I \\ A_2 & A_1 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} 0 & f_{12} & f_{13} & 0 \\ f_{21} & 0 & 0 & f_{24} \\ f_{31} & 0 & 0 & 0 \\ 0 & f_{42} & 0 & 0 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} g_{11} & 0 & 0 & g_{14} \\ 0 & g_{22} & g_{23} & 0 \\ 0 & g_{32} & g_{33} & 0 \\ g_{41} & 0 & 0 & g_{44} \end{bmatrix}$$

and O is the Null matrix of order 4 with

$$f_{12} = \frac{-l\xi m_2}{m_3}, \qquad f_{13} = \frac{m_4}{m_3},$$
$$f_{31} = -m_5, \qquad f_{42} = \varepsilon p\left(1 + \tau_0 p\Xi\right),$$
$$g_{22} = \frac{\left(\mathrm{m}_3\xi^2 + \mathrm{p}^2\right)}{m_1}, \qquad g_{22} = \frac{-\mathrm{lm}_4\xi}{m_1},$$

$$g_{41} = l\varepsilon\xi p\,(1 + \tau_0 p\,\Xi)\,, \qquad g_{44} = \xi^2 + p\,(1 + \tau_0 p)\,,$$

To solve the Eq. (2.14), we take $W(\xi, z, p) = X(\xi, p)\mathrm{e}^{\mathrm{q}z}$ for some q, So we obtain

$$A(\xi, p)W(\xi, z, p) = qW(\xi, z, p),$$

This leads to Eigen value problem. The characteristic equation corresponding to the matrix A is given by

$$det\,(A - q\mathrm{I}) = 0$$

which on expansion provides us

$$q^8 - \sigma_1 q^6 + \sigma_2 q^4 - \sigma_3 q^2 + \sigma_4 = 0$$

where

$$\begin{aligned}
\sigma_1 &= g_{11} + g_{22} + g_{33} + g_{44} + f_{24}f_{42} + f_{12}f_{21} + f_{13}f_{31}\,,\\
\sigma_2 &= g_{11}g_{22} + g_{22}g_{33} + g_{33}g_{11} + g_{44}g_{11} + g_{44}g_{22} + g_{44}g_{33} + f_{12}f_{24}g_{41}\\
&\quad + f_{24}f_{42}g_{11} + f_{24}f_{42}g_{33} + f_{24}f_{42}f_{13}f_{31} - g_{32}g_{23} + f_{12}f_{21}g_{33}\\
&\quad + f_{12}f_{21}g_{44} - f_{12}f_{31}g_{23} + f_{13}f_{31}g_{22} + f_{13}f_{31}g_{44} - g_{14}g_{41} - g_{14}f_{21}f_{42},\\
\sigma_3 &= g_{11}g_{22}g_{33} + g_{22}g_{33}g_{44} + g_{33}g_{44}g_{11} + g_{44}g_{11}g_{22} - g_{11}g_{23}g_{32} - g_{44}g_{23}g_{32}\\
&\quad + f_{24}f_{42}g_{11}g_{33} + f_{13}f_{24}g_{41}g_{32} + f_{31}f_{42}g_{14}g_{23} + f_{12}f_{21}g_{33}g_{44}\\
&\quad - f_{13}f_{21}g_{32}g_{44} - f_{12}f_{31}g_{23}g_{44} - f_{42}f_{21}g_{14}g_{33} + f_{13}f_{31}g_{22}g_{44}\\
&\quad - g_{14}g_{41}g_{22} - g_{14}g_{41}g_{33},\\
\sigma_4 &= g_{11}g_{22}g_{33}g_{44} - g_{23}g_{32}g_{11}g_{44} + g_{14}g_{41}g_{22}g_{33} + g_{23}g_{32}g_{14}g_{41}.
\end{aligned}$$

The eigen values of the matrix A are the characteristic roots of the Eq. (2.19). The vectors X (ξ, p) corresponding to the eigen values q_s can be determined by solving the homogeneous equations

$$[A - qI]\,\mathrm{X_s}\,(\xi, \mathrm{p}) = 0.$$

The set of eigen vectors $\mathrm{X_s}\,(\xi, \mathrm{p})$; s = 1, 2, 3, ..., 8 may be defined as

$$\mathrm{X_s}(\xi, \mathrm{p}) = \begin{bmatrix} \mathrm{X_{s1}}(\xi, \mathrm{p}) \\ \mathrm{X_{s2}}(\xi, \mathrm{p}) \end{bmatrix}$$

where

$$\mathrm{X}_{s1}(\xi, \mathrm{p}) = \begin{bmatrix} \mathrm{a_s q_s} \\ \mathrm{b_s} \\ -\xi \\ \mathrm{c_s} \end{bmatrix}, \qquad \mathrm{X_{s2}}(\xi, \mathrm{p}) = \begin{bmatrix} \mathrm{a_s q_s^2} \\ \mathrm{b_s q_s} \\ -\xi\,\mathrm{q_s} \\ \mathrm{c_s q_s} \end{bmatrix},$$

$$X_{l1}(\xi, p) = \begin{bmatrix} -a_s q_s \\ b_s \\ -\xi \\ c_s \end{bmatrix}, \qquad X_{l2}(\xi, p) = \begin{bmatrix} a_s q_s^2 \\ -b_s q_s \\ \xi q_s \\ -c_s q_s \end{bmatrix},$$

$$a_s = \frac{-\xi}{m_3 \Delta_s}\left[\left\{\xi^2 + p\,(1+\tau_0 p) - q_s^2\right\}\left\{m_4 m_5 + m_2\left(2m_5 + \xi^2 + p^2 m_6 - q_s^2\right)\right.\right.$$
$$\left.\left. + \varepsilon p\left(2m_5 + \xi^2 + p^2 m_6 - q_s^2\right)(1+\tau_1 p)\,(1+\tau_0 p)\,\Xi\right\}\right]$$
$$b_s = \frac{\iota}{m_3 \Delta_s}\left[\left\{\xi^2 + p\,(1+\tau_0 p) - q_s^2\right\}\left\{m_4 m_5 q_s^2 + \left(2m_5 + \xi^2 + p^2 m_6 - q_s^2\right)\right.\right.$$
$$\left.\left. \times\left(m_1 \xi^2 + p^2 - m_3 q_s^2\right)\right\} + \varepsilon p \xi^2\left(2m_5 + \xi^2 + p^2 m_6 - q_s^2\right)(1+\tau_1 p)\,(1+\tau_0 p)\,\Xi\right]$$
$$c_s = \frac{\varepsilon p q_s\,(1+\tau_0 p\,\Xi)\,(\iota \xi a_s + b_s)}{\left[q_s^2 - \left\{\xi^2 + p\,(1+\tau_0 p)\right\}\right]}$$
$$\Delta_s = \frac{m_5}{m_3}[\left\{\xi^2 + p\,(1+\tau_0 p) - q_s^2\right\}\{m_2 q_s^2 - \left(m_1 \xi^2 + p^2 - q_s^2 m_3\right)\}$$
$$+\ \varepsilon p\,(1+\tau_0 p\,\Xi)\,(1+\tau_1 p)\left(q_s^2 - \xi^2\right)]$$

Thus solution of Eq. (2.14) is as given by [23]

$$W(\xi, z, p) = \sum_{s=1}^{4}\left[E_s X_s\,(\xi, p)\,e^{q_s z} + E_{s+4} X_{s+4}\,(\xi, p)\,e^{-q_s z}\right]$$

$E_1, E_2, E_3, E_4, E_5, E_6, E_7$ and E_8 are eight arbitrary constants. The Eq. (2.32) represents a general solution of the plane strain problem for isotropic, micropolar generalized thermoelastic solid and gives the displacement, microrotation and temperature field in the transformed domain.

33.3 Applications

Mechanical Source

We consider an infinite micropolar generalized thermoelastic space in which a concentrated force where F_0 is the magnitude of the force, $F = -F_0 \delta(x)\delta(t)$ acting in the direction of the z-axis at the origin of the Cartesian co-ordinate system as shown in Fig. 33.1. The boundary condition for present problem on the plane $z = 0$ are

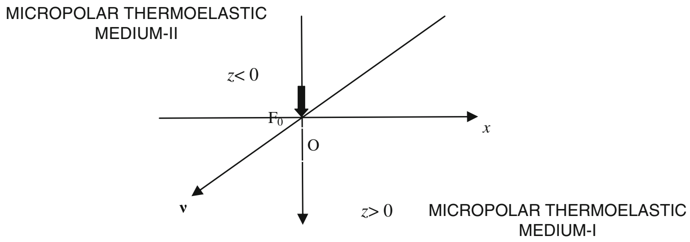


Fig. 33.1 .

$$
\begin{aligned}
&\mathrm{u}_1(\mathrm{x}, 0^+, \mathrm{t}) - \mathrm{u}_1(\mathrm{x}, 0^-, \mathrm{t}) = 0, \mathrm{u}_3(\mathrm{x}, 0^+, \mathrm{t}) - \mathrm{u}_3(\mathrm{x}, 0^-, \mathrm{t}) = 0,\\
&\phi_2(\mathrm{x}, 0^+, \mathrm{t}) - \phi_2(\mathrm{x}, 0^-, \mathrm{t}) = 0, T(\mathrm{x}, 0^+, \mathrm{t}) - T(\mathrm{x}, 0^-, \mathrm{t}) = 0,\\
&\frac{\partial T}{\partial z}\left(\mathrm{x}, 0^+, \mathrm{t}\right) - \frac{\partial T}{\partial z}\left(\mathrm{x}, 0^-, \mathrm{t}\right) = 0, t_{31}\left(\mathrm{x}, 0^+, \mathrm{t}\right) - t_{31}\left(\mathrm{x}, 0^-, \mathrm{t}\right) = 0,\\
&t_{33}\left(\mathrm{x}, 0^+, \mathrm{t}\right) - t_{33}\left(\mathrm{x}, 0^-, \mathrm{t}\right) = -F_0\delta(\mathrm{x})\delta(\mathrm{t}), m_{32}\left(\mathrm{x}, 0^+, \mathrm{t}\right) - m_{32}\left(\mathrm{x}, 0^-, \mathrm{t}\right) = 0
\end{aligned}
$$

Making use of Eq. (2.6)–(2.7) and $F_0' = \frac{F_0}{K}$ in Eq. (2.4)–(2.5), we get the stresses in the non-dimensional form with primes. After suppressing the primes, we apply Laplace and Fourier transforms defined by Eq. (2.8) on the resulting equations and from Eq. (3.1), we get transformed components of displacement, microrotation, temperature field, tangential force stress, normal force stress and tangential couple stress for $z > 0$ are given by

$$
\begin{aligned}
&\tilde{\mathrm{u}}_1(\xi, \mathrm{z}, \mathrm{p}) = -\left\{\mathrm{a}_1\mathrm{q}_1\mathrm{E}_5e^{-\mathrm{q}_1 z} + \mathrm{a}_2\mathrm{q}_2\mathrm{E}_6e^{-\mathrm{q}_2 z} + \mathrm{a}_3\mathrm{q}_3\mathrm{E}_7e^{-\mathrm{q}_3 z} + \mathrm{a}_4\mathrm{q}_4\mathrm{E}_8e^{-\mathrm{q}_4 z}\right\},\\
&\tilde{\mathrm{u}}_3(\xi, \mathrm{z}, \mathrm{p}) = \mathrm{b}_1\mathrm{E}_5e^{-\mathrm{q}_1 z} + \mathrm{b}_2\mathrm{E}_6e^{-\mathrm{q}_2 z} + \mathrm{b}_3\mathrm{E}_7e^{-\mathrm{q}_3 z} + \mathrm{b}_4\mathrm{E}_8e^{-\mathrm{q}_4 z},\\
&\tilde{\phi}_2(\xi, \mathrm{z}, \mathrm{p}) = -\xi\left\{\mathrm{E}_5e^{-\mathrm{q}_1 z} + \mathrm{E}_6e^{-\mathrm{q}_2 z} + \mathrm{E}_7e^{-\mathrm{q}_3 z} + \mathrm{E}_8e^{-\mathrm{q}_4 z}\right\},\\
&\tilde{T}(\xi, \mathrm{z}, \mathrm{p}) = \mathrm{c}_1\mathrm{E}_5e^{-\mathrm{q}_1 z} + \mathrm{c}_2\mathrm{E}_6e^{-\mathrm{q}_2 z} + \mathrm{c}_3\mathrm{E}_7e^{-\mathrm{q}_3 z} + \mathrm{c}_4\mathrm{E}_8e^{-\mathrm{q}_4 z},\\
&\tilde{\mathrm{t}}_{31}(\xi, \mathrm{z}, \mathrm{p}) = \left(\mathrm{m}_3\mathrm{a}_1\mathrm{q}_1^2 + \iota\xi\mathrm{b}_1\mathrm{s}_{10} + \xi\mathrm{m}_4\right)\mathrm{E}_5e^{-\mathrm{q}_1 z} + \left(\mathrm{m}_3\mathrm{a}_2\mathrm{q}_2^2 + \iota\xi\mathrm{b}_2\mathrm{m}_7 + \xi\mathrm{m}_4\right)\mathrm{E}_6e^{-\mathrm{q}_2 z} +\\
&\qquad \left(\mathrm{m}_3\mathrm{a}_3\mathrm{q}_3^2 + \iota\xi\mathrm{b}_3\mathrm{m}_7 + \xi\mathrm{m}_4\right)\mathrm{E}_7e^{-\mathrm{q}_3 z} + \left(\mathrm{m}_3\mathrm{a}_4\mathrm{q}_4^2 + \iota\xi rmb_4\mathrm{m}_7 + \xi\mathrm{m}_7\right)\mathrm{E}_8e^{-\mathrm{q}_4 z},\\
&\tilde{\mathrm{t}}_{33}(\xi, \mathrm{z}, \mathrm{p}) = -[(\iota\xi\mathrm{m}_8\mathrm{a}_1\mathrm{q}_1 + \mathrm{m}_1\mathrm{b}_1\mathrm{q}_1 + \mathrm{c}_1\left(1 + \tau_1 p\right))\mathrm{E}_5e^{-\mathrm{q}_1 z}\\
&\qquad + (\iota\xi\mathrm{m}_8 a_2\mathrm{q}_2 + \mathrm{m}_1\mathrm{b}_2\mathrm{q}_2 + \mathrm{c}_2\left(1 + \tau_1 p\right))\mathrm{E}_6e^{-\mathrm{q}_2 z}\\
&\qquad + (\iota\xi\mathrm{m}_8 a_3\mathrm{q}_3 + \mathrm{m}_1\mathrm{b}_3\mathrm{q}_3 + \mathrm{c}_3\left(1 + \tau_1 p\right))\mathrm{E}_7e^{-\mathrm{q}_3 z}\\
&\qquad + (\iota\xi\mathrm{m}_8 a_4\mathrm{q}_4 + \mathrm{m}_1\mathrm{b}_4\mathrm{q}_4 + \mathrm{c}_4\left(1 + \tau_1 p\right))\mathrm{E}_8e^{-\mathrm{q}_4 z},\\
&\tilde{m}_{32}(\xi, \mathrm{z}, \mathrm{p}) = \xi\mathrm{s}_8\left\{\mathrm{q}_1\mathrm{E}_5e^{-\mathrm{q}_1 z} + \mathrm{q}_2\mathrm{E}_6e^{-\mathrm{q}_2 z} + \mathrm{q}_3\mathrm{E}_7e^{-\mathrm{q}_3 z} + \mathrm{q}_4\mathrm{E}_8e^{-\mathrm{q}_4 z}\right\},
\end{aligned}
$$

for $z<0$, the above expressions get suitably modified, e.g.

$$
\tilde{\mathrm{u}}_1(\xi, \mathrm{z}, \mathrm{p}) = \mathrm{a}_1\mathrm{q}_1\mathrm{E}_1e^{\mathrm{q}_1 z} + \mathrm{a}_2\mathrm{q}_2\mathrm{E}_2e^{\mathrm{q}_2 z} + \mathrm{a}_3\mathrm{q}_3\mathrm{E}_3e^{\mathrm{q}_3 z} + \mathrm{a}_4\mathrm{q}_4\mathrm{E}_4e^{\mathrm{q}_4 z},
$$

Making use of the transformed displacements, microrotation, microstretch and stresses given by (3.6)–(3.12) in the transformed boundary conditions, we obtain eight linear relations between the E_i's, which on solving gives

$$E_1 = E_5 = \frac{F_0}{2q_1\Delta_1}\left[c_2\left(a_3-a_4\right)+c_3\left(a_4-a_2\right)+c_4\left(a_2-a_3\right)\right],$$
$$E_2 = E_6 = \frac{F_0}{2q_2\Delta_1}\left[c_1\left(a_4-a_3\right)+c_3\left(a_1-a_4\right)+c_4\left(a_3-a_1\right)\right],$$
$$E_3 = E_7 = \frac{F_0}{2q_3\Delta_1}\left[c_1\left(a_2-a_4\right)+c_2\left(a_4-a_1\right)+c_4\left(a_1-a_2\right)\right],$$
$$E_4 = E_8 = \frac{F_0}{2q_4\Delta_1}\left[c_1\left(a_3-a_2\right)+c_2\left(a_1-a_3\right)+c_3\left(a_2-a_1\right)\right],$$

where

$$\begin{aligned}\Delta_1 = m_1[&c_1\left\{(a_2b_3-a_3b_2)+(a_3b_4-a_4b_3)+(a_4b_2-a_2b_4)\right\}\\ &+c_2\left\{(a_3b_1-a_1b_3)+(a_1b_4-a_4b_1)+(a_4b_3-a_3b_4)\right\}\\ &+c_3\left\{(a_1b_2-a_2b_1)+(a_4b_1-a_1b_4)+(a_2b_4-a_4b_2)\right\}\\ &+c_4\left\{(a_2b_1-a_1b_2)+(a_1b_3-a_3b_1)+(a_3b_2-a_2b_3)\right\}],\end{aligned}$$

Thus functions $\tilde{u}_1, \tilde{u}_3, \tilde{\varphi}_2, \tilde{T}, \tilde{t}_{31}, \tilde{t}_{33}$ and $\tilde{m}_{32}$ have been determined in the transformed domain and these enable us to find the displacements, microrotation, temperature field and stresses.

Case I : For L-S theory, a_s, b_s and c_s in the expressions (3.5)–(3.12) take the form

$$\begin{aligned}a_s &= \frac{-\xi}{\mathrm{m}_3\Delta_\mathrm{s}}[\left\{\xi^2+p\left(1+\tau_0 p\right)-q_s^2\right\}\{m_4m_5+m_2\left(2m_5+\xi^2+p^2m_6-q_s^2\right)\\ &\quad+\ \varepsilon p\left(2m_5+\xi^2+p^2m_6-q_s^2\right)\left(1+\tau_0 p\right)\}],\\ b_s &= \frac{\iota}{\mathrm{m}_3\Delta_\mathrm{s}}[\left\{\xi^2+p\left(1+\tau_0 p\right)-q_s^2\right\}\{m_4m_5q_s^2+\left(2m_5+\xi^2+p^2m_6-q_s^2\right)\\ &\quad\times\left(m_1\xi^2+p^2-m_3q_s^2\right)\}+\ \varepsilon p\xi^2\left(2m_5+\xi^2+p^2m_6-q_s^2\right)\left(1+\tau_0 p\right)],\\ c_s &= \frac{\varepsilon p q_s\left(\iota\xi a_s+b_s\right)}{\left[q_s^2-\left\{\xi^2+p\left(1+\tau_0 p\right)\right\}\right]},\end{aligned}$$

where

$$\begin{aligned}\Delta_s = \frac{\mathrm{m}_5}{\mathrm{m}_3}[&\left\{\xi^2+p\left(1+\tau_0 p\right)-q_s^2\right\}\{m_2q_s^2-\left(m_1\xi^2+p^2-q_s^2m_3\right)\}\\ &+\ \varepsilon p\left(1+\tau_1 p\right)\left(q_s^2-\xi^2\right)]\ ;\ \mathrm{s}=1,2,3,4\end{aligned}$$

and $\pm q_s$ (s = 1, 2, 3, 4) are roots of the Eq. (2.19) in which $\sigma_1, \sigma_2, \sigma_3$ *and* σ_4 are obtained respectively from expressions(2.20)–(2.23) by taking $\tau_1 = 0$, $\Xi = 1$.

Case II : For G-L theory, as b's and c's in the expressions (3.5)–(3.12) take the form

$$a_s = \frac{-\xi}{\mathrm{m}_3\Delta_\mathrm{s}}[\left\{\xi^2 + p(1+\tau_0 p) - q_s^2\right\}\{m_4 m_5 + m_2\left(2m_5 + \xi^2 + p^2 m_6 - q_s^2\right) + \varepsilon p\left(2m_5 + \xi^2 + p^2 m_6 - q_s^2\right)(1+\tau_1 p)\}],$$

$$b_s = \frac{\iota}{\mathrm{m}_3\Delta_\mathrm{s}}[\left\{\xi^2 + p(1+\tau_0 p) - q_s^2\right\}\{m_4 m_5 q_s^2 + \left(2m_5 + \xi^2 + p^2 m_6 - q_s^2\right)\left(m_1\xi^2 + p^2 - m_3 q_s^2\right)\} + \varepsilon p\xi^2\left(2m_5 + \xi^2 + p^2 m_6 - q_s^2\right)(1+\tau_1 p)],$$

$$c_s = \frac{\varepsilon p q_s(\iota\xi a_s + b_s)}{\left[q_s^2 - \left\{\xi^2 + p(1+\tau_0 p)\right\}\right]},$$

where

$$\Delta_s = \frac{\mathrm{m}_5}{\mathrm{m}_3}[\left\{\xi^2 + p(1+\tau_0 p) - q_s^2\right\}\{m_2 q_s^2 - \left(m_1\xi^2 + p^2 - q_s^2 m_3\right)\} + \varepsilon p(1+\tau_1 p)\left(q_s^2 - \xi^2\right)]\ ;\ \mathrm{s} = 1,2,3,4$$

and $\pm q_s$ (s =1, 2, 3, 4) are roots of the equation (2.19) in which $\sigma_1, \sigma_2, \sigma_3 and \sigma_4$are obtained respectively from expressions (2.20)–(2.23) by taking $\Xi = 0$

Case III : For Green and Naghdi theory (G-N), Eq. (2.1), (2.3) and (2.4) can be written as

$$(\lambda + 2\mu + K)\nabla(\nabla.\mathbf{u}) - (\mu + K)\nabla\times\nabla\times u + K\nabla\times\boldsymbol{\phi} - \nu\nabla\mathrm{T} = \rho\frac{\partial^2\mathbf{u}}{\partial\mathrm{t}^2}$$

$$\mathrm{K}^*\nabla^2\mathrm{T} = \rho\mathrm{C}^*\frac{\partial^2\mathrm{T}}{\partial\mathrm{t}^2} + \nu T_0\frac{\partial^2(\nabla.\mathbf{u})}{\partial\mathrm{t}^2}$$

$$t_{\mathrm{ij}} = \lambda u_{r,r}\delta_{ij} + \mu\left(u_{\mathrm{i,j}} + u_{\mathrm{j,i}}\right) + K\left(u_{\mathrm{j,i}} - \varepsilon_{ij,\mathrm{r}}\phi_r\right) - \nu T\delta_{ij}$$

and K^* is not the usual thermal conductivity but a material characteristics constant in G - N & theory and is given $K^*\left(= \frac{C^*(\lambda+2\mu)}{4}\right)$

With the help of Eq. (3.26)–(3.28) and following the procedure of the previous sections, we get the expressions for displacements, microrotation, temperature, field, force stresses and couple stress by taking in Eq. (3.5)-(3.11).

$$(1+\tau_1 p) = 1, \qquad (1+\tau_0 p) = 4p,$$

Particular Case I : Neglecting micropolarity effect i.e. $\alpha = \beta = \gamma = K = j = 0$ in Eq. (3.5)–(3.12), the expressions for displacement components, force stresses and temperature field are obtained in a thermoelastic medium as

$$\tilde{u}_1(\xi, z, p) = -\left\{a_1^* q_1 E_4^* e^{-q_1^* z} + a_2^* q_2^* E_5^* e^{-q_2^* z} + a_3^* q_3^* E_6^* e^{-q_3^* z}\right\},$$

$$\tilde{u}_3(\xi, z, p) = b_1^* E_4^* e^{-q_1^* z} + b_2^* E_5^* e^{-q_2^* z} + b_3^* E_6^* e^{-q_3^* z},$$

$$\tilde{T}(\xi, z, p) = -\xi\left\{E_4^* e^{-q_1^* z} + E_5^* e^{-q_2^* z} + c_3 E_6^* e^{-q_3^* z}\right\},$$

$$\tilde{t}_{31}(\xi, z, p) = m_7\{\left(a_1^* q_1^{*^2} + \iota\xi b_1^*\right) E_4^* e^{-q_1^* z} + \left(a_2^* q_2^{*^2} + \iota\xi b_2^*\right) E_5^* e^{-q_2^* z} + \left(a_3^* q_3^{*^2} + \iota\xi b_3^*\right) E_6^* e^{-q_3^* z}\},$$

$$\tilde{t}_{33}(\xi, z, p) = -[\left\{m_1^* q_1^* b_1^* + \iota\xi m_8 a_1^* q_1^* - \xi\,(1 + \tau_1 p)\right\} E_4^* e^{-q_1^* z} + \left\{m_1^* q_2^* b_2^* + \iota\xi m_8 a_2^* q_2^* - \xi\,(1 + \tau_1 p)\right\} E_5^* e^{-q_2^* z} + \left\{m_1^* q_3^* b_3^* + \iota\xi m_8 a_3^* q_3^* - \xi\,(1 + \tau_1 p)\right\} E_6^* e^{-q_3^* z}],$$

where

$$E_1^* = E_4^* = \frac{F_0\left(a_3^* - a_2^*\right)}{2q_1^* \Delta_1^*},$$

$$E_2^* = E_5^* = \frac{F_0\left(a_1^* - a_3^*\right)}{2q_2^* \Delta_1^*},$$

$$E_3^* = E_6^* = \frac{F_0\left(a_2^* - a_1^*\right)}{2q_3^* \Delta_1^*},$$

$$\Delta_1^* = m_1^*\left\{\left(a_2^* b_3^* - a_3^* b_2^*\right) + \left(a_3^* b_1^* - a_1^* b_3^*\right) + \left(a_1^* b_2 - a_2^* b_1^*\right)\right\},$$

$$a_s^* = \frac{-\xi}{m_3^* \Delta_s^*}[m_2\left\{\xi^2 + p\,(1 + \tau_0 p) - q_s^{*2}\right\} + \varepsilon p\,(1 + \tau_0 p\,\Xi)\,(1 + \tau_1 p)\}],$$

$$b_s^* = \frac{\iota}{m_3^* \Delta_s^*}[\left\{\xi^2 + p\,(1 + \tau_0 p) - q_s^{*2}\right\}\left(m_1^* \xi^2 + p^2 - m_3^* q_s^{*2}\right)\} + \varepsilon p \xi^2\,(1 + \tau_0 p\,\Xi)\,(1 + \tau_1 p)],$$

$$\Delta_s^* = \frac{\varepsilon p q_s^{*2}\,(1 + \tau_0 p\,\Xi)}{\xi m_3^*}\left\{\left(m_1^* \xi^2 + p^2 - q_s^{*2} m_3^*\right) - \xi^2 m_2\right\},$$

and $\pm q_s^*$ $(s = 1, 2, 3)$ are the roots of the equation

$$q^{*^6} - \sigma_1^* q^{0^4} + \sigma_2^* q^{*^2} - \sigma_3^* = 0$$

$$\sigma_1^* = \left(\frac{\mathrm{m}_1^*\xi^2 + p^2}{\mathrm{m}_3^*}\right) + \left(\frac{\mathrm{m}_3^*\xi^2 + p^2}{\mathrm{m}_1^*}\right) + \left\{\xi^2 + p\,(1 + \tau_0 p)\right\} + \frac{\varepsilon p}{\mathrm{m}_1^*}\,(1 + \tau_0 p\,\Xi)\,(1 + \tau_1 p) - \frac{\xi^2 \mathrm{m}_2^2}{\mathrm{m}_1^* \mathrm{m}_3^*},$$

$$\sigma_2^* = \left(\frac{\mathrm{m}_1^*\xi^2 + p^2}{\mathrm{m}_3^*}\right) + \left(\frac{\mathrm{m}_3^*\xi^2 + p^2}{\mathrm{m}_1^*}\right) + \left(\frac{\mathrm{m}_3^*\xi^2 + p^2}{\mathrm{m}_1^*}\right)\left\{\xi^2 + p\,(1 + \tau_0 p)\right\}$$

$$+ \left\{\xi^2 + p\,(1 + \tau_0 p)\right\}\left(\frac{\mathrm{m}_1^*\xi^2 + p^2}{\mathrm{m}_3^*}\right) + \frac{\varepsilon p}{\mathrm{m}_1^*}\,(1 + \tau_0 p\,\Xi)\,(1 + \tau_1 p)\left(\frac{\mathrm{m}_1^*\xi^2 + p^2}{\mathrm{m}_3^*}\right)$$

$$- \frac{\xi^2 \mathrm{m}_2^2}{\mathrm{m}_1^* \mathrm{m}_3^*}\left\{\xi^2 + p\,(1 + \tau_0 p)\right\} - \frac{2\varepsilon p \xi^2 \mathrm{m}2}{\mathrm{m}_1^* \mathrm{m}_3^*}\,(1 + \tau_0 p\,\Xi)\,(1 + \tau_1 p) + \frac{\varepsilon p \xi^2}{\mathrm{m}_3^*}\,(1 + \tau_0 p\,\Xi)\,(1 + \tau_1 p),$$

with

$$\mathrm{m}_1^* = \frac{\lambda + 2\mu}{\rho \mathrm{C}_1^2}, \qquad \mathrm{m}_3^* = \frac{\mu}{\rho \mathrm{C}_1^2}$$

i. For L-S theory : Taking $\tau_1 = 0$, $\Xi = 1$in expression given by (3.29)–(3.33) of particular case I, we obtain expressions for displacement components, temperature field and force stresses.

ii. For G-L theory : Taking $\Xi = 0$ in expressions given by (3.29)–(3.33)of particular case I, we obtain expressions for displacement components, temperature field and force stresses

iii. For G-N theory : Neglecting micropolarity effect i.e. ($\alpha = \beta = \gamma = K = j = 0$) in subcase III of case I, we get the expressions for displacement components, temperature field and force stresses are obtained in a thermoelastic medium by taking

$$1+\tau_1 p = 1, \quad 1+\tau_0 p = 4p, \quad 1+\tau_0 p\,\Xi = p, \quad \varepsilon = \frac{\varepsilon_1}{4}, \quad \omega^* = \frac{C_1}{h}, \quad \varepsilon_1 = \frac{\mathrm{T}_0 \nu^2}{\rho K^*}$$

in Eq. (3.29)–(3.44) as

$$\tilde{\mathrm{u}}_1(\xi, \mathrm{z}, \mathrm{p}) = -\left\{\mathrm{a}_1^0 \mathrm{q}_1 \mathrm{E}_4^0 e^{-\mathrm{q}_1^0 z} + \mathrm{a}_2^0 \mathrm{q}_2^0 \mathrm{E}_5^0 e^{-\mathrm{q}_2^0 z} + \mathrm{a}_3^0 \mathrm{q}_3^0 \mathrm{E}_6^0 e^{-\mathrm{q}_3^0 z}\right\},$$

$$\tilde{\mathrm{u}}_3(\xi, \mathrm{z}, \mathrm{p}) = \mathrm{b}_1^0 \mathrm{E}_4^0 e^{-\mathrm{q}_1^0 z} + \mathrm{b}_2^0 \mathrm{E}_5^0 e^{-\mathrm{q}_2^0 z} + \mathrm{b}_3^0 \mathrm{E}_6^0 e^{-\mathrm{q}_3^0 z},$$

$$\tilde{T}(\xi, \mathrm{z}, \mathrm{p}) = -\xi\left\{\mathrm{E}_4^0 e^{-\mathrm{q}_1^0 z} + \mathrm{E}_5^0 e^{-\mathrm{q}_2^0 z} + \mathrm{c}_3 \mathrm{E}_6^0 e^{-\mathrm{q}_3^0 z}\right\},$$

$$\tilde{\mathrm{t}}_{31}(\xi, \mathrm{z}, \mathrm{p}) = \mathrm{m}_7\{\left(\mathrm{a}_1^0 \mathrm{q}_1^{02} + \iota\xi \mathrm{b}_1^0\right)\mathrm{E}_4^0 e^{-\mathrm{q}_1^0 z} + \left(\mathrm{a}_2^0 \mathrm{q}_2^{02} + \iota\xi \mathrm{b}_2^0\right)\mathrm{E}_5^0 e^{-\mathrm{q}_2^0 z}$$
$$+ \left(\mathrm{a}_3^0 \mathrm{q}_3^{02} + \iota\xi \mathrm{b}_3^0\right)\mathrm{E}_6^0 e^{-\mathrm{q}_3^0 z}\},$$

$$\tilde{\mathrm{t}}_{33}(\xi, \mathrm{z}, \mathrm{p}) = -[\left\{\mathrm{m}_1^0 \mathrm{q}_1^0 \mathrm{b}_1^0 + \iota\xi \mathrm{m}_8 \mathrm{a}_1^0 \mathrm{q}_1^0 - \xi\,(1 + \tau_1 p)\right\}\mathrm{E}_4^0 e^{-\mathrm{q}_1^0 z}$$
$$+ \left\{\mathrm{m}_1^0 \mathrm{q}_2^0 \mathrm{b}_2^0 + \iota\xi \mathrm{m}_8 \mathrm{a}_2^0 \mathrm{q}_2^0 - \xi\,(1 + \tau_1 p)\right\}\mathrm{E}_5^0 e^{-\mathrm{q}_2^0 z}$$

$$+\left\{m_1^0 q_3^0 b_3^0 + \iota \xi m_8 a_3^0 q_3^0 - \xi\,(1+\tau_1 p)\right\} E_6^0 e^{-q_3^0 z}],$$

where

$$E_1^0 = E_4^0 = \frac{F_0\left(a_3^0 - a_2^0\right)}{2q_1^0 \Delta_1^0},$$
$$E_2^0 = E_5^0 = \frac{F_0\left(a_1^0 - a_3^0\right)}{2q_2^0 \Delta_1^0},$$
$$E_3^0 = E_6^0 = \frac{F_0\left(a_2^0 - a_1^0\right)}{2q_3^0 \Delta_1^0},$$
$$\Delta_1^0 = m_1^0\left\{\left(a_2^0 b_3^0 - a_3^0 b_2^0\right) + \left(a_3^0 b_1^0 - a_1^0 b_3^0\right) + \left(a_1^0 b_2 - a_2^0 b_1^0\right)\right\},$$
$$a_s^0 = \frac{-\xi}{m_3^* \Delta_s^0}[m_2\left\{\xi^2 + p\,(1+\tau_0 p) - q_s^{02}\right\} + \varepsilon_1 p^2,$$
$$b_s^* = \frac{\iota}{m_3^* \Delta_s^0}[\left\{\varepsilon_1 p^2 \xi^2 + \left(\xi^2 + 4p^2 - q_s^{02}\right)\right\}\left(m_1^* \xi^2 + p^2 - m_3^* q_s^{02}\right)\}$$
$$\Delta_s^0 = \frac{\varepsilon_1 p^2 q_s^0}{\xi m_3^*}\left\{\left(m_1^* \xi^2 + p^2 - m_3^* q_s^0 2\right) - \xi^2 m_2\right\}\ , \varepsilon_1 = 4\varepsilon,$$

and $\pm q_s^0$ $(s = 1, 2, 3)$ equation

$$q^{06} - \sigma_1^0 q^{0^4} + \sigma_2^0 q^{02} - \sigma_3^0 = 0,$$
$$\sigma_1^0 = \left(\frac{m_1^* \xi^2 + p^2}{m_3^*}\right) + \left(\frac{m_3^* \xi^2 + p^2}{m_1^*}\right) + \left(\xi^2 + 4p^2\right) + \frac{\varepsilon_1 p^2}{m_1^*} - \frac{\xi^2 m_2^2}{m_1^* m_3^*},$$
$$\sigma_2^0 = \left(\tfrac{m_1^* \xi^2 + p^2}{m_3^*}\right) + \left(\tfrac{m_3^* \xi^2 + p^2}{m_1^*}\right) + \left(\tfrac{m_3^* \xi^2 + p^2}{m_1^*}\right)\left(\xi^2 + 4p^2\right)$$
$$+ \left(\xi^2 + 4p^2\right)\left(\tfrac{m_1^* \xi^2 + p^2}{m_3^*}\right) + \tfrac{\varepsilon_1 p^2}{m_1^*}\left(\tfrac{m_1^* \xi^2 + p^2}{m_3^*}\right)$$
$$- \tfrac{\xi^2 m_2^2}{m_1^* m_3^*}\left(\xi^2 + 4p^2\right) - \tfrac{2\varepsilon_1 p^2 \xi^2 m_2}{m_1^* m_3^*} + \tfrac{\varepsilon_1 p^2 \xi^2}{m_3^*},$$

where $\tilde{f}_e$ and $\tilde{f}_0$ are even and odd parts of the functions $\tilde{f}(\xi, z, p)$ respectively. Thus, expression (37.1) gives us the Laplace transform $\bar{f}(x, z, p)$ of the function f (x, z, t). Following [11], the Laplace transform function $\bar{f}(x, z, p)$ can be inverted to f (x, z, t) .

$$\sigma_3^0 = \left(\tfrac{m_1^* \xi^2 + p^2}{m_3^*}\right)\left(\tfrac{m_3^* \xi^2 + p^2}{m_1^*}\right)\left(\xi^2 + 4p^2\right) + \tfrac{\varepsilon_1 p^2 \xi^2}{m_3^*}\left(\tfrac{m_3^* \xi^2 + p^2}{m_1^*}\right),$$

Thus, the expressions given by equations (3.5)–(3.12) with the help of (3.13)–(3.16) and (3.17) represent the solution of plane strain problem under consideration in the transformed domain using eigen value approach.

33.4 Inversion of the Transforms

To obtain the solution of the problem in the physical domain, we must invert the transforms for three theories that is L-S, G-L and G-N. These expressions are functions of z, the parameters of Laplace and Fourier transforms p and ξ respectively and hence are of the form $\bar{f}(x, z, p)$. To get the function $f(x, z, t)$in the physical domain, first we invert the Fourier transform using

$$\bar{f}(x, z, p) = \int_{-\infty}^{\infty} \exp(i\xi x)\tilde{f}(\xi, z, p)d\xi = \frac{1}{\pi}\int_{0}^{\infty} \left\{\cos(\xi x)\tilde{f}_e + i\sin(\xi x)\tilde{f}_0\right\} d\xi$$

The last step in the inversion process is to evaluate the integral in Eq.(37.1). This was done using Romberg's integration with adaptive step size. This method uses the results from successive refinements of the extended trapezoidal rule followed by extrapolation of the results to the limit when the step size tends to zero. The details can be found in [20].

33.5 Numerical Results and Discussion

Following [5], we take the following values of relevant parameters for the case of Magnesium crystal as

$\rho = 1.74\,\text{gm/cm}^3$,	$j = 0.2 \times 10^{-15}\,\text{cm}^2$,	$\lambda = 9.4 \times 10^{11}\,\text{dyne/cm}^2$,
$\mu = 4.0 \times 10^{11}\,\text{dyne}$,	$K = 1.0 \times 10^{11}\,\text{dyne/cm}^2$,	$C^* = 0.23\,\text{Call/gm}^0 C$,
$\gamma = 0.779 \times 10^{-4}\,\text{dyne}$,	$\varepsilon = 0.073$.	$K^* = 0.6 \times 10^{-2}\,\text{cal/cmsec}$,
$T_0 = 23°\text{C}$,	$\tau_0 = 6.131 \times 10^{-13}\,\text{sec}$,	$\tau_1 = 8.765 \times 10^{-13}\,\text{sec}$
$h = 1\,\text{cm}$,	$z = 1$	

33.6 Discussion

The variations of normal displacement U_3 with distance x for three different theories (L-S, G-L and G-N) in both media after multiplying the original values for G-N theory in MTE medium by 10 are shown in Fig. 33.2 The values of normal displacement due to microrotation effect are less in MTE medium in comparison to TE medium in the $0 \le x \le 0.5$for all three theories, whereas the values of U_3 oscillate as x increases further in the rest of the range for both media. It is also evident that normal displacement decreases for both media for L-S and G-L theories, increases gradually in MTE medium for G-N theory and oscillate in TE medium for G-N theory.

The values of normal force stress T_{33}in magnitude are more for three different theories in MTE medium in comparison to TE medium. It is also noticed that the values of normal force stress oscillate for L-S and G-L theories in MTE and TE media. The values of normal force stress also oscillate for G-N theory in TE medium, whereas

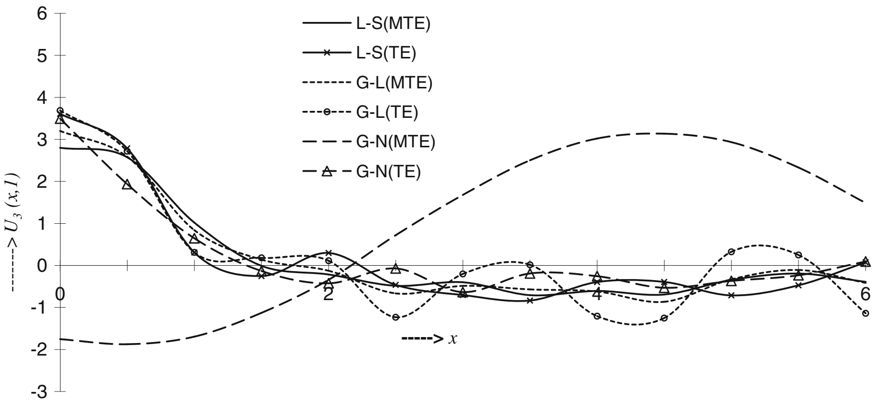


Fig. 33.2 Variation of normal displacement $U_3(x, l)$

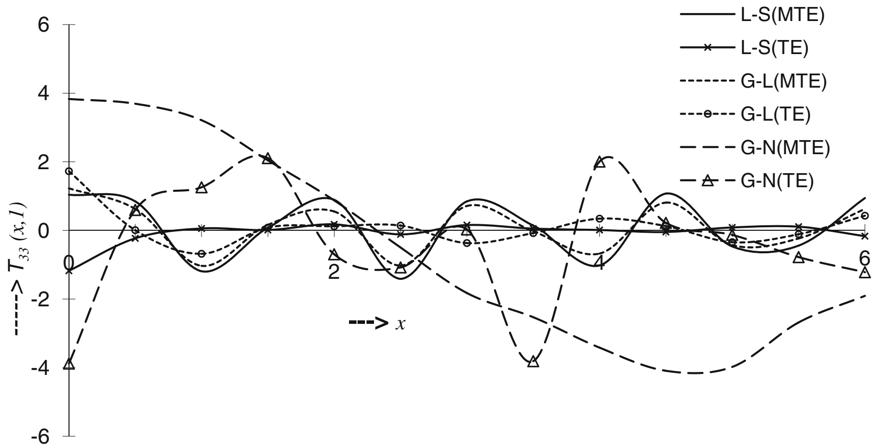


Fig. 33.3 Variation of normal force stress $T_{33}(x, l)$

these decrease gradually with increasing value of x in MTE medium. These variations of normal force stress have been shown in Fig. 33.3 after dividing the original values by 10 in case of G-N theory in MTE medium.

Figure 33.4 depicts the variations of tangential couple stress M_{32} for three different theories in MTE medium after dividing the original values for G-N theory by 10. The behaviour of tangential couple stress is oscillatory for three theories. It is noticed that the value of tangential couple stress for G-N theory are large in comparison to L-S and G-L theories in the range $0 \leq x \leq 2.5$ and the values are small for the rest of the range.

The range of values of temperature field in magnitude is large in case of three theories in MTE medium in comparison to TE medium. It is also observed that temperature field oscillate in TE medium for three different theories but in MTE medium for L-S and G-L theories, the temperature field oscillate. The values of temperature field for

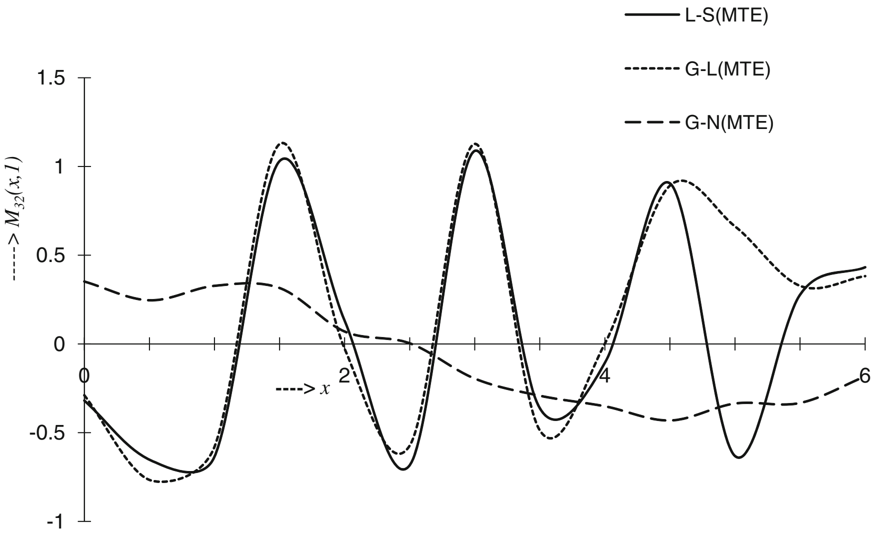


Fig. 33.4 Variation of tangential couple stress $M_{32}(x, l)$

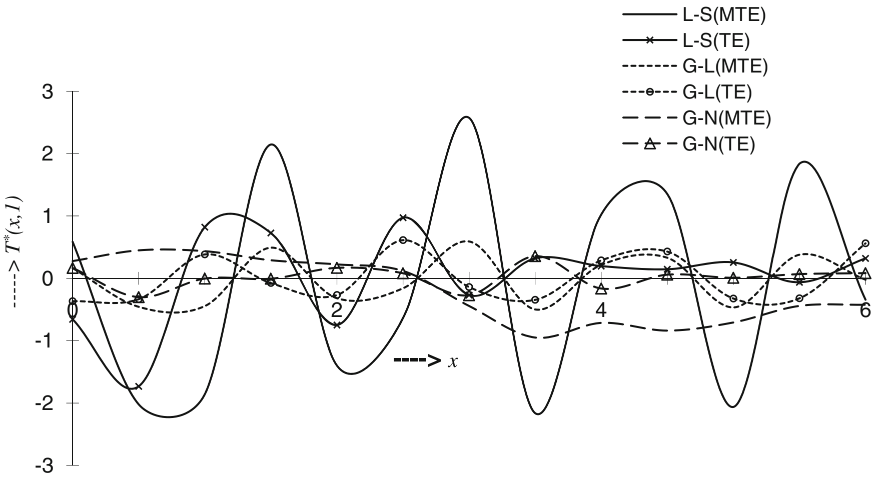


Fig. 33.5 variation of temprature field $T^*(x, l)$

G-N theory decrease gradually with increasing value of x in MTE medium. These variations shown in Fig. 33.5 after multiplying the original values in case of L-S and G-N theories by 10^2 and 10^2 respectively in MTE medium; the original values in case of G-N theory (TE medium) and also magnified by multiplying 10^2.

33.7 Conclusion

From the above numerical results, we conclude that micropolarity has a significant effect on normal displacement, normal force stress and temperature field mechanical-source for three theories.Mcropolar effect is more appreciable for normal displacement and temperature field in, comparison to normal force stress. Application of the present paper may also be found in the field of steel and oil industries. The present Problem is also useful in the field of geomechanics, where, the interest is about the various phenomenon occurring in the earthquakes and measuring of displacements, stresses and temperature field due to the presence of certain sources.

33.8 Nomenclature

λ, μ = Lame's constants
$\alpha, \beta, \gamma, \mathrm{K}$ = Micropolar material constants
$\alpha_0, \lambda_0, \lambda_1$ = Material constants due to the presence of stretch.
$\lambda_I, \mu_I, \mathrm{K}_I, \alpha_I, \nu, \gamma_I, \alpha_{0I}, \lambda_{0I}, \lambda_{1I}$ = Microstretch viscoelastic constants
ρ = Density j = Micro-inertia $\mathbf{u}$ = Displacement vector $\boldsymbol{\phi}$ = Microrotation vector
$\boldsymbol{\phi}^*$ = Scalar microstretch
t_{ij} = Force stress tensor
m_{ij} = Couple stress tensor
λ_l = Microstress tensor
δ_{ij} = Kronecker delta
ε_{ijr} = Alternating tensor
Δ = Gradient operator
ι = Iota
And dot denotes the partial derivative w.r.t. time.

References

1. Boschi E, Jesan D (1993) A generalized theory of linear micropolar thermoelasticity. Meccanica 8(3):154–157
2. Bakshi R, Bera RK, Debnath L (2004) Eigen value approach to study the effect of rotation and relaxation time in two dimensional problems of generalized thermoelastic. Int J Engg Sci 42:1573–1586
3. Cracium L (1990) On the uniqueness of the solution of some boundary value problems of some boundary value problems of the linear micropolar thermoelasticity. Bull Int Politech Lasi Sect 36:57–62
4. Eringen AC (1970) Foundations of micropolar thermoelasticity, vol 23. Springer, International Center for Mechanical Science Courses and Lecturer
5. Eringen AC (1984) Plane waves in a non-local micropolar elasticity. Int J Engg Sci 22:1113–1121
6. Eringen AC (1968) Theory of micropolar elasticity. In: Leibowitz H (ed) Fracture vol II. Academic Press, New York (chapter-7).

7. Green AE, Lindsay KA (1972) Thermoelasticity. J Elast 2:1–5
8. Green AE, Naghdi PM (1977) On thermodynamics and the nature of the second law. Proc Ray Soc Lond A 357:253–270
9. Green AE, Naghdi PM (1991) A Re-examination of the Basic Postulate of Thermomechanics. Proc Ray Soc Lond A 432:171–194
10. Green AE, Naghdi PM (1993) Thermoelastic without energy dissipation. J Elast 31:189–208
11. Honig G, Hirdes U (1984) A method for the numerical inversion of the Laplace transforms. J Comp Appl Math 10:113–132
12. Lord HW, Shulman Y (1967) A generalized dynamical theory of thermoelasticity. J Mech Phys Solids 15:299–306
13. Kumar R, ChadhaT K (1986) On torsional loading in an axismmetricmicropolar elastic medium. Proc Indian Acad Sci (Maths Sci) 95:109–120
14. Kumar R, Deswal S (2001) Mechanical and thermal sources in the micropolargeneralized thermoelastic medium. J Sound Vib 239:467–488
15. Kumar R, Rani L (2004) Deformation due to mechanical and thermal sources in the generalized orthorhombic thermoelastic material. Sadhana 29:429–448
16. Kumar R, RaniL (2005) Deformation due to mechanical and thermal sources in the generalized half-space with void. J Therm Stress 28:123–143
17. Mahalanabis RK, Manna J (1997) Eigen value approach to the problems of linear micropolarthermoelasticity. J Indian Acad Math 19:67–86
18. Marin M, Lupu M (1998) On harmonic vibrations in thermoelasticity of micropolar bodies. J Vib Control 4:507–518
19. Nowacki W (1966) Couple stresses in the theory of thermoelasticity. In: Proceedings of the ITUAM symposia, Springer, pp 259–278.
20. Press WH, Teukolsky SA, Vellerling WT, Flannery BP (1986) Numerical recipes. Cambridge University Press, Cambridge
21. Passarella F (1996) Some results in micropolarthermoelasticity. Mech Res Commun 23:349–359
22. Sturnin DV (2001) On characteristics times in generalized thermoelasticity. J Appl Math 68:816–817
23. Sharma JN, Chand D (1992) On the axisymmetric and plane strain problems of generalized thermoelasticity. Int J Eng Sci 33:223–230

Part VIII
Mathematical Physics

Chapter 34
Relativistic Classical and Quantum Mechanics: Clock Synchronization and Bound States, Center of Mass and Particle Worldlines, Localization Problems and Entanglement

Luca Lusanna

Abstract Parametrized Minkowski theories allow to describe isolated systems in global non-inertial frames in Minkowski space-time (defined as M/oller-admissible $3+1$ splittings with clock synchronization and radar 4-coordinates) , with the transitions among frames described as gauge transformations. The restriction to the intrinsic inertial rest frame allows to formulate a new relativistic classical mechanics for N-particle systems compatible with relativistic bound states. There is a complete control on the relativistic collective variables (Newton-Wigner center of mass, Fokker-Pryce center of inertia, M/oller center of energy) and on the realization of the Poincare' algebra (with the explicit form of the interaction-dependent Lorentz boosts). The particle world-lines are found to correspond to the ones of predictive mechanics and localization problems are clarified. The model can be consistently quantized avoiding the instantaneous spreading of the center-of-mass wave packets (Hegerfeldt theorem), because the non-local non-covariant center of mass is a non-measurable quantity. The basic difference with non-relativistic quantum mechanics is that the composite N-particle system cannot be represented as the tensor product of single particle Hilbert spaces, but only as the tensor product of the center-of-mass Hilbert space with the one of relative motions. This spatial non-separability (due to the Lorentz signature of space-time) makes relativistic entanglement much more involved than the non-relativistic one. Some final remarks on the emergence of classicality from quantum theory are done.

34.1 Introduction

Predictability in classical and quantum physics is possible only if the relevant partial differential equations have a well posed Cauchy problem and the existence and unicity theorem for their solutions applies. A pre-requisite is the existence of a well defined

L. Lusanna (✉)
Sezione INFN di Firenze,Via Sansone 1, 50019 Sesto Fiorentino, FI, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_34,

3-space (i.e. a clock synchronization convention) supporting the Cauchy data. In Galilei space-time there is no problem: time and Euclidean 3-space are absolute.

Instead there is no intrinsic notion of 3-space, simultaneity, 1-way velocity of light (two distant clock are involved) in special relativity (SR): in the absolute Minkowski space-time, only the *conformal structure* (the light-cone) is intrinsically given as the locus of incoming and outgoing radiation. The light postulate says that the 2-way (only one clock is involved) velocity of light c is isotropic and constant. Its codified value replaces the rods (i.e. the standard of length) in modern metrology, where an atomic clock gives the standard of time and a conventional reference frame centered on a given observer is chosen as a standard of space-time (GPS is an example of such a standard).

The standard way out from the problem of 3-space is to choose the Euclidean 3-space of an inertial frame centered on an inertial observer and then use the kinematical Poincaré group to connect different inertial frames. This is done by means of *Einstein convention for the synchronization of clocks*: the inertial observer A sends a ray of light at x_i^o towards the (in general accelerated) observer B; the ray is reflected towards A at a point P of B world-line and then reabsorbed by A at x_f^o; by convention P is synchronous with the mid-point between emission and absorption on A's world-line, i.e. $x_P^o = x_i^o + \frac{1}{2}\,(x_f^o - x_i^o) = \frac{1}{2}\,(x_i^o + x_f^o)$. This convention selects the Euclidean instantaneous 3-spaces $x^o = ct = const.$ of the inertial frames centered on A. Only in this case the one-way velocity of light between A and B coincides with the two-way one, c. However, if the observer A is accelerated, the convention breaks down, because if *only* the world-line of the accelerated observer A (the *1+3 point of view*) is given, then the only way for defining instantaneous 3-spaces is to identify them with the Euclidean tangent planes orthogonal to the 4-velocity of the observer (the local rest frames). But these planes (they are tangent spaces not 3-spaces!) will intersect each other at a distance from A's world-line of the order of the acceleration lengths of A, so that all the either linearly or rotationally accelerated frames, centered on accelerated observers, based either on Fermi coordinates or on rotating ones, will develop *coordinate singularities*. Therefore their approximated notion of instantaneous 3-spaces cannot be used for a well-posed Cauchy problem for Maxwell equations.

Therefore in Refs. [9–11] a general theory of non-inertial frames in Minkowski space-time was developed. As shown in Refs. [10, 14, 11] this formulation allowed to develop a Lagrangian description (*parametrized Minkowski theories*) of isolated systems (particles, strings, fluids, fields) in which the transition from a non-inertial frame to another one is formalized as a gauge transformation. Therefore the inertial effects only modify the appearances of phenomena but not the physics.

Then the restriction to inertial frames allowed to define the Poincaré generators starting from the energy-momentum tensor. The intrinsic inertial rest frame of every isolated system allowed the definition of the *rest-frame instant form of dynamics* where it is possible to study the problem of the separation of the collective relativistic variable of an isolated system from the relative ones living in the so called Wigner 3-spaces in accord with the theory of relativistic bound states [2, 4, 13, 5]. The ordinary world-lines of the particles can then be reconstructed [9].

The rest-frame instant form allows to give a formulation of *relativistic quantum mechanics* (RQM) [6] which takes into account of the problems of relativistic causality and relativistic localization implied by the Lorentz signature of Minkowski space-time. Also a definition of relativistic entanglement can be given: the Lorentz signature implies features of *non-locality and spatial non-separability* absent in the non-relativistic formulation of entanglement.

Elsewhere we extend the definition of non-inertial frames to general relativity (GR) in asymptotically Minkowskian space-times. In GR there is no absolute notion since also space-time becomes dynamical (with the metric structure satisfying Einstein's equations): however in this class of space-times it is possible to make a Hamiltonian formulation and to separate the inertial degrees of freedom of the gravitational field from the tidal ones (the gravitational waves of the linearized theory).

34.2 Non-inertial Frames in Minkowski Space-Time and Radar 4-Coordinates

A metrology-oriented description of non-inertial frames in SR can be done with the $3+1$ *point of view* and the use of observer-dependent Lorentz scalar *radar 4-coordinates*. Let us give the world-line $x^\mu(\tau)$ of an arbitrary time-like observer carrying a standard atomic clock: τ is an arbitrary monotonically increasing function of the proper time of this clock. Then we give an admissible $3+1$ splitting of Minkowski space-time, namely a nice foliation with space-like instantaneous 3-spaces Σ_τ: it is the mathematical idealization of a protocol for clock synchronization. All the clocks in Σ_τ sign the same time of the atomic clock of the observer: it is the non-factual idealization required by the Cauchy problem generalizing the existing protocols for building coordinate system inside the future light-cone of a time-like observer. On each 3-space Σ_τ we choose curvilinear 3-coordinates σ^r having the observer as origin. These are the *radar 4-coordinates* $\sigma^A = (\tau; \sigma^r)$.

If $x^\mu \mapsto \sigma^A(x)$ is the coordinate transformation from the Cartesian 4-coordinates x^μ of a reference inertial observer to radar coordinates, its inverse $\sigma^A \mapsto x^\mu = z^\mu(\tau, \sigma^r)$ defines the *embedding* functions $z^\mu(\tau, \sigma^r)$ describing the 3-spaces Σ_τ as embedded 3-manifold into Minkowski space-time. The induced 4-metric on Σ_τ is the following functional of the embedding ${}^4g_{AB}(\tau, \sigma^r) = [z^\mu_A\, \eta_{\mu\nu}\, z^\nu_B](\tau, \sigma^r)$, where $z^\mu_A = \partial z^\mu / \partial \sigma^A$ and ${}^4\eta_{\mu\nu} = \epsilon\,(+ - --)$ is the flat metric ($\epsilon = \pm 1$ according to either the particle physics $\epsilon = 1$ or the general relativity $\epsilon = -1$ convention). While the 4-vectors $z^\mu_r(\tau, \sigma^u)$ are tangent to Σ_τ, so that the unit normal $l^\mu(\tau, \sigma^u)$ is proportional to $\epsilon^\mu{}_{\alpha\beta\gamma}\, [z^\alpha_1\, z^\beta_2\, z^\gamma_3](\tau, \sigma^u)$, we have $z^\mu_\tau(\tau, \sigma^r) = [N\, l^\mu + N^r\, z^\mu_r](\tau, \sigma^r)$ with $N(\tau, \sigma^r) = \epsilon\, [z^\mu_\tau\, l_\mu](\tau, \sigma^r)$ and $N_r(\tau, \sigma^r) = -\epsilon\, g_{\tau r}(\tau, \sigma^r)$ being the lapse and shift functions.

The foliation is nice and admissible if it satisfies the conditions:

(1) $N(\tau, \sigma^r) > 0$ in every point of Σ_τ (the 3-spaces never intersect, avoiding the coordinate singularity of Fermi coordinates);

(2) $\epsilon\, {}^4g_{\tau\tau}(\tau, \sigma^r) > 0$, so to avoid the coordinate singularity of the rotating disk, and with the positive-definite 3-metric ${}^3g_{rs}(\tau, \sigma^u) = -\epsilon\, {}^4g_{rs}(\tau, \sigma^u)$ having three positive eigenvalues (these are the Møller conditions);
(3) all the 3-spaces Σ_τ must tend to the same space-like hyper-plane at spatial infinity (so that there are always asymptotic inertial observers to be identified with the fixed stars).

These conditions imply that *global rigid rotations are forbidden in relativistic theories*. In Refs. [10, 11] there is the expression of the admissible embedding corresponding to a 3 + 1 splitting of Minkowski space-time with parallel space-like hyper-planes (not equally spaced due to a linear acceleration) carrying differentially rotating 3-coordinates without the coordinate singularity of the rotating disk. It is the first consistent global non-inertial frame of this type.

Each admissible 3 + 1 splitting of space-time allows to define two associated congruences of time-like observers: (a) the Eulerian observers with the unit normal to Σ_τ as 4-velocity; (b) the non surface forming observers with 4-velocity proportional to $z^\mu_\tau(\tau, \sigma^r)$.

Therefore starting from the four independent embedding functions $z^\mu(\tau, \sigma^r)$ we obtain the ten components ${}^4g_{AB}(\tau, \sigma^u)$ of the 4-metric: they play the role of the *inertial potentials* generating the *relativistic apparent forces* in the non-inertial frame (the usual Newtonian inertial potentials can be recovered by doing the non-relativistic limit). The extrinsic curvature tensor ${}^3K_{rs}(\tau, \sigma^u) = [\frac{1}{2N}\,(N_{r|s} + N_{s|r} - \partial_\tau\, {}^3g_{rs})](\tau, \sigma^u)$, describing the *shape* of the instantaneous 3-spaces as embedded 3-sub-manifolds of Minkowski space-time, is a secondary inertial potential functional of the inertial potentials ${}^4g_{AB}$.

34.3 Classical Relativistic Isolated Systems

In these global non-inertial frames of Minkowski space-time it is possible to describe isolated systems (particles, strings, fields, fluids) admitting a Lagrangian formulation by means of *parametrized Minkowski theories*. The matter variables are replaced with new ones knowing the clock synchronization convention defining the 3-spaces Σ_τ. For instance a Klein-Gordon field $\tilde{\phi}(x)$ will be replaced with $\phi(\tau, \sigma^r) = \tilde{\phi}(z(\tau, \sigma^r))$; the same for every other field. Instead for a relativistic particle with world-line $x^\mu(\tau)$ we must make a choice of its energy sign: then the positive- (or negative-) energy particle will be described by 3-coordinates $\eta^{r(\tau)}$ defined by the intersection of its world-line with Σ_τ: $x^\mu(\tau) = z^\mu(\tau, \eta^{r(\tau)})$. Differently from all the previous approaches to relativistic mechanics, the dynamical configuration variables are the 3-coordinates $\eta^{r(\tau)}$ and not the world-lines $x^\mu(\tau)$ (to rebuild them in an arbitrary frame we need the embedding defining that frame!).

Then the matter Lagrangian is coupled to an external gravitational field and the external 4-metric is replaced with the 4-metric $g_{AB}(\tau, \sigma^r)$ of an admissible 3+1 splitting of Minkowski space-time. With this procedure we get a Lagrangian depending on the given matter and on the embedding $z^\mu(\tau, \sigma^r)$, which is invariant under *frame-preserving diffeomorphisms*. As a consequence, there are four first-class constraints (an analogue of the super-Hamiltonian and super-momentum constraints of canonical gravity) implying that the embeddings $z^\mu(\tau, \sigma^r)$ are *gauge variables*, so that *all the admissible non-inertial or inertial frames are gauge equivalent*, namely physics does *not* depend on the clock synchronization convention and on the choice of the 3-coordinates σ^r. Even if the gauge group is formed by the frame-preserving diffeomorphisms, the matter energy-momentum tensor allows the determination of the ten conserved Poincare' generators P^μ and $J^{\mu\nu}$ (assumed finite) of every configuration of the system (in non-inertial frames they are asymptotic generators at spatial infinity like the ADM ones in GR).

If we restrict ourselves to inertial frames, we can define the *inertial rest-frame instant form of dynamics for isolated systems* by choosing the $3 + 1$ splitting corresponding to the intrinsic inertial rest frame of the isolated system centered on an inertial observer: the instantaneous 3-spaces, named *Wigner 3-spaces* due to the fact that the 3-vectors inside them are Wigner spin-1 3-vectors, are orthogonal to the conserved 4-momentum P^μ of the configuration. The embedding corresponding to the inertial rest frame is $z^\mu(\tau, \vec{\sigma}) = Y^\mu(\tau) + \epsilon^\mu_r(\vec{h})\,\sigma^r$, where $Y^\mu(\tau)$ is the Fokker-Pryce center-of-inertia 4-vector, $\vec{h} = \vec{P}/\sqrt{\epsilon\, P^2}$ and $\epsilon^\mu{}_{A=\nu}(\vec{h})$ is the standard Wigner boost for time-like orbits sending $P^\mu = \sqrt{\epsilon\, P^2}\,(\sqrt{1 + \vec{h}^2}; \vec{h})$ to $(1; 0)$.

In Refs. [10, 11] there is also the definition of the admissible *non-inertial rest frames*, where P^μ is orthogonal to the asymptotic space-like hyper-planes to which the instantaneous 3-spaces tend at spatial infinity. This non-inertial family of $3 + 1$ splittings is the only one admitted by the asymptotically Minkowskian space-times without super-translations in GR. Finally in Ref. [1] there is the definition of *parametrized Galilei theories*, non relativistic limit of the parametrized Minkowski theories. Also the inertial and non-inertial frames in Galilei space-time are gauge equivalent in this formulation.

34.4 Relativistic Center of Mass, Particle World-Lines and Bound States

The framework of the inertial rest frame allowed the solution of the following old open problems:

(A) The classification [12] of the relativistic collective variables (the canonical non-covariant Newton-Wigner center of mass (or center of spin), the non-canonical covariant Fokker-Pryce center of inertia and the non-canonical non-covariant Møller center of energy), replacing the Newtonian center of mass (and all tending to it in the non-relativistic limit), that can be built only in terms of the Poincare' generators: they are *non measurable* quantities due to the *non-local* character of such generators (they know the whole 3-space Σ_τ). There is a Møller world-tube around the Fokker-Pryce 4-vector containing all the possible pseudo-world-lines of the other two, whose

Møller radius $\rho = |\vec{S}|/\sqrt{\epsilon P^2}$ ($\vec{S}$ is the rest angular momentum) is determined by the Poincaré Casimirs of the isolated system. This non-covariance world-tube is a non-local effect of Lorentz signature of the space-time absent in Euclidean spaces. The world-lines $x_i^\mu(\tau)$ of the particles are derived (interaction-dependent) quantities and in general they do not satisfy vanishing Poisson brackets [3]: already at the classical level a *non-commutative structure* emerges due to the Lorentz signature!

(B) The description of every isolated system as a decoupled (non-measurable) canonical non-covariant (Newton-Wigner) external center of mass (described by frozen Jacobi data) carrying a pole-dipole structure: the invariant mass and the rest spin of the system expressed in terms of suitable Wigner-covariant relative variables of the given isolated system inside the Wigner 3-spaces (after the elimination of the internal center of mass with well defined rest-frame conditions). The invariant mass is the effective Hamiltonian inside these 3-spaces.

(C) The formulation of classical relativistic atomic physics (the electro-magnetic field in the radiation gauge plus charged scalar particles with Grassmann-valued electric charges to make a ultraviolet and infrared regularization of the self-energies and with mutual Coulomb potentials) and the identification of the Darwin potential at the classical level by means of a canonical transformation transforming the original system in N charged particles interacting with Coulomb plus Darwin potentials and a free radiation field (absence of Haag's theorem at least at the classical level). Therefore the Coulomb plus Darwin potential is the description as a Cauchy problem of the interaction described by the one-photon exchange Feynman diagram of QED (all the radiative corrections and photon brehmstrahlung are deleted by the Grassmann regularization).

In Ref. [15] there are the references for all the interacting systems for which the explicit form of the interaction-dependent Lorentz boosts is known.

34.5 Relativistic Quantum Mechanics

In Ref. [6] there is a new formulation of *relativistic quantum mechanics* in inertial frames englobing all the known results about relativistic bound states due to the use of clock synchronization for the definition of the instantaneous 3-spaces, which implies the absence of relative times in their description.

In Galilei space-time non-relativistic quantum mechanics, where all the main results about entanglement are formulated, describes a composite system with two (or more) subsystems with a Hilbert space which is the tensor product of the Hilbert spaces of the subsystems: $H = H_1 \otimes H_2$. This type of spatial separability is named *the zeroth postulate* of quantum mechanics. However, when the two subsystems are mutually interacting, one makes a unitary transformation to the tensor product of the Hilbert space H_{com} describing the decoupled Newtonian center of mass of the two subsystems and of the Hilbert space H_{rel} of relative variables: $H = H_1 \otimes H_2 = H_{com} \otimes H_{rel}$. This allows to use the method of separation of variables to split the Schroedinger equation in two equations: one for the free motion of the center of mass and another, containing the interactions, for the relative variables (this equation describes both the bound and scattering states). A final unitary transformation of the Hamilton-Jacobi type allows to replace H_{com} with $H_{com,HJ}$, the Hilbert space in which the decoupled center of mass is frozen and described by non-evolving Jacobi data. Therefore we have $H = H_1 \otimes H_2 = H_{com} \otimes H_{rel} = H_{com,HJ} \otimes H_{rel}$.

While at the non-relativistic level these three descriptions are unitary equivalent, this is no more true in RQM. The non-local and non-covariant properties of the decoupled relativistic center of mass, described by the frozen Jacobi data $\vec{z}$ and $\vec{h} = \vec{P}/\sqrt{\epsilon P^2}$, imply that the only consistent relativistic quantization is based on the Hilbert space $H = H_{com,HJ} \otimes H_{rel}$ (in the non-relativistic limit it goes into the corresponding Galilean Hilbert space). We have $H \neq H_1 \otimes H_2$, because, already in the non-interacting case, in the tensor product of two quantum Klein-Gordon fields, $\phi_1(x_1)$ and $\phi_2(x_2)$, most of the states correspond to configurations in Minkowski space-time in which one particle may be present in the absolute future of the other particle. This is due to the fact that the two times x_1^o and x_2^o are totally uncorrelated, or in other words there is no notion of instantaneous 3-space (clock synchronization convention). Also the scalar products in the two formulations are completely different.

We have also $H \neq H_{com} \otimes H_{rel}$, because if instead of $\vec{z} = Mc\,\vec{x}_{NW}(0)$ we use the evolving (non-local and non-covariant) Newton-Wigner position operator $\vec{x}_{NW}(\tau)$, then we get a violation of relativistic causality because the center-of-mass wave packets spread instantaneously as shown by the Hegerfeldt theorem.

Therefore the only consistent Hilbert space is $H = H_{com,HJ} \otimes H_{rel}$. The main complication is the definition of H_{rel}, because we must take into account the three pairs of (interaction-dependent) second-class constraints eliminating the internal 3-center of mass inside the Wigner 3-spaces. When we are not able to make the elimination at the classical level and formulate the dynamics only in terms of Wigner-covariant relative variables, we have to quantize the particle Wigner-covariant 3-variables η_i^r, κ_{ir} and then to define the physical Hilbert space by adding the quantum version of the constraints a la Gupta-Bleuler.

Relativistic quantum mechanics in rotating non-inertial frames by using a multi-temporal quantization scheme is defined in Ref. [8]. In it *the inertial gauge variables are not quantized but remain c-numbers*; the known results in atomic and nuclear physics are reproduced.

34.6 Relativistic Entanglement and Localization Problems

Since we have that the Hilbert space $H = H_{com,HJ} \otimes H_{rel}$ is not unitarily equivalent to the one $H_1 \otimes H_2 \otimes \ldots$, where H_i are the Hilbert spaces of the individual particles, at the relativistic level the zeroth postulate of non-relativistic quantum mechanics does not hold (the necessity to use H_{rel} implies a type of *weak form of relationism* different from the formulations connected to the Mach principle). Since the Hilbert space of composite systems is not the tensor product of the Hilbert spaces of the sub-systems, we need a formulation of *relativistic entanglement* taking into account this spatial non-separability and the non-locality of the center of mass, both coming from the Poincare' group, i.e. from the Lorentz signature of space-time.

Since the center of mass is decoupled, its *non-covariance* is irrelevant. However its *non-locality* is a source of open problems: do we have to quantize it (in the preferred momentum basis implied by the quantum Poincaré algebra) and if yes is it meaningful to consider center-of-mass wave packets or we must add some super-selection rule? The wave function of the non-local center of mass is a kind of wave function of the 3-universe: who will observe it?

As a consequence in SR there are open problems on which type of *relativistic localization* is possible. There are strong indications that the Newton-Wigner position operator *cannot be self-adjoint but only symmetric* with the implication of a bad localization of relativistic particles. The existing problems with relativistic position operators (like the Newton-Wigner one), deriving from the Lorentz signature of Minkowski space-time which forbids the existence of a unique collective variable with all the properties of Newton center of mass , point towards a *non-measurability of absolute positions* (but not of the relative variables needed to describe the spectra of bound states). This type of *un-sharpness* should be induced also in non-relativistic quantum mechanics: in atomic physics the crucial electro-magnetic effects are of order $1/c$. Experiments in atomic and molecular physics are beginning to explore these localization problems, in frameworks dominated by a *Newtonian classical intuition* and taking into account the experimental quantum limits for atom localization.

Therefore SR introduces a *kinematical non-locality* and a *kinematical spatial non-separability*, which reduce the relevance of *quantum non-locality* in the study of the foundational problems of quantum mechanics. Relativistic entanglement will have to be reformulated in terms of relative variables also at the non-relativistic level. Therefore the control of Poincare' kinematics will force to reformulate the experiments connected with Bell inequalities and teleportation in terms of the relative variables of isolated systems containing: (a) the observers with their measuring apparatus (Alice and Bob as macroscopic quasi-classical objects); (b) the particles of the protocol (but now the ray of light, the "photons" carrying the polarization, move along null geodesics); (c) the environment (macroscopic either quantum or quasi-classical object).

34.7 Open Problems

The main open problem in SR is the quantization of *fields* in non-inertial frames due to the no-go theorem of Ref. [16] showing the existence of obstructions to the unitary evolution of a massive quantum Klein-Gordon field between two space-like surfaces of Minkowski space-time. Its solution, i.e. the identification of all the $3+1$ splittings allowing unitary evolution, will be a prerequisite to any attempt to quantize canonical gravity taking into account the equivalence principle (global inertial frames do not exist!). Moreover entanglement in non-inertial frames without Rindler observers is still to be formulated.

References

1. Alba D (2006) Quantum mechanics in noninertial frames with a multitemporal quantization scheme: II. Nonrelativistic particles. Int J Mod Phys A21:3917 (hep-th/0504060).
2. Alba D, Crater HW, Lusanna L (2001) The semiclassical relativistic darwin potential for spinning particles in the rest frame instant form: two-body bound states with spin 1/2 constituents. Int J Mod Phys A16:3365 (arXiv hep-th/0103109).
3. Alba D, Crater HW, Lusanna L (2007) Hamiltonian relativistic two-body problem: center of mass and orbit reconstruction. J Phys A40:9585 (arXiv gr-qc/0610200).
4. Alba D, Crater HW, Lusanna L (2010), Towards relativistic atom physics. I. The rest-frame instant form of dynamics and a canonical transformation for a system of charged particles plus the electro-magnetic field. Canad, J Phys 88:379 (arXiv 0806.2383).
5. Alba D, Crater HW, Lusanna L (2010), Towards relativistic atom physics. II. Collective and relative relativistic variables for a system of charged particles plus the electro-magnetic field. Canad, J Phys 88:425 (arXiv 0811.0715).
6. Alba D, Crater HW, Lusanna L (2011) Relativistic quantum mechanics and relativistic entanglement in the rest-frame instant form of dynamics. J Math Phys 52:062301 (arXiv 0907.1816).
7. Alba D, Lusanna L (1998) The Lienard-Wiechert potential of charged scalar particles and their relation to scalar electrodynamics in the rest-frame instant form. Int J Mod Phys A13:2791 (arXiv hep-th/0708156).
8. Alba D, Lusanna L (2006) Quantum mechanics in noninertial frames with a multitemporal quantization scheme: I. Relativistic particles. Int J Mod Phys A21:2781 (arXiv hep-th/0502194).
9. Alba D, Lusanna L (2007) Generalized radar 4-coordinates and equal-time cauchy surfaces for arbitrary accelerated observers. Int J Mod Phys D16:1149 (arXiv hep-th/0502194).
10. Alba D, Lusanna L (2010) Charged particles and the electro-magnetic field in non-inertial frames: I. Admissible 3+1 splittings of Minkowski spacetime and the non-inertial rest frames. Int J Geom Methods Phys 7:33 (arXiv 0908.0213).
11. Alba D, Lusanna L (2010) Charged particles and the electro-magnetic field in non-inertial frames: II. Applications: rotating frames, Sagnac effect, Faraday rotation, wrap-up effect. Int J Geom Methods Phys 7:185 (arXiv 0908.0215).
12. Alba D, Lusanna L, Pauri M (2002) Centers of mass and rotational kinematics for the relativistic N-body problem in the rest-frame instant form. J Math Phys 43:1677–1727 (arXiv hep-th/0102087)
13. Crater HW, Lusanna L (2001) The rest-frame Darwin potential from the Lienard-Wiechert solution in the radiation gauge. Ann Phys (NY) 289:87
14. Lusanna L (1997) The N- and 1-time classical descriptions of N-body relativistic kinematics and the electromagnetic interaction. Int J Mod Phys A 12:645
15. Lusanna L (2011) Canonical gravity and relativistic metrology: from clock synchronization to dark matter as a relativistic inertial effect (arXiv 1108.3224).
16. Torre CG, Varadarajan M (1999) Functional evolution of free quantum fields. Class Quantum Grav 16:2651

Chapter 35
A New Computational Approach to Infinity for Modelling Physical Phenomena

Yaroslav D. Sergeyev

Abstract A new computational methodology for computations with infinite and infinitesimal quantities is described. It is based on the principle 'The part is less than the whole' introduced by Ancient Greeks and observed in the physical world. It is applied to all sets and processes (finite and infinite) and all numbers (finite, infinite, and infinitesimal). It is shown that it becomes possible to work with all of them in a unique framework (different from non-standard analysis) allowing one to easily manage mathematical situations that traditionally create difficulties (divergences of various kind, indeterminate forms, etc.) and to construct mathematical models of physical phenomena of a new type.

35.1 Introduction

Physicists use (and create) Mathematics as an instrument that allows them to construct and to study mathematical models describing the physical world. As a consequence, each concrete mathematical language used for this purpose reflects in some sense the opinion of physicists that they have about the world at that concrete historical period. This opinion is based on a sum of knowledge obtained from experiments that, in their turn, are made by scientific instruments having their accuracy that bounds our possibility of the observation of physical phenomena. As a result, also mathematical theories describing physical phenomena have their accuracy determined by the accuracy of physical experiments. Obviously, a good theory can also show directions for

Y. D. Sergeyev (✉)
University of Calabria, Rende, Italy

Y. D. Sergeyev
N.I. Lobatchevsky State University, Nizhni Novgorod, Russia

Y. D. Sergeyev
Institute of High Performance Computing and Networking of the National Research Council of Italy, Roma, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_35,

new experiments and new tools that would confirm or refuse theoretical hypotheses. Thus, there exists a strong link between the mathematical theories describing physical world and the current level of accuracy of physical instruments.

In different historical periods, various mathematical disciplines were used for constructing physical models. Mathematical Analysis is one of them and, obviously, it is widely used in Physics. However, the foundations of Analysis have been developed more than 200 years ago with the goal to develop mathematical tools allowing one to solve problems arising in the real world at that time. Thus, Analysis that we use now does not reflect in itself numerous achievements of Physics of the twentieth century.[1] For instance, in Physics, the same object can be viewed as either discrete or continuous in dependence of the instrument used for the observation (we see a table continuous when we look at it by eye and we see it discrete (consisting of molecules, atoms, etc.) when we observe it at a microscope). Moreover, physicists together with the result of the observation supply the accuracy of the instrument used for this observation.

In Mathematics, both facts are absent: each mathematical object (e.g., function) is either discrete or continuous and nothing is said about the accuracy of the observation of the mathematical objects and about tools used for these observations. Mathematical notions have an absolute character and the ideas of relativity are not present in it. In some sense, there exists a gap between the physical achievements made in the last two hundred years and their mathematical models that continue to be written using the mathematical language developed two hundred years ago on the basis (among other things) of physical ideas of that remote time.

The point of view presented in this paper tries to fill up this gap. It uses strongly such methodological ideas borrowed from Physics and applied to Mathematics as: the distinction between the object (we speak here about a mathematical object) of an observation and the instrument used for this observation; interrelations holding between the object and the tool used for this observation; the accuracy of the observation determined by the tool. In particular, from this new physical point of view the ways to deal with infinities and infinitesimals are studied. The main attention is dedicated to mathematical languages, more precisely, to numeral systems[2] that we use to write down numbers, functions, models, etc. and that are among our tools of investigation of mathematical objects. It is shown that numeral systems strongly influence our capabilities to describe the inhabitants of the mathematical world.

---

[1] This is true also for the modern non-standard analysis [1] that re-writes the standard one in terms of infinitesimals and realizes the program of Leibniz. For instance, one of the basic concepts in the non-standard analysis is *monad*, the notion taken from Physics that, however, is not used in Physics for a long time.

[2] We remind that a *numeral* is a symbol or group of symbols that represents a *number*. The difference between numerals and numbers is the same as the difference between words and the things they refer to. A number is a concept that a numeral expresses. The same number can be represented by different numerals written in different numeral systems. For example, the symbols ‘4’, ‘four’, and ‘IV’ are different numerals, but they all represent the same number. Rules used in different numeral systems to execute arithmetical operation can be also different.

In particular, a new numeral system (see [2–11]) for performing computations with infinite and infinitesimal quantities is used for the observation of mathematical objects and modeling physical phenomena. It is based on the principle 'The part is less than the whole' introduced by Ancient Greeks and observed in the physical world. It is applied to all sets and processes (finite and infinite) and all numbers (finite, infinite, and infinitesimal). The new methodology has allowed the author to introduce the Infinity Computer (see the USA patent [11]) working numerically with infinite and infinitesimal numbers. The introduced computational paradigm both gives possibilities to execute computations of a new type and simplifies fields of Mathematics and Computer Science where infinity and/or infinitesimals are required. In order to see the place of the new approach in the historical panorama of ideas dealing with infinite and infinitesimal, see [12] and [13].

The new methodology has been successfully applied for studying fractals [6, 14], percolation [14, 15], optimization algorithms [3, 16], hyperbolic geometry [17], Turing machines [18], cellular automata [19] , infinite series [6–9, 20], etc.

35.2 A New Methodology for Performing Computations with Infinite and Infinitesimal Quantities

Traditionally, there exist different ways to work with mathematical objects connected to the concepts of infinite and infinitesimal (see, e.g., [1, 21–24] and references given therein). However, it is well known that we work with infinite objects in a way different with respect to the rules that we are used to deal with finite quantities. In fact, there exist undetermined operations (for example, $\infty - \infty$, $0 \cdot \infty$, etc.), divergences, etc. It is worthwhile to mention also that the philosophical principle of Ancient Greeks 'The part is less than the whole' observed in the physical world around us does not hold for infinities, including infinite cardinals introduced by Cantor, e.g., it follows $x+1 = x$, if x is an infinite cardinal, although for any finite x we have $x+1 > x$.

In order to understand how it is possible to look at the problem of infinity in a new way, let us consider a study published in Science (see [25]) where the author describes a primitive tribe living in Amazonia—Pirahã—that uses a very simple numeral system for counting: one, two, many. For Pirahã, all quantities larger than two are just 'many' and such operations as 2+2 and 2+1 give the same result, i.e., 'many'. Using their weak numeral system Pirahã are not able to see, for instance, numbers 3, 4, 5, and 6, to execute arithmetical operations with them, and, in general, to say anything about these numbers because in their language there are neither words nor concepts for that.

It is important to emphasize that the answer 'many' is not wrong, it is just *imprecise*. Analogously, the answer 'many' to the question 'How many trees are there in the garden in front of our house?' is correct, but its precision is low. Already these first considerations show us that numeral systems have their accuracy like physical instruments have. This means that when one uses a numeral system, this system defines the accuracy of mathematical results that can be obtained with its help. As a consequence, mathematical truths are not absolute; they are relative to the precision of the numeral systems (and, in general, to the mathematical language) used for their formulation. The understanding of the fact that Mathematics, as all natural sciences, depends on its instruments and is able to provide results that depend on the accuracy of the instruments used for their description is very important.

Pirahã's numeral system is interesting for us also because its weakness leads to such results as

$$\text{'many'} + 1 = \text{'many'}, \ \text{'many'} + 2 = \text{'many'}, \ \text{'many'} + \text{'many'} = \text{'many'}, \quad (35.1)$$

which are very familiar to us in the context of views on infinity used in the traditional calculus:

$$\infty + 1 = \infty, \qquad \infty + 2 = \infty, \qquad \infty + \infty = \infty. \qquad (35.2)$$

This similarity leads us to the following idea: *Probably our difficulty in working with infinity is not connected to the nature of infinity but is a result of inadequate numeral systems used to express infinite numbers*. Analogously, Pirahã do not distinguish numbers 3 and 4 not due to the nature of these numbers but due to the weakness of their numeral system.

When we have such mathematical objects as infinite numbers, then even the modern numeral systems used for the observation are not sufficiently precise. The instruments of Pirahã do not allow them to distinguish 3, 4, and other numbers higher than 2. Our modern numeral systems are more precise, they allow us distinguish various finite numbers, but they fail when it is necessary to work with infinite quantities.

The observation made above is important from several points of view; in particular, with respect to the mathematical instruments developed by Cantor (see [21]) who has shown that there exist infinite sets having different number of elements. Cantor has proved, by using his famous diagonal argument, that the cardinality, $\aleph_0$, of the set, $\mathbf{N}$, of natural numbers is less than the cardinality, C, of real numbers at the interval [0, 1]. Cantor has also developed an arithmetic for the infinite cardinals. Some of the operations of this arithmetic including $\aleph_0$ and C are given below:

$$\aleph_0 + 1 = \aleph_0, \qquad \aleph_0 + 2 = \aleph_0, \qquad \aleph_0 + \aleph_0 = \aleph_0, \qquad (35.3)$$

$$\mathrm{C} + 1 = \mathrm{C}, \qquad \mathrm{C} + 2 = \mathrm{C}, \qquad \mathrm{C} + \aleph_0 = \mathrm{C}, \mathrm{C} + \mathrm{C} = \mathrm{C}. \qquad (35.4)$$

Again, it is possible to see a clear similarity with the arithmetic operations used in the numeral system of Pirahã (a detailed discussion related to these issues dealing also with the Continuum Hypothesis see in [10]). This prompts us that, probably, Cantor's numeral system used to measure infinite sets can be also improved. If we were able to distinguish more infinite numbers we could understand better many processes and objects dealing with the concepts of infinite and infinitesimal (remind the famous phrase of Ludwig Wittgenstein: ‘The limits of my language are the limits of my world.’).

In [2–11] a new numeral system has been developed for performing computations with infinite and infinitesimal quantities using the principle ‘The part is less than the whole’ introduced by Ancient Greeks and observed in the physical world (it is discussed and compared with other approaches in [12] and [13]). The main idea of the new approach consists of the possibility to measure infinite and infinitesimal quantities by different (infinite, finite, and infinitesimal) units of measure.

For this purpose, a new infinite unit of measure expressed by the numeral CD called *grossone* is introduced as the number of elements of the set, N, of natural numbers. Concurrently with the introduction of grossone in our mathematical language such symbols like ∞, Cantor’s ω, C from (35.4), all Alephs $\aleph_0$,$\aleph_1$, etc. are excluded from the language because grossone ant other numbers constructed with its help not only can be used instead of all of them but can be used with a higher accuracy. Grossone is introduced by describing its properties postulated by the Infinite Unit Axiom [4] added to axioms for real numbers (similarly, in order to pass from the set, N, of natural numbers to the set, Z, of integers a new element—zero expressed by the numeral 0—is introduced by describing its properties).

It is necessary to notice that grossone has both cardinal and ordinal properties as usual finite natural numbers have. The new numeral allows us to construct different numerals expressing different infinite and infinitesimal numbers and to calculate the number of elements of certain infinite sets. For instance, it becomes possible to see that the sets of even and odd numbers have CD/2 elements each. The set, Z, of integers has 2CD + 1 elements (CD positive elements, CD negative elements, and zero). The set B = N \{b}, b $\in$ N, has CD$-$1 elements and the set A = N U { a1; a2 }, a1 $\notin$ N, a2 $\notin$ N, has CD + 2 elements.

Note that positive integers larger than grossone do not belong to N but they can be also easily interpreted as the number of elements of certain infinite sets. For instance, CD^{3}is the number of elements of the set V, where

$$V = \{(a_1; a_2; a_3) : a_1 \in N, a_2 \in N, a_3 \in N\}.$$

It is worthy to mention that these results do not contradict Cantor’s one-to-one correspondence principle (see [13] and [9, 10] for a detailed discussion). Both Cantor’s numeral system and the new one give correct answers, but their answers have different accuracy. We just use a stronger (with respect to cardinals of Cantor) tool, CD, for the observation of infinite sets that allows us to give more accurate answers than those of Cantor. By using the physical analogy we can say that the lens of our microscope is stronger and we are able to see many different dots where Cantor’s microscope allows him to observe just one dot—countable sets.

Note, that the new numeral system using grossone allows us to avoid records of the type (1)–(4). In fact, it can be easily shown (see [4, 9]) that, for example,

$$\text{CD}/2 < \text{CD} < 2\text{CD} < 2\text{CD}+1 < 2\text{CD}+2 < \text{CD}^3/2 < \text{CD}^3-1 < \text{CD}^3 < \text{CD}^3+1 < \text{CD}^3+2.$$

Within the sets having cardinality of the continuum it is also possible to distinguish infinite sets having different number of elements expressible in the numeral system using grossone (see [4, 10]):

$$2^{\text{CD}}-1 < 2^{\text{CD}} < 2^{\text{CD}}+1 < 10^{\text{CD}}-1 < 10^{\text{CD}} < 10^{\text{CD}}+1 < \text{CD}^{\text{CD}}-1 < \text{CD}^{\text{CD}} < \text{CD}^{\text{CD}}+1 < \text{CD}^{\text{CD}}+2.$$

The rules we use to execute arithmetical operations with 0 and 1 work with grossone, as well:

$$0\cdot\text{CD}=0,\ \text{CD}\cdot 0=0,\ \text{CD}-\text{CD}=0,\ \text{CD}:\text{CD}=1,\ \text{CD}^0=1,\ 0^{\text{CD}}=0,\ 1^{\text{CD}}=1.$$

Since CD^0=1, a finite number c can be represented in the new numeral system simply as cCD^0=c, where the numeral c is written down by any convenient numeral system used to write down finite numbers.

The simplest infinitesimal numbers are represented by numerals having only negative powers of grossone that are finite or infinite. The following two numerals are examples of infinitesimals expressed in the new numeral system: $4.2\text{CD}^{-3.21}$, $74.56\text{CD}^{-33.85\text{CD}}$. The simplest infinitesimal is CD^{-1}being the inverse element with respect to multiplication for CD:

$$\text{CD}\cdot\text{CD}^{-1}=1.$$

Note that all infinitesimals are not equal to zero. Particularly, it follows $\text{CD}^{-1} > 0$ because CD^{-1}= 1 / CD, i.e., it is a result of division of two positive numbers.

It is necessary to mention that it is not easy to interpret grossone in the framework of the traditional mathematics (including the non-standard analysis). When one tries to compare two languages, it can often happen that their accuracies are different and the translation becomes possible only partially. Moreover, different languages represent the world in different ways (any person knowing more than one language knows that there exist things that can be described better in one language than in another). In linguistics, the relativity of the language with respect to the world around us is well known. This fact has been formulated in the form of the Sapir–Whorf thesis of the linguistic relativity (see [26, 27]). For example, it is impossible to translate to the language of Pirahã the word 'four' from English maintaining the same accuracy in their language as in English. The same thing happens when we compare the mathematical languages using, on the one hand, such symbols like ∞, Cantor's ω, C from (35.4), all Alephs $\aleph_0$,$\aleph_1$, etc. and, on the other hand, grossone and other numerals constructed with its help. The accuracies of the two languages are different.

Another crucial problem related to such a translation consists of the fact that in the traditional mathematics (even in the non-standard analysis) very often there is no a clear distinction between numbers (and sets of numbers) and numerals and sets of numerals used to represent numbers.

However, as it happens in Physics, in Mathematics it is also always necessary to indicate the instrument one uses for an observation in order to understand what can be observed. If such a clarification is absent, then ambiguities can easily be encountered. Let us illustrate this affirmation by considering the following set defined by the phrase: the set of all numbers less than three. This phrase seems to identify clearly a set because, without saying it explicitly, we keep in mind that we are speaking about real numbers. So, it is supposed implicitly that the instrument of the observation of the set is a positional numeral system.

However, in different historical periods such implicit suppositions were different. Before discovering negative numbers (for instance, Roman and Greek numeral systems do not include zero and are not able to express negative numbers; in these systems, expression 3–7 was an undetermined form) this was the set of *positive* numbers less than three. Moreover, before discovering irrational numbers this phrase was defining *rational* positive numbers less than three. In the language of Pirahã, this set cannot be even defined because they do not know what 'three' is. Then, if we use the new approach working with grossone, it can be shown (see [10]) that in dependence on the radix of the positional system used to write down numerals, different sets of real numbers can be observed.

In particular, since the traditional mathematics very often does not pay any attention to the distinction between numbers and numerals, many theories dealing with infinite and infinitesimal quantities have a symbolic (not numerical) character. For instance, many versions of the non- standard analysis are symbolic, since they have no numeral systems to express their numbers by a finite number of symbols (this is necessary for organizing numerical computations). Namely, if we consider a finite n than it can be taken n = 6, or n = 12 or any other numeral consisting of a finite number of symbols. If we consider an infinitesimal h then it is not clear which numerals consisting of a finite number of symbols can be used to write h = . . . In fact, very often in non-standard analysis texts, *a generic* infinitesimal h is used and it is considered as a symbol, i.e., only symbolic computations can be done with it. Approaches of this kind leave unclear such issues, e.g., whether the infinite 1/h is integer or not or whether 1/h is the number of elements of an infinite set.

35.3 Examples of the Usage of the New Computational Methodology

In this section, we give several examples showing how to use the new computational methodology with the new numeral systems using grossone. Examples are chosen in such a way that they can be directly used in modeling physical phenomena.

We start by showing that, in the new language, infinite (both convergent and divergent) series can be substituted by sums with the precisely defined infinite number of items. Since in the new numeral system we have many different infinite integers and we have seen that the symbol ∞ is a kind of 'many', such records as

$$S = a_1 + a_2 + a_3 + a_4 \ldots \qquad \text{or} \qquad S = \sum_{i=1}^{\infty} a_i$$

become imprecise. In the new language they are just something like "Calculate the sum of a_i where i goes from 1 to 'many"'. This means that when we have infinitely many items in a sum, we are in the same situation as with sums having a finite number of items: it is not sufficient to say that the number, n, of items in the sum is finite, it is necessary to fix explicitly the value of n using for this purpose numerals available in a chosen traditional numeral system used for expressing finite numbers, e.g. n = VII or n = 2,000. Analogously, now we should decide (in dependence on the problem we deal with) how many items should be in our sum with an infinite number of items by choosing an appropriate value for n, e.g., n=10CD or n=2000CD^{12}, see [4, 5, 28] for a detailed discussion.

In problems where the traditional language using ∞ fails and is not able to provide any answer, the new numeral system using grossone allows us to work with expressions involving infinite numbers and to obtain, where it is possible, as results infinite, finite, and infinitesimal numbers. For instance, let us consider two divergent series

$$S_1 = 5 + 5 + 5 + \cdots, \qquad S2 = 3 + 3 + 3 + \cdots$$

Then, by using the traditional language, we are not able to execute such operations like, e.g., S1–S2 or S2 / S1 since both S1 and S2 are divergent series. In the new language, both expressions S1 and S2 are not well defined since the number of items in them is not specified. Instead of series, we should consider the sums S1(k1) and S2(k2), to fix the infinite number of items, k1, in the first sum and the infinite number of items, k2, in the second sum, to calculate the respective results, and then to execute the further required operations, see [4, 5, 28]. For instance, if k1 = 2CD + 2 and k2 = 4CD − 1 then we are able to calculate easily both S1(2CD + 2) and S2(4CD − 1) and to execute arithmetical operations with the results:

$$S_1(2CD + 2) = 5(2CD + 2) = 10CD + 10, \quad S_2(4CD - 1) = 3(4CD - 1) = 12CD - 3,$$
$$S_1(2CD + 2) - S_2(4CD - 1) = 10CD + 10 - (12CD - 3) = -2CD + 13,$$
$$S_1(2CD + 2) + S_2(4CD - 1) = 10CD + 10 + 12CD - 3 = 22CD + 7,$$
$$S_1(2CD + 2) \cdot S_2(4CD - 1) = 120CD^2 + 90CD - 30,$$
$$S_2(4CD - 1)/S_1(2CD + 2) = 1.2 - 1.5CD^{-1} + 1.5CD^{-2} - 1.5CD^{-3},$$

where the result of division has been taken with the error that is not higher than $O(CD^{-4})$.

The same situation we have with divergent integrals. Namely, records of the type $\int_a^b f(x)$ become imprecise if either a = $-\infty$ or b = ∞, or both. The limits of integration equal to $-\infty$ or ∞ should be substituted by exact infinite numbers. For instance, if a = −6CD, b = 3CD^2, and f(x) =x^2, then

$$\int_a^b f(x) = ①^6 + 2①^3,$$

where the result is the infinite number. By changing any limit of integration (as it happens in the case of finite limits a and b) the result will also change. For instance, by taking the infinitesimal a = 3CD^{-2}, we have that the integral is equal to CD^6–CD^{-6}. We are able to find easily difference of the two obtained infinite numbers:

$$\mathrm{CD}^6 + 2\mathrm{CD}^3 - (\mathrm{CD}^6 - \mathrm{CD}^{-6}) = 2\mathrm{CD}^3 + \mathrm{CD}^{-6},$$

that is the result of integration for a = −6CD, b = 3CD^{-2}. By taking the limits a = 3CD^{-2}, b = 3CD^2 and by taking f(x) = CDx^2 we have that the integral, since grossone is a constant and can be put out of the sign of integral, is equal to CD^7–CD^{-5}. Note also that we can easily work with derivatives that can assume infinite (and infinitesimal) values; e.g., if f(x) = CDx^2 then f'(x) = 2CDx, analogously, if g(x) = CD^{-6}x^2 then g'(x) = 2CD^{-6} x.

In general, it becomes possible to use infinite and infinitesimal numbers as constants to construct new mathematical objects that are not visible when the traditional mathematical language is used, see [4, 5, 28]. Let us consider, for example, the following discontinuous function that cannot be described by the traditional mathematical language, see Fig. 35.1.

If we try to describe this function by using the traditional mathematical language working with finite numerals and the symbol ∞, we shall see immediately that we are not able to do this. To be more precise, we can give only a rough qualitative description of this function similar to the following one: The function is ∞ everywhere and it

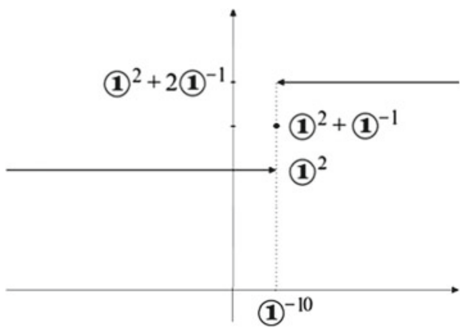


Fig. 35.1 Example of a discontinuous function having the point of discontinuity at the infinitesimal point CD^{-10} and assuming three different infinite values at this point, on its *left*, and on its *right*

has an infinitesimal jump infinitesimally close to zero. This description has just a qualitative character since in the traditional language there are no numerals allowing us to express different infinite and infinitesimal numbers. In contract, the description presented in Fig. 35.1 is quantitative; it uses new numerals that allow us to write the function easily as follows

$$f(x) = \begin{cases} \text{①}^2, & x < \text{①}^{-1}, \\ \text{①}^2 + \text{①}^{-1}, & x = \text{①}^{-1}, \\ \text{①}^2 + 2\text{①}^{-1}, & x > \text{①}^{-1}, \end{cases}$$

The usage of new numerals allows us to quantify even infinitesimal changes in the structure of objects under consideration. For instance, the following function g(x) is clearly different from the function f(x)

$$g(x) = \begin{cases} \text{①}^2, & x < \text{①}^{-1}, \\ \text{①}^2 + \text{①}^{-1,2}, & x = \text{①}^{-1}, \\ \text{①}^2 + 2\text{①}^{-1}, & x > \text{①}^{-1}, \end{cases}$$

even though the difference is infinitesimal; it holds at the point x = CD^{-1}. Note that the traditional language is able again to give only a qualitative description of g(x); moreover, this description will be the same as for function f(x).

The passage from a qualitative description to quantitative ones is very important when we speak about natural sciences. It allows us to measure infinite and infinitesimal quantities and to distinguish many different objects in cases that are difficult for the traditional mathematics because it is able to see just ∞ or $\aleph_0$ and $\aleph_1$. Let us illustrate this fact by considering what the traditional language and the new one can say with respect to the delta-function—the object widely used in Physics.

Informally, it is a generalized function depending on a real x such that it is zero for all values of the x except when x = 0, and its integral from $-\infty$ to ∞ is equal to one. The new numeral system allows us to distinguish many different

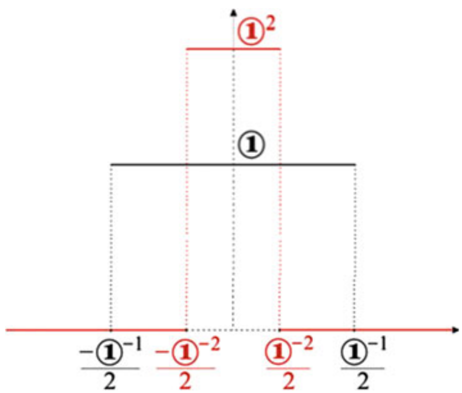


Fig. 35.2 Examples of two different delta-functions

delta-functions assuming a concrete infinite value over a concrete infinitesimal interval. Two of them are shown in Fig. 35.2. The first one is equal to CD over the interval $[-0.5\text{CD}^{-1}, 0.5\text{CD}^{-1}]$ and the second one is equal to CD^2over the interval $[-0.5\text{CD}^{-2}, 0.5\text{CD}^{-2}]$.

We conclude this paper with the hope that the new numeral system allowing us to work numerically with different infinite and infinitesimal numbers on the Infinity Computer (see [11]) will allow physicists and mathematicians to construct new models describing the physical world aroungd us better than it is done actually when the traditional numeral systems are applied for this purpose.

References

1. Robinson A (1996) Non-standard analysis. Princeton University Press, Princeton
2. Sergeyev Ya D (2003) Arithmetic of infinity. Edizioni Orizzonti Meridionali, CS
3. Sergeyev Ya D (2006) Misuriamo l'infinito. Periodico di Matematiche 6(2):11–26
4. Sergeyev Ya D (2008) A new applied approach for executing computations with infinite and infinitesimal quantities. Informatica 19(4):567–596
5. Sergeyev Ya D (2009a) Numerical point of view on Calculus for functions assuming finite, infinite, and infinitesimal values over finite, infinite, and infinitesimal domains. Nonlinear Anal Ser A Theor Methods Appl 71(12):e1688–e1707
6. Sergeyev Ya D (2009b) Evaluating the exact infinitesimal values of area of Sierpinski's carpet and volume of Menger's sponge. Chaos, Solitons Fractals 42:3042–3046
7. Sergeyev Ya D (2009c) Numerical computations and mathematical modelling with infinite and infinitesimal numbers. J Appl Math Comput 29:177–195
8. Sergeyev Ya D (2009d) Computer system for storing infinite, infinitesimal, and finite quantities and executing arithmetical operations with them. EU patent 1728149(03):06
9. Sergeyev Ya D (2010a) Lagrange lecture: methodology of numerical computations with infinities and infinitesimals. Rendiconti del Seminario Matematico dell'Università e del Politecnico di Torino 68(2):95–113
10. Sergeyev Ya D (2010b) Counting systems and the first Hilbert problem. Nonlinear Anal Ser A Theor Methods Appl 72(3–4):1701–1708
11. Sergeyev Ya D (2010c) Computer system for storing infinite, infinitesimal, and finite quantities and executing arithmetical operations with them. USA patent 7,860,914. 28.12.
12. Lolli G (2012) Infinitesimals and infinites in the history of Mathematics: a brief survey. Appl Math Comput (in press).
13. Margenstern M (2011) Using Grossone to count the number of elements of infinite sets and the connection with bijections, p-Adic numbers. Ultrametric Anal Appl 3(3):196–204
14. Vita M C, De Bartolo S, Fallico C and Veltri, M (2012) Usage of infinitesimals in the Menger's Sponge model of porosity. Appl Math Comput (in press).
15. Iudin DI, Sergeyev Ya D and Hayakawa M (2012) Interpretation of percolation in terms of infinity computations. Appl Math Comput (in press).
16. De Cosmis S, De Leone R (2012) The use of Grossone in mathematical programming and operations research. Appl Math Comput (in press).
17. Margenstern M (2012) An application of Grossone to the study of a family of tilings of the hyperbolic plane. Appl Math Comput (in press).
18. Sergeyev Ya D, Garro A (2010) Observability of turing machines: a refinement of the theory of computation. Informatica 21(3):425–454
19. D'Alotto L (2012) Cellular automata using infinite computations. Appl Math Comput (in press).
20. Zhigljavsky A (2012) Computing sums of conditionally convergent and divergent series using the concept of grossone. Appl Math Comput (in press).

21. Cantor G (1955) Contributions to the founding of the theory of transfinite numbers. Dover Publications, New York
22. Conway JH, Guy RK (1996) The book of numbers. Springer, New York
23. Levi-Civita T (1898) Sui numeri transfiniti. Rend Acc Lincei 5a(113):7–91.
24. Mayberry JP (2001) The foundations of mathematics in the theory of sets. Cambridge University Press, Cambridge
25. Gordon P (2004) Numerical cognition without words: evidence from Amazonia. Science 306:496–499
26. Carroll JB (ed) (1956) Language, thought, and reality: selected writings of Benjamin Lee Whorf. MIT Press, Cambridge
27. Sapir E (1958) Selected writings of Edward Sapir in language, culture and personality. University of California Press, Princeton
28. Sergeyev Ya D (2011a) On accuracy of mathematical languages used to deal with the Riemann zeta function and the Dirichlet eta function, p-Adic numbers. Ultrametric Anal Appl 3(2):129–148
29. Sergeyev Ya D (2011b) Higher order numerical differentiation on the infinity computer. Optimization Letters 5(4):575–585
30. The Infinity Computer web page, http://www.theinfinitycomputer.com

Part IX
Computational Physics

Chapter 36
Seismic Hazard Assesment: Parametric Studies on Grid Infrastructures

Andrea Magrin, Cristina La Mura, Franco Vaccari, Giuliano F. Panza, Alexander A. Gusev, Iztok Gregori and Stefano Cozzini

Abstract Seismic hazard assessment can be performed following a neo-deterministic approach (NDSHA), which allows to give a realistic description of the seismic ground motion due to an earthquake of given distance and magnitude. The approach is based on modelling techniques that have been developed from a detailed knowledge of both the seismic source process and the propagation of seismic waves. This permits us to define a set of earthquake scenarios and to simulate the associated synthetic signals without having to wait for a strong event to occur. NDSHA can be applied at the regional scale, computing seismograms at the nodes of a grid with the desired spacing, or at the local scale, taking into account the source characteristics, the path and local geological and geotechnical conditions. Synthetic signals can be produced in a short time and at a very low cost/benefit ratio. They can be used as seismic input in subsequent engineering analyses aimed at the computation of the full non-linear seismic response of the structure or simply the earthquake damaging potential. Massive parametric tests, to explore the influence not only of deterministic source parameters and structural models but also of random properties of the same source model, enable realistic estimate of seismic hazard and their uncertainty. This is particular true in those areas for which scarce (or no) historical or instrumental information is available. Here we describe the implementation of the seismological codes and the results of some parametric tests performed using the EU-India Grid infrastructure.

A. Magrin (✉) · C. La Mura · F. Vaccari · G. F. Panza
Dipartimento di Geoscienze, Università di Trieste, Trieste, Italy

A. A. Gusev
Kamchatka Branch, Geophysical Service, Russian Academy of Sciences, Institute of Volcanology and Seismology, Petropavlovsk-Kamchatsky, Russia

I. Gregori · S. Cozzini
CNR/IOM c/o Sissa, Trieste, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_36,

36.1 Neo-Deterministic Seismic Hazard Assessment

The typical seismic hazard problem lies in the determination of the ground motion characteristics associated with future earthquakes, at both the regional and the local scale. Seismic hazard assessment can be performed in various ways, e.g. with a description of the groundshaking severity due to an earthquake of a given distance and magnitude ("groundshaking scenario"), or with probabilistic maps of relevant parameters describing ground motion.

Seismic hazard assessment can be performed following a neo-deterministic approach (NDSHA), which allows to give a realistic description of the seismic ground motion due to an earthquake of given distance and magnitude. The procedure for neo-deterministic seismic zoning [8] is based on the calculation of synthetic seismograms. It can be applied also to areas that have not yet been hit by a catastrophic event in historical times, but are potentially prone to it. The neo-deterministic method allows to quantitatively model the effects of an earthquake which may happen in the future and therefore is a very effective technique in seismic hazard assessment, even in the regions with scarce or no historical or instrumental information available.

Starting from the available information on the Earth's structure, seismic sources, and the level of seismicity of the investigated area, it is possible to compute complete synthetic seismograms and the related estimates on peak ground acceleration (PGA), velocity (PGV) and displacement (PGD) or any other parameter relevant to seismic engineering (such as design ground acceleration, DGA) which can be extracted from the computed theoretical signals.

NDSHA can be applied at the regional scale, computing seismograms at the nodes of a grid with the desired spacing, or at the local scale, taking into account the source characteristics, the path and local geological and geotechnical conditions.

In the NDSHA approach, the definition of the space distribution of seismicity accounts only for the largest events reported in the earthquake catalogue at different sites, as follows. Earthquake epicenters reported in the catalogue are grouped into $0.2^\circ \times 0.2^\circ$ cells, assigning to each cell the maximum magnitude recorded within it. A smoothing procedure is then applied to account for spatial uncertainty and for source dimensions [8]. Only cells located within the seismogenic zones are retained. This procedure for the definition of earthquake locations and magnitudes for NDSHA makes the method pretty robust against uncertainties in the earthquake catalogue, which is not required to be complete for magnitudes lower than 5. A double-couple point source is placed at the center of each cell, with a focal mechanism consistent with the properties of the corresponding seismogenic zone and a depth, which is a function of magnitude.

To define the physical properties of the source-site paths, the territory is divided into $1^\circ \times 1^\circ$ cells, each characterized by a structural model composed of flat, parallel anelastic layers that represent the average lithosphere properties at regional scale [2]. Synthetic seismograms are then computed by the modal summation technique for sites placed at the nodes of a grid with step $0.2^\circ \times 0.2^\circ$ that covers the national territory, considering the average structural model associated to the regional polygon

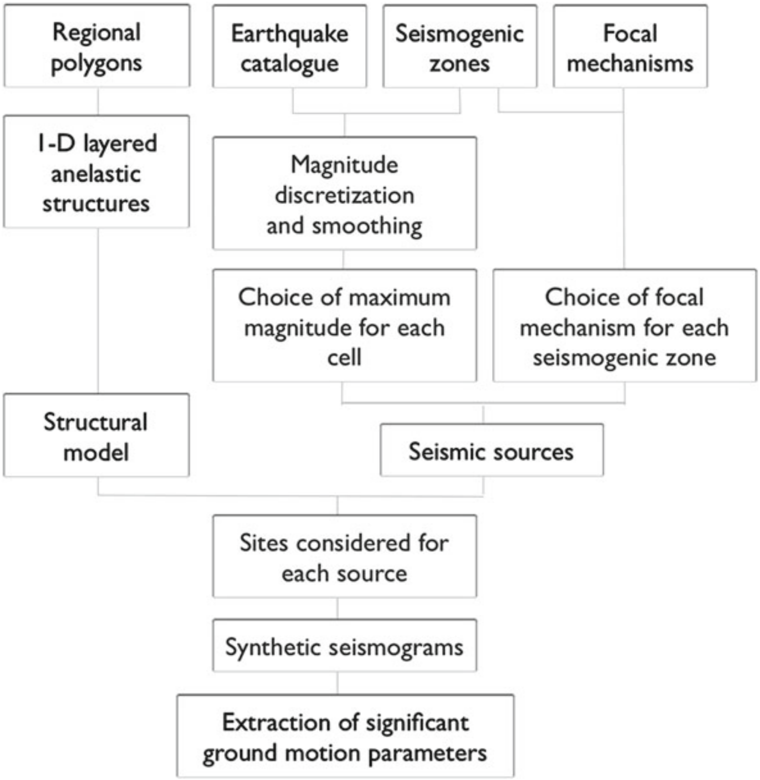


Fig. 36.1 Flow chart of national scale hazard package

that includes the site. Seismograms are computed for an upper frequency content of 1 Hz, which is consistent with the level of detail of the regional structural models, and the point sources are scaled for their dimensions using the spectral scaling laws proposed by GUSEV [3], as reported in AKI [1] (Fig. 36.1).

From the set of complete synthetic seismograms, various maps of seismic hazard describing the maximum ground shaking at the bedrock can be produced. The parameters representative of earthquake ground motion are maximum displacement, velocity, and acceleration. The acceleration parameter in the NDSHA is given by the design ground acceleration (DGA). This quantity is obtained by computing the response spectrum of each synthetic signal for periods of 1 s and longer (the periods considered in the generation of the synthetic seismograms) and extending the spectrum, at frequencies higher than 1 Hz, using the shape of the Italian design response spectrum for soil A [6], which defines the normalized elastic acceleration response spectrum of the ground motion for 5 % critical damping. For more details see PANZA et al. [7].

The neo-deterministic method has been recently adapted to account for the extended source process by including higher frequency content (up to 10 Hz) as well as the rupture process at the source and the consequent directivity effect (obtained by means of the PULSYN algorithm by GUSEV [4]).

The most common approach to the simulation of the earthquake source consists in adopting a kinematic model; the kinematic description is merely phenomenological but, also their simplest versions (e.g. [5]) are able to describe the gross features of the rupture process by simply using five source parameters: the fault dimensions (length, L, and width, W), the amount of slip at any point of the fault, the rise-time, and the rupture velocity. In order to describe the source microstructure, i.e. its roughness, one can use a stochastic or a deterministic (or a combination of both) distribution of barriers and asperities on the fault surface resulting in a non-uniform distribution of slip. Actually dynamic considerations should underlie this choice, but a kinematical model simply takes into account the existence of fault irregularities. Thus, at a given site the motion is dependent on the size and the duration of each of the sub-sources, and on their distribution in space and in time. Furthermore, at epicentral distances comparable with the dimensions of the fault, the relative positions of the sub-sources with respect to a site receiver can play a fundamental role in the interference of the different wavetrains, resulting in the so-called directivity effect. For an earthquake with a given radiated energy, the decay of the source spectrum at high frequencies shows a strong azimuthal dependence, since the corner frequency at a given receiver in the near-field is a function of all the kinematic source parameters.

The seismic waves due to an extended source are obtained by approximating the fault with a rectangular plane surface, on which the main rupture process is assumed to occur. The source is represented as a grid of point subsources, and their seismic moment rate functions are generated considering each of them as realizations (sample functions) of a non-stationary random process. Specifying in a realistic way the source length and width, as well as the rupture velocity, one can obtain realistic source time functions, valid in the far-field approximation. Furthermore, assuming a realistic kinematic description of the rupture process, the stochastic structure of the accelerograms can be reproduced, including the general envelope shape and peak factors. The extended earthquake source model allows us to generate a spectrum (amplitude and phase) of the source time function that takes into accounts both the rupture process and directivity effects, also in the near source region.

Slip on the fault varies in space and increases weakly monotonously in time, so that the dislocation rate is non-negative. Its unit moment tensor (defined by slip direction and fault-normal direction) does not vary over this area or in time. Therefore, the description of the source in space-time is essentially scalar, in terms of the distribution over the fault area of the seismic moment (and its time rate). To specify temporal and spectral properties of the simulated sources, the equivalent point source (SSPS) moment rate time history $\dot{M}_0(t)$, its Fourier transform $\dot{M}_0(f)$ and corresponding amplitude spectrum ("source spectrum") $\left|\dot{M}_0(f)\right|$ have been widely used (e.g. [8]).

36.2 Uncertainties in Hazard Maps

In NDSHA, the treatment of uncertainties is performed by sensitivity analyses for key modelling parameters. Fixing the uncertainty related to a particular input factor is an important component of the procedure. The input factors must account for the imperfectness in the prediction of fault radiation, and for the use of Green functions, for a given medium, that are only imperfectly known. At present, the following factors are selected as assumedly dominating ones.

- Fault radiation factors
  - parameters related to the point source position/orientation representing the fault in world coordinates: centroid depth, strike/dip/rake. The point source horizontal position within the cell is not perturbed.
  - intrinsic fault parameters describing the extension of the fault, like fault length and width vs. Mw (moment magnitude) relationship, Brune's stress drop, fault-average rupture velocity, directivity effects.
  - random seeds that define space-time structure of a particular realization of the random source, including random subsource time functions.
- Path factors
  - parameters that specify uncertainties in the assumed 1D bedrock velocity-density-Q profiles.

36.3 Implementation of Seismological Codes on Grid

The use of the EU-India Grid infrastructure (http://www.euindiagrid.eu/) allows to conduct massive parametric tests, to explore the influence not only of deterministic source parameters and structural models but also of random properties of the same source model, to enable realistic estimate of seismic hazard and their uncertainty. The random properties of the source are specially important in the simulation of the high frequency part of the seismic ground motion.

We have ported and tested seismological codes for national scale on the Grid infrastructure. The first step was the speed optimization of this package by the identification of the critical programs and of the hot spots within programs. The critical point in the algorithm was the computation of synthetic seismograms. The optimization was performed in two ways: first by removing of repeated formatted disk I/O, second by sorting of seismograms by source depth, to avoid the repeated computation of quantities that are depth-dependent.

The second step was porting of the national scale hazard package on EU-IndiaGRID infrastructure. Two different types of parametric tests were developed: on the deterministic source parameters and on the random properties of the source model. The first experiment is performed by perturbing the properties of the seismic

Table 36.1 Performance of the three test runs

	Seed1 Hz	Persut Italy	Seed10 Hz
Type of submission	Direct to CE	WMS	WMS
Total no of job	216	526	600
% of successful job	26	65	80
Total time of computation of successful job	948 h	1200 h	791 h
Average time of computation for one job	17 h	3.5 h	1.6 h
Number of computed seismograms for one job	197720	91398	29232

sources selected by the algorithm before the computation of synthetic seismograms. In the second test different sets of curves of source spectrum generated by a Monte-Carlo simulation of the source model are used for scaling the seismograms. In both cases there are many independent runs to be executed, so a script for generation of the input and other scripts for checking the status of jobs, retrieving the results and relaunching aborted jobs were developed.

36.4 Preliminary Results

One preliminary test over deterministic source parameter for whole Italy ("persut Italy"), and two different tests over random properties ("seed1 Hz" and "seed10 Hz") for the whole Italian territory, with different frequency content and different maximum distance for the computation of seismograms, were conducted. The performance of the package over the grid in terms of computational time and number of successful jobs was tested, and submission of job and retrieval of its output were refined.

The number of seismograms that must be computed determines the duration and the storage requirement of the run. This parameter seems critical for the success of the job. The test runs on the random component of the source gave an indication on the effective number of jobs that must be computed to have a good estimate of the distribution of the ground shaking peaks at each receiver (Table 36.1).

The first runs have provided a preliminary evaluation of the uncertainty of the hazard maps due to the random representation of the source and to the uncertainty on source parameter. Figure 36.2 shows an example of results of the test on the random component of the source model. The variability on the different random realizations of the source model (right) is shown in terms of ratio between standard deviation and average at each receiver.

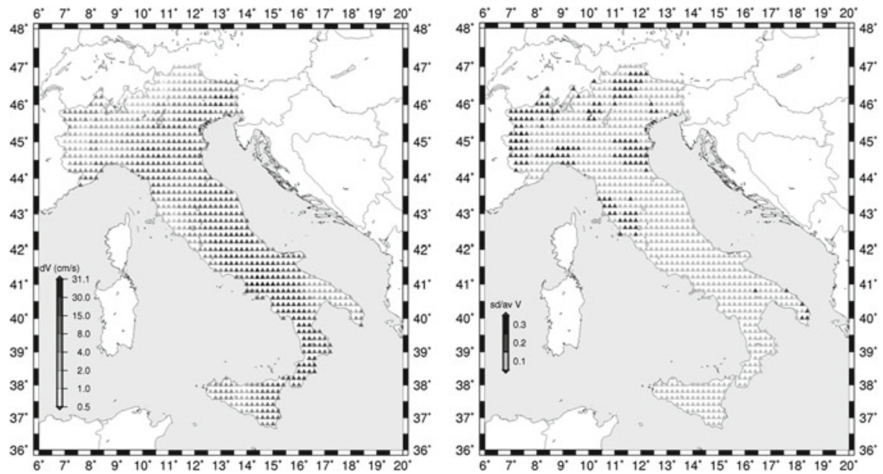


Fig. 36.2 Maps of average of PGV (peak ground velocity) on different random realizations of source model (*left*) and variability of the PGV in terms of ratio between standard deviation and average of the maximum peaks at each receiver

36.5 Conclusions and Perspectives

We have ported and tested seismological codes for seismic hazard assessment at national scale on the Grid infrastructure. The use of the EU-India Grid infrastructure allows to conduct massive parametric tests for evaluating the uncertainties in the computed hazard maps. Two different types of parametric tests were developed: on the deterministic source parameters and on the random properties of the source model. The performance of the package over the grid in terms of computational time and number of successful jobs were tested and submission of job and retrieval of its output were refined. The tests on the random component of the source gave an indication on the effective number of jobs that must be executed to have a good estimate of the distribution of the ground shaking peaks at each site. At the same time, they provided a preliminary estimate of the uncertainty of the hazard maps due to the random representation of the source.

The procedure followed for porting the package on the grid infrastructure can be implemented for other seismological programs as well.

A Cooperation Project, aimed at the definition of seismic and tsunami hazard scenarios by means of indo-european e-infrastructures in the Gujarat region (India), has been recently funded by the Friuli Venezia Giulia Region. This two-years project, starting in November 2011, involves three Italian partners (DiGeo, University of Trieste; ICTP SAND Group; CNR/IOM uos Democritos) and two Indian partners (ISR, Gujarat; CSIR C-MMACS, Bangalore). The project aims to set up a system for the seismic characterization, integrated with the e-infrastructures distributed amongst India and Europe, to allow for the optimization of the computation of the ground shaking and tsunami scenarios. This goal will be attained thanks to the strict connection with the European project EU-IndiaGrid2, that will provide the necessary infrastructure. Thus, the project will permit developing an integrated system, with high scientific and technological content, for the definition of scenarios of ground shaking, providing in the same time to the local community (local authorities and engineers) advanced information for seismic and tsunami risk mitigation in the study region.

References

1. Aki K (1987) In: Erdik M, Toksoz M (eds) Strong ground motion seismology, NATO ASI Series, Series C: Mathematical and Physical Sciences, vol 204, D. Reidel Publishing Company, Dordrecht, pp 3–39
2. Brandmayr E, Raykova RB, Zuri M, Romanelli F, Doglioni C and Panza GF (2010) The lithosphere in Italy: structure and seismicity. In: Beltrando M, Peccerillo A, Mattei M, Conticelli S, Doglioni C (eds) The Geology of Italy Journal of the Virtual Explorer, Electronic edn, vol 36, paper 1. ISSN 1441–8142. http://virtualexplorer.com.au/article/2009/224/lithosphere-structure-seismicity
3. Gusev AA (1983) Descriptive statistical model of earthquake source radiation and its application to an estimation of short period strong motion. Geophys J R Astron Soc 74:787–800
4. Gusev AA (2011) Broadband kinematic stochastic simulation of an earthquake source: a refined procedure for application in seismic hazard studies. Pure Appl Geophys 168(1–2): doi:10.1007/s00024-010-0156-3
5. Haskell NA (1964) Total energy and energy spectral density of elastic wave radiation from propagating faults. Bull Seismol Soc Am 54:1811–1841
6. Norme Tecniche per le costruzioni (D.M. 14/09/2005), published on G.U. 23/09/2005.
7. Panza GF, Vaccari F, Costa G, Suhadolc P, Fah D (1996) Seismic input modelling for zoning and microzoning. Earthq Spectra 12:529–566
8. Panza GF, Romanelli F, Vaccari F (2001) Seismic wave propagation in laterally heterogeneus anelastic media: theory and applications to seismic zonation. Adv Geophys 43:1–95
9. Zuccolo E, Vaccari F, Peresan A, Panza GF (2010) Neo-deterministic and probabilistic seismic hazard assessments: a comparison over the Italian territory. Pageoph. doi:10.1007/s00024-010-0151-8

Part X
Physics Teaching/Learning and Teachers Formation

Chapter 37
Learning Scenarios for a 3D Virtual Environment: The Case of Special Relativity

Cécile de Hosson, Kermen Isabelle, Maisch Clément, Parizot Etienne, Doat Tony and Vézien Jean-Marc

Abstract Special Relativity, as introduced by Einstein, is regarded as one of the most important revolutions in the history of physics. Nevertheless, the observation of direct outcomes of this theory on mundane objects is impossible because they can only be witnessed when travelling at relative speeds approaching the speed of light c. These effects are so counterintuitive and contradicting with our daily understanding of space and time that physics students find it hard to learn special relativity beyond mathematical equations and to understand the deep implications of the theory. Although we cannot travel at the speed of light, Virtual Reality (VR) makes it possible to experiment the effects of relativity in a 3D immersive environment (a CAVE: Cave Automatic Virtual Environment). The use of the immersive environment is underpinned by the development of dedicated learning scenarios created through a dialectic between VR-related computational constraints and cognitive constraints that include students' difficulties.

37.1 Introduction

This research takes place within the context of the EVEILS research project (French acronym for Virtual Spaces for the Education and Illustration of Science). This project aims at exploring the innovating potential of Virtual Reality (VR) in several areas of science through an interdisciplinary approach involving physicists, VR

EVEIL is supported by the French National Research Agency (ANR) and conducted under the responsibility of Pr E. Parizot (APC Université Paris Diderot-Paris 7).

C. de Hosson (✉) · K. Isabelle · M. Clément
LDAR, Université Paris Diderot-Paris 7, Paris, France

P. Etienne
APC, Université Paris Diderot-Paris 7, Paris, France

D. Tony · V. Jean-Marc
LIMSI CNRS/Orsay, Orsay, France

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_37,

specialists and physics education researchers. The project exploits advanced interfaces in order to confront a student with unusual phenomena otherwise inaccessible to human experience. The exploration of the cognitive modifications and pedagogical advantages associated with the 'immersion' is part of the main goals of EVEILS. This educational aspect makes EVEILS quite specific among the research programs devoted to computer simulations associated with VR [1].

Special Relativity, as introduced by Einstein, is regarded as one of the most important revolutions in the history of physics. Nevertheless, the observation of direct outcomes of this theory on mundane objects is impossible because they can only be witnessed when travelling at relative speeds approaching the speed of light c. The theory of Special Relativity teaches us that space and time are neither absolute, that is, independent of the observer (or the reference frame associated with the observer), nor independent from one another. Instead, they make up a global geometric structure with 4 dimensions, called space-time, whose "time" and "space" components depend on the reference frame used to describe physical bodies and events in terms of positions and instants. In particular, the length of a given object, as well as the duration of a phenomenon (between two well-defined events) will be—not only appear—different for two observers moving with respect to one another. These effects are so counterintuitive and contradicting with our daily understanding of space and time that physics students find it hard to learn relativity beyond mathematical equations and to understand the deep implications of the theory. Although macroscopic objects can not travel at the speed of light, Virtual Reality (VR) makes it possible to experiment the effects of relativity in a 3D immersive environment (a CAVE: Cave Automatic Virtual Environment, see Fig. 37.1)[1] where the speed of the light is simulated to a reduced value. The EVEILS project is a framework designed to merge advanced 3D graphics with Virtual Reality interfaces in order to create an appropriate environment to study and learn relativity as well as to develop some intuition of the relativistic effects and the quadri-dimensional reality of space-time. Of course, very early, mathematics were used to predict what objects would look like in relativistic motions and this have been applied to VR computing. Nevertheless, the specificity of the EVEILS project is to make physics education research actively involved in both conception and evaluation of the scenarios to be implemented into the CAVE.

37.2 Overview of the Research

The use of the immersive environment is underpinned by the development of dedicated learning scenarios created through a dialectic between VR-related com-

[1] A CAVE is a surround-screen, surround-sound, projection-based virtual reality system. The illusion of immersion is created by projecting 3D computer graphics into a cube composed of display screens that completely surround the viewer. It is coupled with head and hand tracking systems to produce the correct stereo perspective and to isolate the position and orientation of a 3D input device. The viewer explores the virtual world by moving around inside the cube and grabbing objects with a appropriate device.

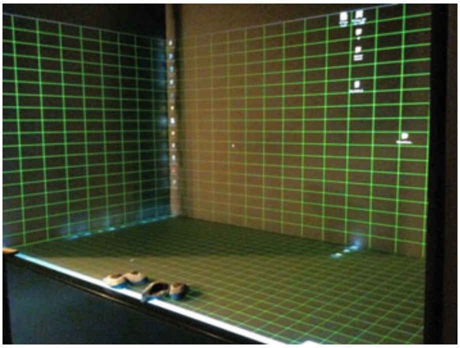


Fig. 37.1 Picture of the CAVE used for the EVEILS project (LIMSI, CNRS/Orsay, France)

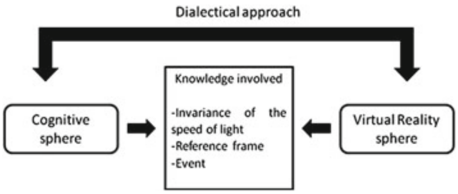


Fig. 37.2 Diagram of the research process

putational constraints and cognitive constraints that include students' difficulties. Investigating students' understanding of relativistic situations (that involve speeds closed to c) led to the typifying of a cognitive profile that orientated the situations to be implemented into the CAVE and the associated learning scenarios (see Fig. 37.2).

These scenarios aim at approaching the consequences of the invariance of the speed of light and more specifically the relativity of the simultaneity but also a deeper understanding of the concepts of "reference frame" and "event" (in physics). Here we will present the results of the characterization of the cognitive profile and its consequences on the development of the scenarios.

37.3 Students' Difficulties in Special Relativity: Elements of the State of the Art

The transition from classical to relativistic kinematics requires a radical change in the conceptual framework. In the theory of special relativity, c is a constant that connects space and time in the unified structure of space-time. The speed of light is equal to that constant and thus is invariant with respect to any inertial reference frame. Besides, the simultaneity of two events is not absolute (two events at different locations that occur at the same time in a given reference frame are not simultaneous in all other reference frames). Assuming this change in the conceptual framework requires a sound knowledge of the concepts of *reference frame* and *event* that underpin the laws of classical kinematics.

Studies conducted in order to characterize student's difficulties in special relativity are not very numerous. Nevertheless, from what have been explored we can detain that students use 'spontaneous' kinematics lines of reasoning (such as absolute motion, distances and velocities) to explain mechanical phenomena in both classical and special relativity frameworks [2, 3]. Students think that *simultaneity* is absolute and independent of relative motions [3, 4]. Students fail in understanding the concept of *reference frame* confusing "reference frame" and "point of view". Thus, each observer constitutes a distinct *reference frame* [4]. Moreover, they fail in defining and using the concept of *event* and thus confuse the instant of an *event* and the instant of the perception of that *event* by an observer [4, 5].

By connecting students with visual consequences of movements whose speed appears close to the speed of light we also hope we can reach a better understanding of concepts involved in the classical kinematics framework. This echoes various researches that pointed out the benefits of introducing modern physics in high school in order to improve fundamental concepts comprehension even in classical physics [6]. This could also contribute to update the physics content taught in the secondary school context.

37.4 Confronting Students to the 3D Virtual Environment

Considering both VR-related computational and cognitive constraints we designed learning scenarios to be implemented into the CAVE. These scenarios aim at giving direct access to:

- lengths and durations are not invariant and depend on the relative velocity between the objects and the reference frames involved. Thus, there is a priori a conflict between the intrinsic definition of the objects in their own reference frame and their actual occurrence in other reference frames, with respect to which they are moving. More precisely, Special Relativity teaches us that, in these other reference frames, the (instantaneous) lengths between two given points of the object are generally not the same. For instance, a billiard puck that is intrinsically a sphere, is no longer a sphere when described in the rest frame of the billiard board (see Fig. 37.3, with respect to which it is moving. This calls for a consistent description of the objects in any reference frame, i.e. in the 4D space-time reality itself.
- the speed of light c is finite (and invariant), so we do not see the objects where they are now, but where they were when they emitted the photons that we perceive now. The determination of what a given observer effectively sees at a given location at a given time (i.e. at a given point in space-time), requires a framework in which the whole history containing the past positions of the various objects of the scene is accessible to find the emission event.

The general idea of the scenarios is to confront users immersed into the CAVE with objects (billiard pucks) moving on a billiard table (without frictions) at a speed approaching that of the light (Fig. 37.4). Interaction with the simulation is made possible by applying impulsions to the pucks, to observe the effect of the limited

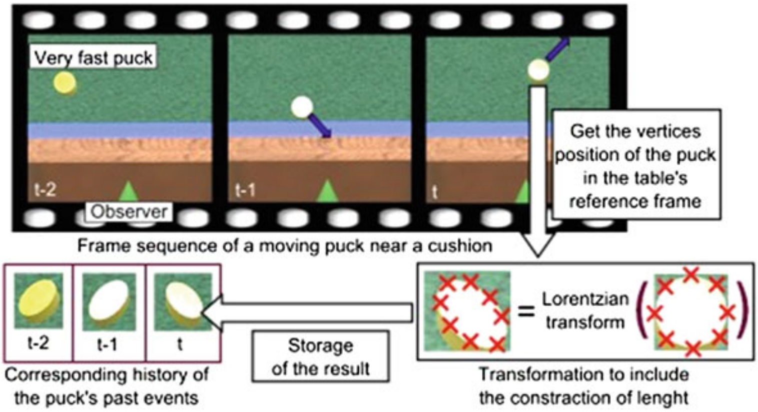


Fig. 37.3 The rendering of a relativistic scene involves a search in the history of each vertex (point) of the billiard pucks. Each time an image is generated for the user, we store the position of each vertex at the current simulation time, after applying Lorentz contraction to the intrinsic definition of the objects. The history table built in this way can then be deep-searched through by dichotomy to find the emission event associated with each vertex, at any later observation event [7]

speed of light and the Lorentzian contraction of length in the carom billiard. [2] Indeed, our application allows observing some subtle consequences of the theory of Special Relativity which are particularly important for physics education:

- The changes in the puck shape
- The apparent non-simultaneity of the bounces
- The apparent acceleration and deceleration of the pucks
- The aberration of the light

With respect to students' spontaneous lines of reasoning (see above) we focused on the concept of *event* as defined by its coordinates in space and time in a given reference frame. Since the perception of an event (e.g. emission of light) by an observer depends on the time taken by the light to travel from the location of the event to the eyes of the observer, the signal travel time of light will impact the observations of the users immersed in the simulator. In everyday life the light signal travel time does not matter because it is close to zero whatever the location of the emission of the event. But when the students take it into account, they consider that the order in which two events occur is a consequence of the order in which these two events are perceived as if causality would apply from future to past [5]. In our environment

[2] To avoid overloading the simulation and affecting the understanding of the scene by superimposing effects of very different nature, we have not implemented the Doppler effect nor the effects of changes in light intensity. Indeed, we initially limited the rendering to purely geometrical effects (space-time and related concepts). The Doppler effect is certainly an effect of space-time, but its manifestation depends on the physical nature of light (electromagnetic wave which actually has a frequency...). In our approach, for now, we do not question the luminous phenomenon itself but the space-time architecture of the physical reality.

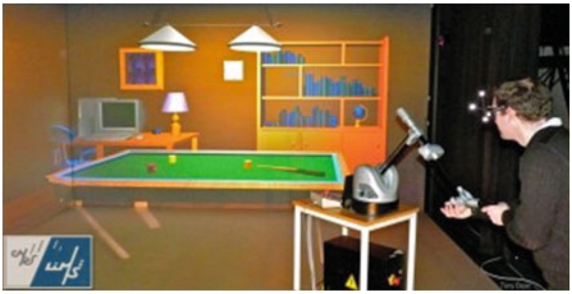


Fig. 37.4 The carom billiard running in the EVE CAVE (CNRS/LIMSI, Orsay, France)

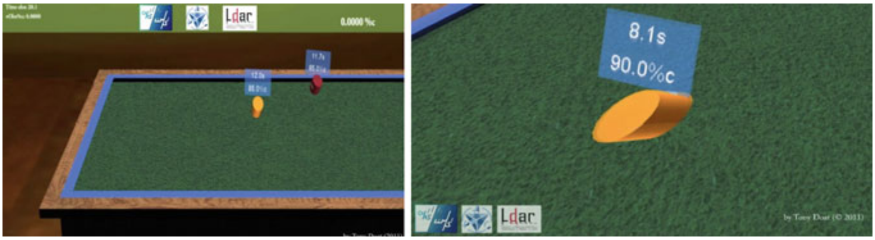


Fig. 37.5 On the *left* screen-capture the apparent collision delay is observed; one the *right* screen-capture the observer can see the object deformation as an effect of the consideration of the finite value of c. The photons that arrive into the eye of the observer have not been emitted at the same time. The image seen is not the result of the emission of photons emitted simultaneously

the signal travel time cannot be neglected (c is fixed to 1m/s), thus mundane objects are not seen as in everyday life. The simulator we designed takes into account the relativistic effects (simulation algorithms based on Lorentz transform, [7]) and also those due to the light propagation. Then we make the hypothesis that the users of our simulator will take into account the time delay between emission and perception in a relevant way.

According to our first scenarios, users (who also are observers) are asked about changes in the shape of the pucks and about the changes in their velocity. We also question them about the instant of the contact of the pucks with the billiard table. The corrected proper time of each puck is visible. It represents the time measured by a clock located in the puck itself. The delay of reception of the photons by the observer (who is actually not located where the time is measured) is taken into account. Thus, according to the movement of the puck, the perceived time seems to pass faster or slower. Moreover, the movement of the puck can be "freeze" so that the puck is seen using the "Matrix" effect: a camera turns around the puck showing each part of it without changing the initial point of view which is the observer's one.

The same scene can be replayed but as seen by an observer shifted on the left (or on the right). Then, two pucks (of two different colors) are in movement perpendicularly to the billiard table (Fig. 37.5). The observer can be located at equal distances between the red puck and the orange one. He can also moves to the left (or to the right) breaking the symmetry of the distances. All the effects observed are discussed.

37.5 Conclusion and Perspectives

Five users were immersed in the simulator and had to explain what they saw when confronted to the learning pathway. We undertook a lexical analysis searching for some specific conceptual elements that we consider as key points for the users' understanding. These elements are the following: (a) the incoming of light in the eye, (b) the finite nature of the speed of light, (c) the distance between the pucks and the user, (d) the object discretization (a set of points as punctual sources of photons), (e) the geometrical relativistic effects explained by the Lorentz transform. Following users' reactions during the immersion, we detected which situations favour the emergence of the conceptual elements that allow a relevant interpretation of the situation. As a conclusion the immersion of users in the 3D environment where they are confronted to relativistic phenomena favours lines of reasoning that take into account the light travel time and the arrival of light into the users' eyes.

We believe that using our application to experience these effects "without thinking" will help to develop intuition on relativistic behaviours while trying to play billiard properly at relativistic velocities. It is expected to help students in their efforts to understand Einstein's theory from a practical point of view. This will be tested by the EVEILS group through a dedicated research work in formal evaluations on physics students.

References

1. Savage CM, Searle A, McCalman L (2007) Real time relativity: exploration learning of special relativity. Am J Phys 75:791–798
2. Saltiel E, Malgrange JL (1980) Spontaneous ways of reasoning in elementary kinematics. Eur J Phys 1:73–80
3. Villani A, Pacca JLA (1987) Students' spontaneous ideas about the speed of light. Int J Sci Educ 9(1):55–66
4. Scherr R, Schaffer P, Vokos S (2001) Student understanding of time in special relativity: simultaneity and references frames. Am J Phys 69:24–35
5. de Hosson C, Kermen I, Parizot E (2010) Exploring students' understanding of reference frames and time in Galilean and special relativity. Eur J Phys 31(12):1527–1538
6. Biseci E, Michelini M (2008) Comparative teaching strategies in special relativity, GIREP. Conference, August 18–22, Nicosia
7. Doat T, Parizot E, Vezien J-M (2011) A carom billiard to understand special relativity. Virtual reality conference (VR), IEEE, pp 201–202
8. Hewson PW (1982) A case study of conceptual change in special relativity: the influence of prior knowledge in learning. Int J Sci Educ 4(1):61–78

Chapter 38
Stories in Physics Education

Federico Corni

Abstract Narratives are at the basis of every effort of making sense of human experience. Stories are a particular type of narrative particularly suitable for children because they involve them affectively and cognitively. This contribution deals with the opportunity of using stories for physics education in primary school. It is introduced a disciplinary approach suitable to be shaped into stories, relying on cognitive linguistics results, as well as some elementary features of the stories are compared to the fundamental caracteristics of scientific inquiry. Hints are proposed for possible research programs and some examples of stories, coherent with this program, are presented and discussed.

38.1 Introduction

It is out of doubts that children love stories. A fact presented to children in form of a story gains a great attraction and becomes very interesting to their eyes. The emotions that stories stimulate make the children fully involved, not only affectively, but also cognitively [1]. Stories, however, are not only to be mentioned when speaking about children, since they represent, in a more general meaning, the way humans express their thought through language. As the stories frame the narrated facts into a certain structure and with a certain sequence, in an analogous way also concepts and thoughts are explicated within a certain structure and sequence. When we explain a scientific theory or a model, we organize our talk in a suitable way, with assumptions, logical connections, references, and support of experimental observations: everything follows a coherent sequence, necessary and sufficient to support the thesis. A mathematical theorem is expressed as a logical path we have to follow, starting from the hypotheses, to show the evidence of the thesis. From this point of

F. Corni (✉)
Department of Education and Humanities, University of Modena and Reggio Emilia, Emilia-Romagna, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_38,

view, stories and theories or models in physics are not two distinct species; rather they are the ends of a spectrum of how humans create meanings.

From this qualitative consideration, we could start to fund the use of stories in physics education as form in which to shape the disciplinary contents and the way they are presented and taught to children.

Not every type of story, however, is adequate to host science contents, as well as not every approach to science is suitable to be framed into a story form. On one side, the contents to be delivered have to be suitable to be shaped in narrative form; on the other side, the stories have to meet certain conditions to be suitable for the construction of meanings.

In this contribution both aspects will be treated, that of the stories and that of the contents: in the next two sections will be introduced a disciplinary framework coherent with the basic forms of language as they are recognized and studied in cognitive linguistics, then will be evidenced features of stories that make them akin to the process of scientific inquiry; in the further next section will be proposed general hints for possible research programs and experimentations of the production and use of stories for physics education; finally, will be illustrated and discussed examples of stories developed according to this program.

38.2 The Disciplinary Framework for Physics Education in Primary School

Cognitive sciences and cognitive linguistics in particular [2, 3] evidence recursive and pervasive patterns, called image schemata, we use to express our thought with language and to understand the world. An image schema is a recurrent pattern, shape and regularity in, or of, our actions, perceptions and conceptions [2]. We extract these structures from bodily experiences: they are developed very early in the life and are already present in a rudimental form also in small children. They are simple yet have enough internal structure so that more elaborate mental structures can be built upon them. In other words, they make reasoning possible [2].

Image schemata are more general than every particular mental image we form; this means that any effort to represent them, even schematically, trivializes them. We have the concept of triangle, but whenever we draw a triangle, we reduce its generality. Moreover, image schemata are more basic than any proposition, explainable with a definite statement; rather they emerge with analog fashion from our thought.

These recurrent patterns are relatively few in number and we use them in a variety of situations. In synthesis, image schemata are real abstractions, gestalts: structures simpler than the sum of their parts. Table 38.1. reports the most important image schemata evidenced by Johnson [2], Croft and Cruse [4], Evans and Green [5].

Children, in this sense, are abstract thinkers, since they make sense of reality employing image schemata, the same used by adults, just less featured or differentiated. They and we speak about natural, emotional, and social phenomena using

Table 38.1 Main image schemata form specific literature

Polarity	light-dark, warm-cold, female-male, good-bad, just-unjust, slow-fast, high-low
Space	up-down, front-back, left-right, near-far, center-periphery, other:contact, path
Process	process, state, eyele
Container	containment, in-out, surface, full-empty, content
Force/Causation	balance, counterforce, compulsion, restraint, enablement, blockage, diversion, attraction
Unity/Multiplicity	merging, collecting, splitting, iteration, part-whole, mass-count, link
Identity	matching, superimposition
Existence	removal, bounded space, object, substance, fluid substance

metaphorical projection of image schemata [6]. The abstract elementary concepts developed by the children' minds should not be regarded as obstacles for scientific thought; on the contrary, these concepts share the same seeds of adults' formal science and are to be regarded as resources to build upon. Fuchs [7] has identified an important gestalt, named Force Dynamic Gestalt (FDG), particularly relevant to scientific understanding. It is structured with three main aspects that we use to understand the most diverse complex phenomena, from psychological ones-fear, happiness, pain-to social ones-justice, the market-, and that can be strictly related to basic and simple concepts useful to explain the behaviour of various natural phenomena, such as those involving fluids, electric charge, heat, motion, and chemicals.

The FDG has the aspects of fluid substance, polarity of verticality, and force. The image schema of fluid substance has the character of an amount or quantity, and is related to the scientific concept of extensive quantity. The verticality image schema has the character of intensity or scale, and corresponds to the scientific concept of intensive quantity or generalized potential. Finally, the image schema of force, that summarizes various more simple schemata of direct manipulation (see Table 38.1), is related to causation and the scientific transversal concept of energy [7]. So, we think to and speak about the increase of temperature of a pot of water while it is heated on a heater in the same way we think to and speak about the increase of the level in a glass while we pour water, or in the same way we think to the rise of electrical potential while we charge a capacitance. Again, we refer to heat as a fluid substance that flows from hot to cold things in contact, in a similar way as we speak about water or air that flow from a high-pressure vessel connected to low-pressure one. Water, heat and electric charge, in these examples, are conceptualized as fluid substance-like quantities with a qualitative vertical scale of intensity.

The physics contexts that are usually regarded as distinct fields of knowledge, especially in the school curricula, with this approach are seen in a unitary fashion, evidencing their analogical structure of contents. Every context has its characteristic extensive quantity that accumulates into containers; the current of such quantity flows

Table 38.2 Correspondence between the fundamental concepts of the contexts relevant for primary school

Context	Substance-like	Potential	Capacitance	Current	Resistance
Fluids	Volume	Pressure	Section	Current	Hydraulic impedance
Heat	Entropy	Temperature	Specific heat	Thermal current	Thermal resistance
Electricity	Electric charge	Electric potential	Capacitance	Electric current	Electric resistance
Motion	Momentum	Velocity	Mass	Force	…

through the cotainer boundaries directed by a negative difference of the conjugate generalized potential; the scientific concept of capacitance relates to the way in which the potential increases when the extensive quantity accumulates into a container; and the concept of resistance relates to the way in which the potential difference increases to increase the current intensity. Table 38.2 shows the analogical correspondence between the main experience contexts relevant for primary school physics education.

At primary school level, physics education has the general aim to help children to identify, differentiate, and recognize and master the relationships among the aspects of the FDG.

38.3 Stories and Scientific Inquiry

The connecton between image schemata and phenomena are the metaphors we produce, through imagination, expressing the affinity among the various experience contexts. When we want to express our thought to ourselves or to someone else we are forced to put in a narrative form (verbal and sequential) what is an imaginative pattern in our mind. At the basis of every effort of making sense of human experience there is the narrative thought, which searches for meanings and establishes relations [8].

Narratives are spontaneous and natural tools, very appropriate to the human being. Historically they appeared with the development of oral cultures, well after writing. In the child they appear very early, without external influence or instruction. Over time the child refines his ability of employing this tool and learns the mechanisms that regulate narratives in order to produce them more and more effectively. Many research efforts are devoted to the study of narratives, especially by humanists, but also by scientists.

A particular kind of narrative is the story form. We know that children love stories: let's discuss if and how such a tool may be useful for physics education purposes.

The most common and basic features of stories can be summarized as follows [9, 10].

1. Stories have a general structure called story grammar or story schema. If we listen to someone speaking, not necessarily we feel the speech as meaningful and

interesting for us, and this even if he is treating a particular argument and does not beat around the bush. A narrative results meaningful for us when it is dressed in a form that attracts our attention, that is affectively attractring. Stories build upon, engage, and strengthen affective meaning. Basically, a story consists in a beginning that is set up by a problem created by a polarity, a middle part where the problem is elaborated, and an ending where the problem is solved. It is ultimately a technique for organizing events, facts, ideas, characters, and so on, whether real or imagined, into meaningful units that shape our affective responses. As the events of the story proceed, the story tells us how to feel about the whole picture.
2. A story tells about a limited world, with a created and given context. It is a simplified and delimited environment where events stand out and are clearly focused so as we can grasp the necessary information and search for their meaning. The story form provides a suitable environment to exercize imagination. Things become meaningful within contexts and limited spaces.
3. Stories are narratives that have an end. A story does not necessarily end when we are told that all lived happily ever after, but when we have grasped all the presented events and situations and we know how to feel about them. A narrative is a story because its ending balances the initial tension and completes all what was elaborated in the middle. Stories grant us the satisfaction of being sure how to feel about events and characters.

To answer the question about the utility of stories for phyisics education we can draw a sort of parallel between the peculiarities of the story form listed above and those of scientific inquiry.

1. An investigation is an action that developes in time. It starts from a problem, it provides for the elaboration of information, and ends with a solution of the problem. Not all data are appropiate for a given inquiry, but only those ones that are coherent with the need of clarification and rational understanding the problem rises. An inquiry is the story of the information that involves us cognitively. Affectivity and cognition are two ways of grasping things.
2. Within an investigation, the interest is focused and the object of the study is limited. In order to converge to a solution, the investigation field must be defined and limited, the phenomenon object of inquiry have to be insulated form the rest of the world, and, in some cases, it can be reproduced and studied in the laboratory. A circumscribed world and the laboratory work favor the research.
3. An investigation does not end arbitrarily, but when all needed information is known and it is organized in a coherent and meaningful picture. We do not feel to have finished an inquiry until we have arrived to a satisfactory solution of the initial and of all open problems. Completeness and coherence are the final gaols of any inquiry.

As a story is a search for affective meaning in a context, analogously, science is a search for rational meaning of a phenomenon. As, over the primary school years, children interests evolve from affective to rational values, the story form can fairly be regarded as the suitable environment to match both kinds of meaning. Stories could be used as a vehicle to guide the children in their growth between the affective needs of early childhood and the developing interest in rational and scientific understanding of the world of higher ages. Stories can be employed at all grades of primary school as transversal educational tools that develops together with children.

38.4 Hints for Education Research

Joining the disciplinary approach to physics and the methodological value of stories discussed in the previous two sections, we may argue that well crafted stories can be tools useful to clarify and make apparent the aspects of the FDG and their relationships about an object of interest. This makes our thought clear to ourselves and to the listeners of our stories.

Telling good stories about natural phenomena, i.e., using a correct natural language for it, is the first step toward a useful conceptualization of the processes of nature. A story is in general narrated in natural language, but if one seeks to clarify the aspects of the FGD, he needs and makes use of a language which becomes increasingly clear, accurate, unambiguous, and, finally, formal.

For the purpose of science education, stories can be either told by the teacher, or told by the children, or both. The fauncions are different, of course, and fairly important. In the rest of this paper stories will be treated in their particular value for the teacher side.

Just a little parenthesis about children storytelling. We have observed that children are competent narrators even in the early childhood: stories can be used by the teacher as instrument for the assessment of the children learnings. This is a valuable function of stories. In fact, a scientific education driven by the cognitive processes that occur in the children minds rather than oriented to the mere acquisition of contents, as outlined in the second section, cannot be assessed with the conventional instruments (tests, prolems, …) suitable to quantify notion knowledge and scientific competence. The children thought can be accessed and qualitatively evaluated through the analysis of their discourses, through the observation of the language naturally employed to explain their stories of any type (concerning fantastic facts, or explaining an experimental inquiry). In this sense, grids of language analysis, either from a syntax point of view, or from a semantic point of view, can be useful [11, 12].

Back to our focus, stories to be told by the teacher in primary school have to be suitably designed, with developing structure, content, characters and language according to the children mind development and age.

Research efforts should be addressed to the development of the story form from early childhood up to the last years of primary school. With increasing age of the children:

1. the story structure should change from simple and elementary to articulated and complex, hosting pupils activities (experiments, texts, drawings, games, …) and teacher intervention (discussions, explanations, …)

Fig. 38.1 Scenes of the three stories of the "Piccoli Scienziati" project

2. the content should concern increasningly abstract topics (water, food, …, heat, …, energy)
3. in the first years, the affective characters of the story can be fantastic characters, then, in the last years, some of them could remain as friends of children, but the effective central characters should become the natural phenomena and the physical quantities
4. language should become more and more precise and specific to the particular cases, as well as decontextualized referring to general conceptual organizers.

This research program contributes to overcome the notion of a dichotomy between narrative and paradigmatic thought. In contrast to Bruner [8], narrative and paradigmatic forms of understanding represent a polarity that opens up a spectrum between the poles of narrative and paradigmatic thought. At the same time, this program works in the direction of giving a unitary fashion to primary school instruction, and of strengthening the marriage between sciences and humanities.

In the next section, will be synthetically presented the stories produced within the project "Piccoli Scienziati" (Little Scientists), developed at the University of Modena and Reggio Emilia in the last three years, proposed to a large number of teachers in training courses in various Italian Regions, and experimented by the teachers themselves in their classes (see Corni et al. [13]).

38.5 The Pico's Stories

Lots of examples of use of stories are present in science education literature, most of them with historical background [14–23].

The Pico's stories [13]—*Pico and his friends at the Luna Park*, *Rupert and the dream of a swimming pool*, and *The Rupert efforts and the mountain trip* (Fig. 38.1)—are designed to be used at the various levels of primary school according to the above research program.

They are animated stories in slideshow format that tell the adventures of some friends in various situations. They are corredated by two cases full of materials–toys, simple models, devices of everiday use–for experimental and laboratory activities. Each character of the stories has evident features in which a child, according to his temperament and disposition, can easily identify with. Pico is the archetypal child, lively and creative, who may become a scientist. He acts as mediator between the story world of characters and the real world of children. He is always present in the story scene and promptly intervenes in the problematic situations. He poses the right question at the right time and involves children in the solution inviting them to make hypotheses, draw pictures and schemes, perform experimental activities with the help of the materials of the cases. A second key character is Blackbird that flies high and looks at the situations from different points of view. Rupert, a frog, pretty, ingenuous and a bit unlucky, is the carefree child that, together with his bear friends Thomson and Aielmo, undergo to various vicissitudes and adventures.

In the folowing, the features of the three stories will be synthetically presented with reference to the variables evidenced in the previous section concerning the education research. The structure concerns the complexity of the scene and the presence of children and teacher activities, the content relates to the aspects of the FDG, the characters are the effective entities over wich the children attention is focused, and the language concerns the linguistic competencies needed.

38.5.1 Pico and His Friends at the Luna Park (1st–2nd Grades)

The friends buy ice creams, exchange ice cream balls and reflect on the quantities and flavors. Then they buy drinks and reflect on their quantities and on their levels in various glasses.

Structure. The scene is still and and centered on the friends.

Content. Identification and differentiation of the aspects of quantity and intensity of substances.

Characters. The friends who want to have fun at the Luna Park. The story is centered on their wishes.

Language. Elementary and with common terms.

38.5.2 Rupert and the Dream of a Swimming Pool (3rd–4th Grades)

Rupert wishes to place a swimming pool in his garden. It positions the pool at various heights and tries to fill it by connecting with pipes to a nearby aqueduct.

Structure. The scene is structured and focused on essential elements (aqueduct, pool, pipe). Children are invited to do experiments with a model apparatus.

Content. Identification and differentiation of difference of potential, current, resistance and their relationships in the context of water.

Characters. The Rupert's problem to be solved.

Language. Elementary with the introduction of terms of common language like to empty the pool, water flow, water level.

38.5.2.1 The Rupert Efforts and the Mountain Trip, Part I (4th–5th Grades)

Rupert struggles to push a wheelbarrow full of flower pots to emebllish the pool. He tries various strategies, using the wind and an inflated balloon.

Structure. The scene is complex and offers various cues. Children are invited by Pico to study and explore (following some worksheets) a hairdryer and toys (a car and a balloon) to understand their functioning and to find the way to better employ them to obtain a result.

Content. Identification, differentiation of differences of intensive quantities and currents of substances in interactions (wind generated by the hairdryer with the car; air blowon by an on board inflated balloon with the car); relationships between current and changes in potentials of the cause (wind) and of the effect (motion).

Characters. The wind and the car in the experimental activities. Rupert and his wishes are on the background.

Language. Precise and specific to express clearly the various conditions met.

38.5.2.2 The Rupert Efforts and the Mountain Trip, Part II (Second Part)

Rupert and his friends go for a mountain trip and see and experience various natural phenomena (rivers flowing downhills, a melting glacier, a dam with an artificial lake, a landslide, operating wind mills, a lightning during a storm, food, photovoltaic panels, …). The story is an adventure to discover nature.

Structure. The numerous scenes are rich and offer many cues. Children are invited by Pico to discuss about natural phenomena and their effects on one another.

Content. Decontextualization and generalization of the relationships between differences of potential and currents in the interactions.

Characters. Natural phenomena. Rupert and his frined are on the background.

Language. Specific to every phenomenon, but more and more decontextualized referring to potencials and currents.

38.5.3 Results of Experimentations with Children

The results of the use of the swimming pool story in fourth grade classes have been evaluated [24] with reference to the following research issues:

1. children involvement in problem solving,
2. transition from the description to the interpretation levels,
3. identification of the relevant variables,
4. generalization of meanings.

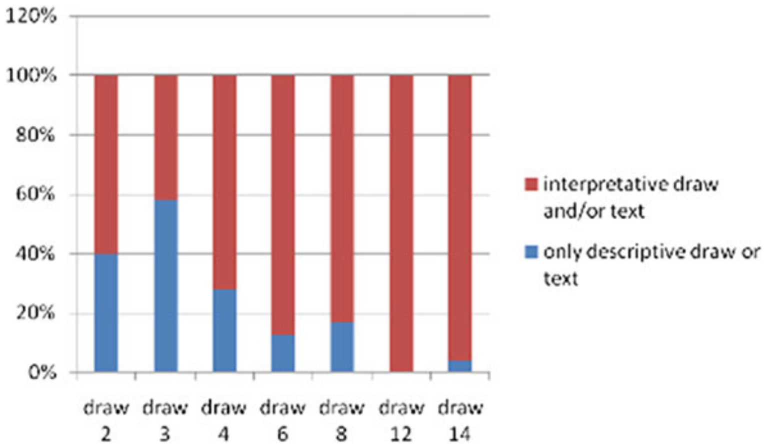


Fig. 38.2 Percentages of descriptive and interpretative children' drawings and texts

38.5.3.1 Children Involvement in Problem Solving

By analyzing the children's drawings evolution, context items are less and less present and sidelined in the background, while elements that relate to the problematic situation are highlighted and centered. The drawings made by each child evidence a gradual ability to identify and focus on significant story objects.

38.5.4 Transition from the Description to the Interpretation Levels

The analises of children drawings and texts show (see Fig. 38.2) a gradual shift from phenomenological description to formulation of hypotheses and interpretations.

38.5.5 Identification of the Relevant Variables

During the evolution of the story, children begin to identify some concepts, mainly those of water level difference, current and resistance. The learning analysis of every child shows a generally increasing trend with some fluctuations.

38.5.6 Generalization of Meanings

Generalization is a high goal and it cannot be reached effectively with a single activity or series of activities limited in time. However, two children in particular have developed a remarkable level. They represent and explain by words that if the pipe is full of water, its shape does not influence the water flow, even if in some points the pipe is lifted above the water level of the aqueduct. This conclusion is not directly deducible from the story and the provided activities with the hydraulic model.

38.6 Conclusions

Support for the use of stories in physics education has been outlined. Research in this direction must be conducted in order to define how to design stories, how to employ them in the didactical practice, and what results can be obtained. The project "Piccoli scienziati" has started a research program since three years ago and some encouraging results are now being obtained.

References

1. Egan K (1986) Teaching as story telling an alternative approach to teaching and curriculum in the elementary school. The University of Chicago Press
2. Johnson M (1987) The body in the mind. University of Chicago Press, Chicago
3. Lakoff G, Johnson M (1999) Philosophy in the flesh. Basic Books, New York, NY
4. Croft W, Cruse DA (2004) Cognitive Linguistics. Cambridge University Press, Cambridge, UK
5. Evans V, Green M (2006) Cognitive Linguistics.An Introduction. Edinburgh University Press, Edinburgh
6. Lakoff G, Johnson M (1980) Metaphors we live by, University of Chicago Press
7. Fuchs HU (2007) From image schemas to dynamical models in fluids, electricity, heat, and motion. ZHAW, Winterthur, Switzerland. https://home.zhaw.ch/~fusa/LITERATURE/Literature.html, accessed december 2011
8. Bruner J (1986) Actual Minds. Cambridge, Harvard University Press, Possible Worlds
9. Mandler JM (1984) Stories, scripts, and scenes. Psychology Press, New York, NY
10. Egan K (1988) In: De Castell S C, Luke A, Luke C (eds.) Language, authority, and criticism: readings on the school textbook. London, Falmer
11. Laurenti E, Mariani C, Corni F (2011) A qualitative look on children's conversations about processes involving the concept of energy. In: GIREP 2011 Proceedings
12. Mariani C, Laurenti E, Corni F (2011) Hands-on, minds-on activities to construct the concept of energy in primary school. In: Experiments, games and group discussions, ICPE 2011 Proceedings
13. Corni F, Mariani C, Laurenti E (2011) Innovazione nella didattica delle scienze nella scuola primaria: al crocevia fra discipline scientifiche e umanistiche. Artestampa, Modena
14. Jung W (1994) Toward preparing students for change: a critical discussion of the contribution of the history of physics in physics teaching. Sci Educ 3(2):99–130
15. Kipnis N (2002) A history of science approach to the nature of science: learning science by rediscovering it. Nat Sci Sci Educ 5(II) 177–196
16. Tselfes V, Paroussi A (2009) Science and theatre education: a cross-disciplinary approach of scientific ideas addressed to student teachers of early childhood education. Sci Edu 18(9):1115–1134
17. Adúriz-Bravo A, Izquierdo-Aymerich M (2009) A research-informed instructional unit to teach the nature of science to pre-service science teachers. Sci Educ 18:1177–1192

18. Klassen S (2010) The relation of story structure to a model of conceptual change in science learning. Sci Educ 19:305–317
19. Kubli F (2000) Can the theory of narratives help science teachers be better storytellers? Sci Educ 10:595–599
20. Lehane S, Peete M (1977) The amazing adventures of erik stonefoot. Language Arts 54(4):395–400
21. Lebofsky LA, Lebofsky NR (1996) Celestial storytelling. Science Scope 20(3):18–21
22. Ellis B (2001) Cottonwood: how i learned the importance of storytelling in science education. Sci Child 38(4):42–46
23. Sallis DA (2008) Humorous cartoons for teaching science concepts to elementary students: Process and products. First Annual Graduate Student Research Symposium. University of Northern Iowa, Cedar Falls
24. Corni F, Giliberti E, Mariani C (2010) A story as innovative medium fro science education in primary school. In: GIREP 2010 proceedings

Chapter 39
How Physics Education Research Contributes to Designing Teaching Sequences

Jenaro Guisasola

Abstract Over the last decades, a growing number of physicists have taken up the challenge of implementing the same research discipline in physics learning and teaching as applied to traditional research in the field. This commitment is widely known as "Physics Education Research" (PER). In this paper, I will single out some of the directions taken by present and emerging research that I deem promising. I will also discuss the impact of research on the educational practice of physics teaching at university level. This paper presents evidence from different studies to demonstrate the potential positive impact of research into teaching and learning physics on students' understanding of physics. Finally, I will show some practical challenges and propose some steps that could be taken to ensure PER growth and productivity.

39.1 Introduction

In recent decades a growing number of physicists have taken up the challenge of applying an approach to physics education research, just as rigorous as traditional physics research, concerning problems relating to learning and teaching physics. This commitment is widely known as "Physics Education Research" (PER). PER concentrates on understanding and improving how physics is learnt by studying the contents of the physics curriculum and what teachers and students do when teaching and learning in schools. The research field relating to teaching science, and in particular Physics Education, has been well-established for some decades. It attempts to integrate knowledge from different fields of research; such as physics, the psychology of learning, the epistemology of science or the pedagogies of the teaching–learning process [1, 13] in a non-mechanical way.

J. Guisasola (✉)
Applied Physics Department, University of the Basque Country, Donostia, Spain

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_39,

The objective of "improving" student learning lies at the heart of all education research. As Ref. [9] says 'there is a little reason to do research, unless there is a pay-off in the classroom'. PER has developed around intensifying the belief that this goal is possible. Its objective is usually formulated as "reforming or changing the teaching of physics" causing PER to go beyond identifying student learning difficulties found in traditional teaching. Research has developed didactic materials and strategies, which have been repeatedly submitted to examination, assessment and redesign. A number of different studies have had a positive impact on physics teaching and learning. One example is research into student's difficulties in terms of learning physics concepts which resulted in designing new instruments to assess students' knowledge and the effectiveness of teaching. Halloun and Hestanes [8] used the results obtained from an investigation into university students' ideas, within the field of Mechanics, to design the "Force Concept Inventory" formative assessment test. Since its design, there has been an increase in the number of physics programmes and textbooks, which pay greater attention to conceptual differences. Recently, in the field of electromagnetism, Maloney et al. [12] have developed formative assessment instruments with similar aims.

However, difficulties in carrying out research into teaching science, aimed at improving the practice of teaching, must not be underestimated. Lijsen and Klaassen [11] argue that designing learning sequences requires a complex process of applying the general principles of didactics to specific teaching contexts for teaching the subjects on the curriculum. They point out that this task is not linear but rather a cyclical process with the aim of generating knowledge about teaching and learning, implementing improved teaching methods appropriately in the classroom. Designing teaching sequences is not a mechanical process involving transferring pedagogical principles and research results to teaching specific science subjects. On the contrary, teaching sequence design is a creative process, which considers not only research but also the classroom culture and the circumstances of both teachers and pupils.

In this paper I will discuss the impact of research covering educational practice of teaching physics on university courses. I will look at some practical challenges and propose some steps that could be taken to ensure PER growth and productivity.

39.2 Implications of P.E.R. to Designing Teaching Sequences

Designing research-based teaching sequences takes into account two kinds of research recommendations: (a) results of empirical studies on students' ideas and reasoning; (b) technical contributions connected to the nature of science and how it is learnt and taught. Both contributions are connected, as the principles deriving from the latter influence how empirical studies are analysed on the former. Analysing students' ideas involve not only conceptual aspects but also epistemological and ontological aspects. On this point, it is necessary to analyse the historical development of the topic to be taught, the difficulties that the scientific community had to overcome and the arguments used to construct new concepts and explanatory models. Working from this epistemological analysis on the scientific content of school curriculum, it is possible to define the teaching–learning aims in a well-founded way. In other words, it is possible to justify choosing these aims based on epistemological evidence of the discipline and not idiosyncratically or based on the educational programme's tradition.

The research recommends sequencing the main stages that teacher must work through when designing the teaching programme. In our research group we use the so-called "learning indicators" to specify what students should learn on the topic in accordance with the school curriculum [4, 5]. The learning indicators include ontological aspects (values and attitudes) that must be consciously taken into account. Research into teaching sciences shows that the emotional and value-related aspects cannot be considered without making a close connection to cognitive processes when students are working on their activities in science classes [19]. In this respect, designing activities that relate aspects of science, technique, society and the environment to each other means supporting a presentation of socially contextualised science that encourages students' interest in the scientific topic being taught. Including activities related to Science-Technology-Society-Environment generates interest among students on the study topics, encouraging them to get involved in the solving task [18].

The design and development of the research-based teaching sequences imply that creative integration must be carried out taking into account the teaching difficulties, learning indicators and ontological aspects. This creative integration leads to a close analysis of the differences between the learning difficulties and the teaching targets set in the curriculum. This analysis should build bridges between activity design and research work. In this respect we can say that we are using evidence from the empirical research when designing the sequence. It is necessary to highlight that the curriculum standards provide information on "what should be taught" generally. On the contrary, teaching aims for the panel that we have defined using the research evidence specify even further what students should learn and justify why they should do it.

Although many of the activities from the innovating educational sequences are common to the questions and exercises from the text books used in regular teaching, it is necessary to highlight that they are used differently. Maybe the most significant differences revolve around the time dedicated to student difficulties, plus activities that aim to interest the students on the topic and justify introducing new models and concepts. The sequence activities are designed with the aim of providing students with opportunities to understand and apply the same model repeatedly. On the other hand, the activities also tackle epistemological aims by getting students to appreciate the power of a scientific model capable of explaining a great number of experimental cases. As a conclusion, all the above can be organised into Table 39.1.

Table 39.1 Use of research evidence to design teaching sequences

Students' ideas and reasoning	Epistemological analysis of the contents of the school curriculum	Interests, attitudes, values and standards
Difficulties in learning	Learning indicators	S-T-S-E Aspects
Teaching goals		
Set out specific problems and aims in a sequence		
Interactive learning environment		
Teaching strategies		

39.3 Teaching Sequence of Electromagnetic Induction (EMI)

Over recent years, within the Physics Education Research Group of the University of the Basque Country (PERG-UBC), I have carried out different research projects, which have developed teaching sequences for classroom implementation and subsequent assessment as one of their main objectives. See, for example, [2, 4–6]. In this paper, I will describe the processes involved in designing teaching sequences following the instruments within chart 1, with specific reference to teaching Electromagnetic Induction (EMI) in an Introductory Physics Course at university level.

Regarding students' ideas, previous works [3] have identified their conceptual and epistemological difficulties in understanding the theory of EMI. On this point, it is necessary to analyse the historical development of the topic to be taught, the difficulties that the scientific community had to overcome and the arguments used to construct new concepts and explanatory models. Working from this epistemological analysis of the scientific content of the school curriculum, it is possible to define the learning indicators and justify their choice based on epistemological evidence from the discipline and not idiosyncratically or based on the tradition of the education programme. We will present the learning indicators drawn up for teaching EMI given in Table 39.2 below:

Analysing the differences between the learning difficulties and the teaching aims set provides us with bridges between sequence design and the research work. In this respect we can say that we are using evidence from the empirical research when designing the sequence. Table 39.3 below shows the sequence for EMI topic.

Although the sequence design is strongly supported by evidence from research, it is necessary to assess how it is implemented in relation to learning indicators. This means that talking about teaching sequences based on the evidence from research involves assessing their implementation. This aspect will be mentioned in the next section.

Table 39.2 Description of learning objectives and difficulties in EMI topic

Learning indicators	Students' difficulties
(i1) Be able to weigh up how useful it is to solve the proposed problem	- EMI presentation is not significant to involve students in the study
(i2) Be familiar with the experimental phenomena of electromagnetic induction	- Lack of familiarity with the scientific lab methodology (compiling data, handling apparatus, analysing results)
(i.2.1) Find out about electromagnetic induction phenomena in spirals and solenoids crossed by variable magnetic fields	- Difficulties to distinguish between the empirical level (use of multi-meters and interpretation of measurements) and the interpretative level that uses concepts such as variable magnetic and electric fields over time, Lorentz Force, magnetic flow and electromotive force
(i.2.2) Find out about electromagnetic induction in circuits that are moving within a stationary magnetic field	-Not used to working in a group
(i.2.3) Find out about electromagnetic induction phenomena caused by a combination of the aforementioned effects	
(i.2.4) Be able to tell the difference between the phenomenon of producing induced electromotive force and generating induced current intensity	
(i3) Be able to analyse electromagnetic induction phenomena qualitatively and later quantitatively	- Students find it hard to: a) Interpret simple induction phenomena properly. b) Attribute EMI to the presence of a variable magnetic field or to the variation of magnetic flux c) Apply Faraday's Law correctly, interpreting motional electromotive force phenomena
(i.3.1) Using macroscopic modelling correctly	- Students tend to use a field model more frequently than a force model when explaining EMI phenomena
(i.3.2) Using microscopic modelling correctly	-Few students show that they know how to explain the phenomena from a microscopic and macroscopic point of view
(i4) Repeatedly use scientific work strategies to find the solution to the set problems	- Lack of familiarity with the scientific methodology in problem solving. Difficulties to: ○ Carry out qualitative analysis ○ Make a hypothesis ○ Draw up alternative strategies ○ Analyse results ○ Handle apparatus - Not used to working in a group

Table 39.3 Map of the didactic sequence for the study on EMI

Problem sequence	Science Procedure to be learnt by the students	Explanations for the students to understand
What is the interest and/or use of the EMI study	- Science is interested in natural phenomena and their social implications. The origin of the interest in a phenomenon, situation or fact can vary depending on: the science topic, technological problems, local or global problems, etc	- Acquiring a preliminary conception of the study that is going to be carried out and promoting interest on the EMI study
When does an EMI phenomenon occur and when does it not	- Make empirical observations and take down information on the phenomena that occur and make predictions on what might happen	- Macroscopic and qualitative study of EMI. EMI occurs when there is a variable B over time and/or when a conductor moves in a magnetic field. There is no induction if the conductor is in a stationary magnetic field
How can the EMI be quantified	- Science is able to measure the phenomena that are seen and give quantitative answers from the phenomena seen	Macroscopic perspective of EMI
		- Faraday's Law
		- Induced emf
		- Induced emf $\neq$ induced I
		- Energy conservation law (Lenz Law)
Is there another way of measuring induced emf	- Science can solve the same problem using different laws and points of view. A problem can be solved with different procedures and get the same results	Microscopic point of view for EMI
	- Find out about the field of application for the laws	- Lorentz Force
		- Magnetic force
		Non conservative electric field
		- Relationship between emf with Lorentz force and the conservative electric field
		- Reference systems and the fields
Applications for EMI	- The applications of science and technology in everyday life meet our needs but they are also present in our leisure. Science and its technological applications are all around us	Know how to apply Faraday's law within the context of your everyday life beyond the school context

39.4 Assessment of Teaching Sequence and Conclusions

Physics education research, generating relevant knowledge about teaching science subjects, presents significant difficulties. Firstly, considering that one proposal is "better" than another involves agreeing with the aims used to assess the quality of the proposal. These quality criteria may be based on the percentage of students who pass official internal and external tests which may have some correlation with the students' results in conceptual comprehension tests. In other words, quality may be measured in terms of the number of students preparing to following science and engineering studies and in the proper preparation of this elite group. An alternative method of measuring quality revolves around how effective it is at generating better scientific literacy and increasing the wider understanding of basic scientific theories. Finally, the nature of the teaching quality is a question of values, concerning educational administration and, finally, the teachers responsible for implementing it. In any event, if different quality criteria are applied in different situations, it is not possible to identify a single "best teaching practice'.

Another important difficulty in generating knowledge relevant to teaching is demonstrated by evidence from research [17] showing that teachers make changes when implementing the curriculum which may affect original intentions, and envisaged goals. This may lead to teaching veering off track from its "official" goals. In addition, extensive research into teachers' thinking [15] shows that teachers have a positive attitude towards the results of didactic research, but are not prepared to change how they teach if the actions proposed are not consistent with their teaching practice. Teachers point out that their educational practice is strongly influenced by their school colleagues and by textbooks and didactic materials used in classroom practice.

Studies carried out by PERG-UBC bear in mind the aforementioned difficulties when analysing teaching sequence implementation. They are usually assessed in three ways. Firstly, we are interested in the effectiveness of the sequence compared to the traditional approach to teaching. Pre-test and post-test analysis is used for this, consisting of a questionnaire with questions related to the learning indicators specified for the sequence. In addition, the students' conceptual understanding in the experimental groups is compared with the Control group. These results are used to judge the sequence's effectiveness in terms of improving the students' understanding, compared to traditional teaching of the subject. We know the methodological difficulties of making these kinds of comparisons, but we agree with Leach and Scott [10] that if they are made in accordance with the conditions imposed by the quantitative methodology of research, they are at least as valid as any others.

Secondly, a group of tasks is usually used to assess the experimental groups' conceptual and methodological understanding. These tasks are carried out by the experimental group students throughout the sequence implementation. The task structure meant that students had to explain their decisions and their results, as well as predicting how situations would develop following the scientific model studied in class (assessment of epistemological aspects). Student responses are recorded on audio or video for later analysis.

Thirdly, our goals demand that students should be interested in the tasks and acquire greater interest in the scientific content of the subject. We wish to assess the sequence's influence on student activities. To do so, a Likert scale questionnaire was designed scoring from 1 to 10. It consisted of 13 questions divided into three sections on: the contents, the method of working in class and the satisfaction with which the work was done. The students in the experimental groups completed the questionnaire after finishing the course. A teacher who had not taught the sequence supervised questionnaire completion, which was done anonymously.

Results from the Teaching and learning EMI project shows that the majority of students (between 50 and 70 %) in the experimental groups demonstrate correct understanding of the studied scientific model. It would actually have been surprising if all the students had answered all the questions correctly. This would have meant that all the students had acquired all the knowledge and skills proposed in the indicators. Our, no less idealistic, intention was for the vast majority of student answers to lie between the "correct" and "incomplete" codes and this was achieved for three quarters of the pupils, in all questions. Similarly, the vast majority of the student groups, who answered tasks where it was necessary to apply the scientific model they had studied, did so correctly. In the case of control group students, the percentage of correct and incomplete replies did not reach 25 % in any of the post-test questions.

The experimental group students also showed a (more) positive attitude to the contents of the experimental teaching sequence. Connections with the concepts studied beforehand and the method of working on the contents in the sequence were particularly emphasised.

What evidence do these results contribute to teaching the topics in question? When drawing conclusions and looking at implications for teaching, it is necessary to bear in mind that the teaching sequences designed in the different projects were implemented in two or three groups of students. In addition, teachers who implemented the sequence are experts in the teaching strategies used and have helped to design some aspects of the sequence. So, we cannot present evidence for more general contexts or teachers untrained in the use of the sequences. However, we have found that similar research on teaching sequences carried out by international groups [14, 16] has also achieved a significant improvement in teaching the specified indicators.

Our projects are not designed to provide conclusive evidence on why students might improve their learning and, in fact, there may be improvements in learning due to other features of the teaching process. However, we think that the existence of a connection between students' learning improvements in the specified indicators and the implementation of these teaching sequences may be a plausible explanation: the sequence and its implementation having been assessed in accordance with the research methodology into science teaching.

The results contributed by our projects show that, for whatsoever reason, students who follow the sequences are capable of obtaining a significantly better understanding of the scientific models proposed in the learning indicators than students who receive traditional teaching. So, teachers who decide to use these sequences in the future seem likely to be able to help their students learn more effectively than with the traditional teaching approach.

Continuing with the design of materials and strategies, as well as their assessment in extensive samples of schools and at different educational levels seems crucial to me, as in research we base ourselves on the fact that if the science teaching (physics) were as it should be, it would not be necessary to spending time getting a better understanding of "how", "when" or "why" students learn.

References

1. Abell SK, Lederman NG (2007) Handbook of research on science education. Lawrence Erlbaum, Mahwah
2. Furió C, Guisasola J, Almudi JM, Ceberio M (2003) Learning the electric field concept as oriented research. Sci Educ 87(5):640–662
3. Guisasola J, Almudi JM, Zuza K (2011) University students' understanding of electromagnetic induction. Int J Sci Educ. doi: 10.1080/09500693.2011.624134. http://dx.doi.org/10.1080/09500693.2011.624134. Accessed November 2011
4. Guisasola J, Almudí JM, Zuza K (2010a) The design and evaluation of an instructional sequence on Ampere's law. Am J Phys 78(11):1207–1217
5. Guisasola, J, Zubimendi, J L and Zuza K (2010b) How much have students learned? Research-based teaching on electrical capacitance. Phys Rev Spec Topics Phys Educ Res 6(2):020102-1–020102-10
6. Guisasola J, Furió C, Ceberio M (2009) Science education based on developing guided research. In: Thomase MV (ed) Science education in focus. Nova Science Publisher, New York, pp 55–85
7. Guisasola J, Almudi JM, Zuza K, Ceberio M (2008) The Gauss and Ampere laws: different laws but similar difficulties for students learning. Eur J Phys 29:1005–1009
8. Halloun I, Hestane D (1985) Common sense conceptions about motion. Am J Phys 53:1056–1065
9. Hurd de Hart P (1991) Issues in linking research to science teaching. Sci Educ 75:723–732
10. Leach J, Scott P (2002) Designing and evaluating science teaching sequences: an approach drawing upon the concept of learning demand a social constructivism perspective on learning. Stud Sci Educ 38:115–142
11. Lijnse P, Klaassen K (2004) Didactical structures as an outcome of research on teaching-learning sequences? Int J Sci Educ 26(5):537–554
12. Maloney DP, O'kuma TL, Hieggelke CJ, Van Heuvelen A (2001) Surveying students' conceptual knowledge of electricity and magnetism. Am J Phys Suppl 69(7):12–23
13. McDermott LC (1997) Bringing the gap between teaching and learning: the role of research. In: The changing role of physics departments in modern universities. American Institute of Physics, New York
14. McDermott LC, Heron PRL, Shaffer PS, Stetzer MR (2006) Improving the preparation of K-12 teachers through physics education research. Am J Phys 74(9):763–767
15. Mellado V (2003) Cambio didáctico del profesorado de ciencias experimentales y filosofía de la ciencia. Enseñanza de las Ciencias 21(3):343–358
16. Michelini M (2003) New approach in physics education for primary school teachers: experimenting innovative approach in Udine University, Inquiries into European Higher Education in Physics, H.Ferdinande et al. eds. EUPEN 7:180–185

17. Pintó R (2005) Introducing curriculum innovations in science: identifying teachers' transformations and the design of related teacher education. Sci Educ 89(1):1–12
18. Simpson RD, Oliver JE (1990) A summary of major influences on attitude toward and achievement in science among adolescent students. Sci Educ 74(1):1–18
19. Zembylas M (2005) Three perspectives on linking the cognitive and the emotional in science learning: conceptual change, socio-constructivism and poststructuralism. Stud Sci Educ 41:91–115

Chapter 40
Quantum Physics in Teacher Education

Gesche Pospiech and Matthias Schöne

Abstract Teacher education forms the decisive link between physics education research and the realization of research results in school. The actions of teachers are influenced by their own experience from school, from study at university and the need to follow the school curriculum. Discrepancies between these factors have to be overcome. In quantum physics, the gap between teaching traditions at university, recent experimental results and the possibilities of teaching at school is especially big. Therefore it seems desirable to study this gap in detail in order to diminish it within given resources.

40.1 Introduction

Quantum physics is the basic theory of physics. During the last decades there has been made significant progress in understanding its fundamental features as opposed to classical physics. Especially experiments concerning uncertainty, entanglement and quantum information allow for deeper insight and have found their way into popular science (e.g. Zeilinger 2003). In this way the fascination of youth towards quantum physics is raised.

However, in general these new developments do not yet find their way into the school curricula and also not into teacher education. On the contrary the teaching at school is strongly influenced by the historical developments, such as the atomic model, the representation of e.g. an electron by the so-called wave-function or the statistical interpretation of the uncertainty relation of position and momentum as central parts. Whereas it is undoubted that quantum physics as a basic theory has to be taught at school it is not equally clear which aspects to choose as the most relevant content. Therefore the implementation of a modern course in quantum physics be at school or at university requires big efforts in several aspects: first it needs the

G. Pospiech (✉) · M. Schöne
TU Dresden, Dresden, Germany

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_40,

definition of an agreed upon “canon” for school and university alike and secondly identifying appropriate learning pathways.

The decision of teachers regarding the focus of their lessons relies on several factors: their own experience as students in school, their own convictions about good physics teaching and its related goals, what they have learned at university and last but not least the school curriculum with the expected teaching. In addition, teaching quantum physics is strongly connected with the personal views on the role of physics and its statements it can make as a science and where its borders are.

Teacher education has to cover mastering the basics of quantum physics—including some mathematical formalism—and on the other hand being able to explain it to students. In university lectures on quantum physics these both combined goals mostly are not reached in such a way that the teachers transfer their knowledge into the classroom. So the adequate fitting of content knowledge and pedagogical content knowledge remains an open problem. The goal of the present study is to go a first step towards its solution in describing the relation between the conceptions of the lecturers at university and the students’ conception of the relevant quantum physics terms and the interrelations.

40.2 Theoretical Framework

Teacher education lies in the triangle between the professional knowledge teachers are expected to acquire at university, their own experiences from the physics lessons they attended themselves at school and the demands of school curricula, lesson plans and school authorities, [1]. In order to devise their teaching teachers have to combine their content knowledge, the pedagogical knowledge and the pedagogical content knowledge. This is seen through the glasses of personal convictions, teaching habitudes and cornered by the school curriculum.

40.2.1 Teaching of Quantum Physics

Quantum physics proves to be a special area in physics teaching because of its abstract formalism, the differences to classical physics, the nearly complete absence of real experiments suitable for school and the strong traditions in teaching partly due to the long unclear history of interpretation. In addition, in quantum physics as a vivid research area with numerous experiments concerning its fundamentals, there is an exceptional discrepancy between concepts revealed by recent research and the more traditional views of teaching at school. Many courses at school (and at university) start with the photoeffect in order to introduce light as particle and in the sequence guiding the students towards quantum objects and the double slit experiment. During this learning path students tend to retain classical notions; they might still think: “finally, the electron must take some path or other” [2].

In modern presentations of quantum physics the terms of superposition, uncertainty and entanglement play a central role, [3, 4]. Michelini and her group have developed a didactical path along the core concepts of quantum theory: the superposition principle and uncertainty, strongly related to phenomenology treating the polarization of photons interacting with polarizers and birefringent crystals, in the end leading to the formalism. For the use of teachers a web site has been constructed [5]. But this approach is not typical for university lectures on quantum physics.

In each learning pathways learning difficulties of students can be observed. A broad range of studies on high school students and university students analyzes conceptions of quantum mechanics (for a short overview see [6]). Students' difficulties in understanding and acquiring an adequate quantum view depend not only on the subject in question—atomic model, uncertainty relation or wave function—but also on preliminary instruction, [7]. Many students do not exhibit a consistent perspective on uncertainty and measurement across multiple contexts. A more refined study [8] focuses on the change of students perspective depending on the views of the instructor as well as on own personal convictions about the "real world", (see also [9]). Baily and Finkelstein find that the students' views on quantum mechanical phenomena can be significantly changed by instruction, e.g. by explicitly teaching the quantum perspective, but may stay mixed between classical and quantum interpretation.

However, there are nearly no studies taking teacher students into the view. One exception is a study concerning the conception of atomic model [10]. The problems of teacher students mostly are the conceptual understanding and the interpretation of quantum physics. Often students calculate without a mental concept [11], the indeterminism and the measuring process are not understood and the classical views, like trajectories or fixed properties remain unchanged [12]. Mannila et al (2001) found differences between students preparing for teacher and future physicists with teacher students tending more to a quantum view. Because of the small number of participants it is not clear whether this is a general feature.

Concerning teachers researchers find that they have limited viewpoints about quantum mechanics. The teachers relate between quantum mechanic concepts and classical ideas in an undifferentiated way. This is a big obstacle in the teaching process in school in the light of a recent study showing that—quite independent from the teaching approach—especially concepts taught in quantum physics are very stably retained, [13]. This hints to the importance of the above mentioned triangle of previous experiences of teacher students for their future learning and directs the focus onto teacher education.

40.2.2 Teacher Professionalisation

Studies show that the quality of teachers' explanations strongly depends on their content knowledge, [14, 15]. It has to be flexible and adjusted to their students needs. An effective teacher is the single most important factor of student learning [16].

Demtröder et al [17] argue that teaching physics is an own profession, not a part of the education for future physicists and has to be a study course of its own. Besides theoretical instruction Pedagogical Content Knowledge (PCK), like methodical didactical knowledge, discussion of historical and philosophical question, interpretation of quantum physics concepts and main modern applications, is necessary to teach quantum physics well in school.

So adequate teacher formation at university has to be developed, especially as it turns out that the most relevant learning period concerning content knowledge seems closed after the "Referendariat" [18]. In an approach to teacher preparation [10] five interactive components of an instructional model are identified: exploring learners pre-instructional knowledge, the content analysis from the educational perspective, identifying the learners needs, fostering the reconstructing of knowledge and meta-cognitive activities. Because in quantum physics the concepts are retained exceptionally well [13] this procedure is of great importance. A study from Bagno et al (2010) showed that teachers need extensive qualitative discussions besides a founding in the formal aspects of quantum theory. But in many courses teachers learn quantum physics with a strong focus on mathematical formalism. As a consequence teacher students often do not know the didactical possibilities to teach quantum physics.

To explain the context of the study teacher education in Germany will be explained in a few remarks. Teachers study two subjects, e.g. physics and mathematics, with equal weight. Part of the study are lectures and seminars on general pedagogy. During their physics study they learn experimental physics, theoretical physics and attend laboratory work. In addition there are special seminars and lectures on pedagogical content knowledge. The physics course covers all the basic areas of physics. Generally only few of these lectures are specially prepared for teacher students but mostly are designed for students choosing physics as their major. However, in Saxony the lectures in theoretical physics are given separately for teacher students. With respect to quantum physics this opens up special possibilities for analyzing the gap between university and school curriculum.

40.3 Design of Study

In this study we concentrate on the relation between the actual teaching and learning of quantum physics at university in the special case of teacher education. On the basis of these results an additional course with emphasis on PCK in quantum physics will be developed. As a first step the status quo is being described.

40.3.1 Research Questions

First we analyze the fit between lecturer views and the view and demands of teacher students concerning lectures in quantum physics. The questions are:

- Which are the central concepts taught by lecturers?
- How do students grasp and understand these concepts?
- Which of these concepts are suited for teaching at school according to the view of teacher students?

40.3.2 Methodology

The goals and main content of the quantum physics course given by the lecturers were collected by a written questionnaire with open-ended questions. They were also asked about the obstacles in teaching, relevant topics and concepts, they wanted to teach to students, a rating of students pre-knowledge and if they make any differences between teacher students and future physicists.

For the teacher students' views a pilot interview study was started with five students to find the content of theoretical lectures, demands for an additional course and suitable topics for teaching quantum physics at school. Also a test of understanding was made with a german translation of tests of Robertson/Kohnle [11] and Baily/Finkelstein [7] as an anchoring test with respect to the learning difficulties described in the studies above. The topics of this test were quantum mechanical measurement, probabilities, possibilities oft interpretation, entanglement and EPR (Einstein, Podolsky, Rosen) Experiment, non determinism and uncertainty principle. In addition the interviewed student should rate the sufficiency of the content and the pedagogical content knowledge and their own understanding after their standard course in quantum physics.

The answers were of both groups to open questions or interviews were analyzed with the qualitative content analysis after Mayring [19] to find the main categories. To compare the interview results concept maps of the categories were built with the program Atlas.ti©. The main categories were defined as best linked with other categories or often mentioned in the interviews.

Based on this analysis we created a main quantitative questionnaire containing questions on the main contents of the lectures, the demanded content for an additional didactical course and suitable topics for a school course in quantum physics, a rating of imparted PCK and CK knowledge after a standard course and a self rating of the quantum physical understanding.

40.4 Results

40.4.1 Results of Analysis of Lectures

The main concepts of the standard courses in quantum physics from lecturers' point of view are shown in Fig. 40.1 as a concept map.

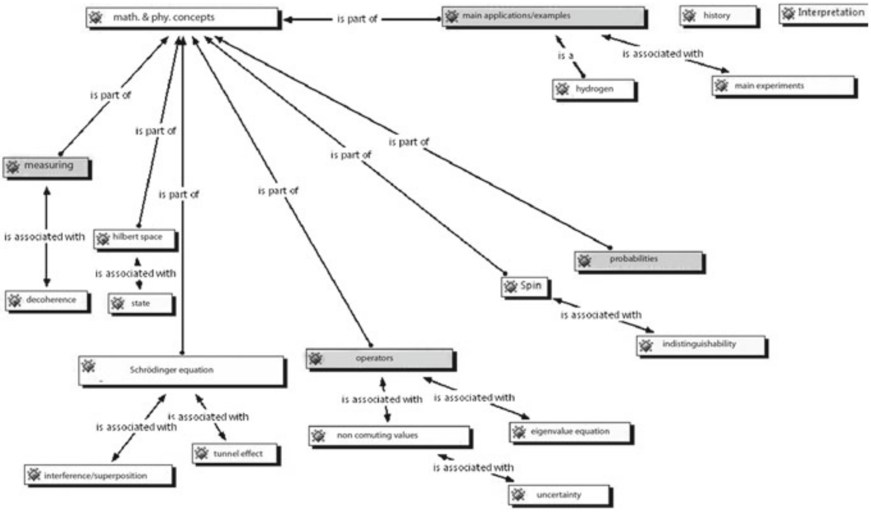


Fig. 40.1 Main concepts according to lecturers. The most important topics are marked grey

Lecturers put a great focus on mathematical techniques. One possible reason may be that lecturers often rate students' mathematical knowledge as insufficient and that it has to be enlarged. So the ability for doing formal calculations and mathematical analysis of quantum physics are two of the main goals of the lecturers. Other important goals are the knowledge of basic concepts and example applications. But no lecturer mentioned methodical or didactical concepts as important. Lecturers often feel there is too few time for these additional questions and/or they are not trained in this respect.

40.4.2 Results of Pilot Study with Teacher Students

As a first result of the pilot study the five different interviews with the students were analyzed and different categories for the content of theoretical physics courses, the demands for an additional didactical course and possible topics for teaching quantum physics in school were created. The main content of the lectures from the students point of view are shown in Fig. 40.2 as a concept map.

The identified categories were used as different choices in the quantitative main questionnaire, so a rank order of these topics could be constructed. In contrast to the lecturers' view students see more details like the Compton or photo effect than main principles. They separate more between physical model and mathematical formalism. Also students connect general quantities to concrete examples, e.g. Spin is related to hydrogen atom, but indistinguishability not at all. Also the main demands of

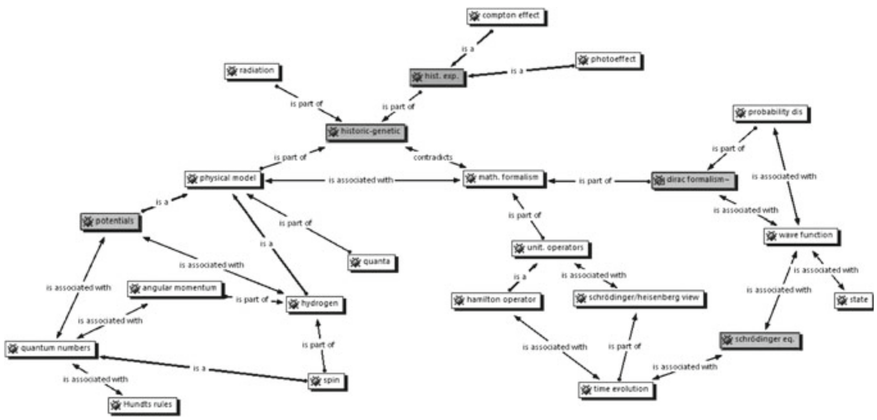


Fig. 40.2 Main contents of lectures—seen by students. The most important topics are marked grey

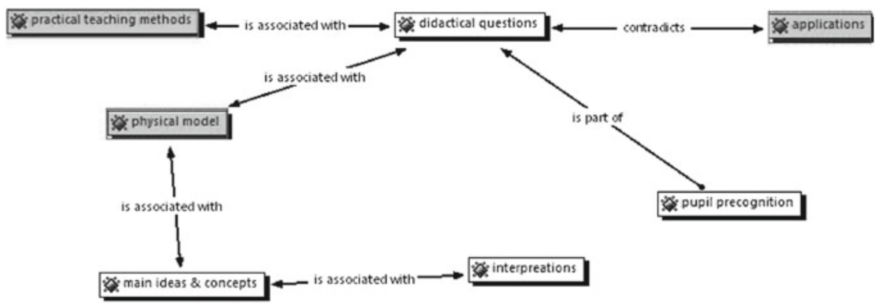


Fig. 40.3 Students demands for an additional didactical course

the students for an additional course that improves the didactical education were categorized and are shown in Fig. 40.3, the main demands also marked grey.

As a second part the answers in the test for understanding were analyzed. Three of five students formally knew the problems of the measuring process in quantum physics, but none respectively only one could apply this knowledge to two questions about energy measurement (see Robertson/Kohnle [11], question 1.1 and 1.2). In contradiction 90 % of the comparable experimental group of Robertson/Kohnle could give the right answer to question no. 1.1 and 35% to question no. 1.2.

Concerning the explanation of quantum properties the uncertainty could be described by four of five students, but none could describe the difference between the uncertainty in a classical Galton Board experiment (for further details see Baily/Finkelstein [7], essay question no. 2) and quantum uncertainty. This is almost in agreement to the result of Baily/Finkelstein with 15 % right answers.

Only two of five students in our survey could describe the indeterminism and the entanglement correctly. None of the students knew the significance of the EPR Experiment.

40.4.3 Results of Questionnaire

In the main study 20 master students participated, mainly from the fifth year at universities in Dresden and Leipzig.

The content view of the master students in the quantitative survey is similar to the qualitative one. Main subjects like the Dirac formalism, the Schrödinger equation and calculus of potential wells are mentioned often, but historical experiments seem not so important for this group. In addition they put more attention to angular momentum/spin and mathematical formalism, e.g. the description of the time evaluation, so they are closer to the lecturers view.

The main demands of the students are also a stronger explanation of the physical model in interpretation questions but mainly PCK questions in practical teaching methods (18 of 20 students), concepts for a teaching unit (17 of 20 students), application examples (14 of 20 students) and general educational viewpoints (16 of 20 students). This is in accordance to the rating of students own learned PCK in quantum physics lectures. 70 % of the students think the given knowledge is not sufficient for teaching quantum physics well in school. In contradiction students feel well prepared in their content knowledge. 16 of the 20 interviewed students think the main concepts of quantum physics became clear. But not all subdomains were equally rated, only 11 of 20 of the students feel well prepared in the topic of entanglement, 12 of 20 in uncertainty and only 8 of 20 in the understanding of the measurement process. In accordance to the lecturers' view only 12 of 20 students feel sufficiently prepared in the mathematical formalism.

The analysis of suitable topics for school courses in quantum physics show that the uncertainty principle, application of laser and properties of quantum objects are recommended by all students. Also the photo effect (19 of 20 students) and interference experiments (17 of 20) were seen as suitable topics.

Many topics from the university lectures like angular momentum, Schrödinger equation, Bose Einstein Condensation, probability distribution and equations of radiation, which are difficult, were rejected by the students for school education.

Visionary application topics like quantum computer or cryptography are rated ambiguous. This shows that students often reconstruct their own school experience, especially in historical experiments, but are open for new subjects like applications if they are well constructed. The analysis of the central points of the saxonian school curriculum (2011) shows similar topics with photon as a quantum object, the photo effect, interference of photon, electron and neutron and Heisenberg's uncertainty principle. Only the measurement process was not mentioned by the students, but is a main topic in the curriculum. It is remarkable that lecturers are focused on mathematical formalism, which cannot be used in school, because only basics of vector algebra are known by the pupils.

40.5 Future Work

In order to compare content structures of university and school lectures it is planned to analyze two different concept maps. One is built up from lecture scripts and exercises of university lectures, the other from school books. Hendrik Härtig has shown in his dissertation (2010), that school books are with high significance a good indicator for the average school lesson. To validate both concept maps experts will analyze and verify them.

To improve the pedagogical content education of teacher students we are planning an additional course. Main goals of this course is the discussion of basic ideas of quantum physics, fundamental experiments and possible problems in the teaching learning process. But also teaching concepts in quantum physics and practical developments by students will be included.

The course will be evaluated with a pre- and posttest with qualitative instruments and a questionnaire.

40.6 Conclusion

This study shows that lecturers and teacher students have different views and demands for a reasonable course in quantum physics. Lecturers see basic mathematical techniques and quantum physics formalism as main goals. Teacher students want to understand the basic principles, interpretational problems and like to discuss PCK relevant questions, like practical teaching methods, teaching concepts and applications.

It was shown that teache students feel well prepared in the content knowledge but poorly in pedagogical content knowledge with the consequence that they fell not prepared to teach quantum physics in school. We will analyze the effects of an additional course especially for teacher students in this respect.

References

1. Etkina E (2010) Pedagogical content knowledge and preparation of high school physics teachers. Phys Rev Spec Top Phys Edu Res 6:020110
2. Fanaro M, Otero MR, Arlego M (2009) Teaching the foundations of quantum mechanics in secondary school: a proposed conceptual structure. Investigações em Ensino de Ciências 14(1):37–64
3. Michelini M, Ragazzon R, Santi L, Stefanel A (2000) Proposal for quantum physics in secondary school. Phys Educ 35(6):406–410
4. Pospiech G (1999) Teaching the EPR-Paradox at high school? Phys Educ 34:311–316
5. Santi L, Michelini M, Stefanel A, Meneghin G (2011) A resource environment for preservice teacher education to introduce quantum physics in secondary school. In: MPTL conferrence 2010

6. Akarsu B (2011) Instructional designs in quantum physics: a critical review of research. Asian J Appl Sci 4(2):112–118
7. Baily C, Finkelstein N (2009) Development of quantum perspectives in modern physics. Phys Rev Spec Top Phys Edu Res 5:010106
8. Baily C, Finkelstein N (2010) Refined characterization of student perspectives on quantum physics. Phys Rev Spec Top Phys Edu Res 6:020113
9. Levrini O, Fantini P, Pecori B (2008) The problem is not understanding the theory, but accepting it: a study on students' difficulties in coping with quantum physics, in GIREP-EPEC conference. Frontiers of Physics Education Selected contributions, RIJEKA Zlatni rez
10. Mannila K, Koponen IT, Niskanen JA (2002) Building a picture of students' conceptions of wave- and particle-like properties of quantum entities. Eur J Phys 23:45–53
11. Robertson E, Kohnle A (2010) Testing the development of student conceptual understanding of quantum mechanics, GIREPEPEC PHEC 2009 conference (2010), pp. 261–273
12. Müller R, Wiesner H (1998) Vorstellungen von lehramtsstudenten zur interpretation der quantenmechanik. In: Behrendt H (Hrsg): Zur didaktik der physik und chemie, Alsbach S, p. 382
13. Deslaurier L, Wieman C (2011) Learning and retention of quantum concepts with different teaching methods. Phys Rev Spec Top Phys Edu Res 7:010101
14. Baumert J, Kunter M, Blum W, Brunner M, Voss T, Jordan A et al (2010) Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. Am Educ Res J 47(1):133–180
15. Riese J (2010) Empirische erkenntnisse zur wirksamkeit der universitären lehrerbildung - indizien für notwendige veränderungen der fachlichen ausbildung von physiklehrkräften. Physik und Didaktik in Schule und Hochschule 9:25–33
16. Darling-Hammond L (2000) Teacher quality and student achievement: a review of state policy evidence. Educ Policy Anal Arch 8(1):1–44
17. Demtröder et al. (2006) Thesen für ein modernes Lehramtsstudium im Fach Physik, DPG
18. Borowski A, Kirschner S, Liedtke S, Fischer HE (2011) Vergleich des fachwissens von studierenden. Referendaren und Lehrenden in der Physik, Physik und Didaktik in Schule und Hochschule 10(1):1–9
19. Mayring P (2010) Qualitative inhaltsanalyse: grundlagen und techniken, Beltz
20. Eylon B, Cohen E, Bagno E (2010) The interplay of physics and mathematics in a graduate quantum mechanics course for physics teachers, talk given in GIREP-conference 2010 in the Symposium "Addressing the role of mathematics in physics education
21. Kalkanis G, Hadzidaki P, Stavrou D (2003) An instructional model for a radical conceptual change towards quantum mechanics concepts. Sci Educ 87:257–280

Chapter 41
Using a Sociocultural Approach in Teaching Astronomy Concepts with Children of Primary School

Rocco Servidio, Marcella Giulia Lorenzi and Mauro Francaviglia

Abstract This chapter describes the use of a social collaborative learning approach to teach astronomical concepts. We used a repeated measure design to test the effectiveness of a teaching intervention aimed at improving basic concepts of Astronomy. Results show that children after the didactical experience were able to give a detailed explanation of the acquired knowledge. These results suggest that the approach adopted is fruitful and helps children to learn and understand scientific concepts, revisiting their pre-existing knowledge, often influenced by naïve explanations.

41.1 Introduction

Several studies highlighted that primary school age children often build their scientific concepts of the physical world on the basis of their own everyday experiences [2, 3, 7, 9, 11, 12] . Some of children's Astronomy ideas are [5, 10]: the "Earth is flat" (children imagine the Earth shape as a disk); the "Earth is hollow" (the earth is similar to a rectangle or to a sphere). Yet, children think that some natural phenomena can represent negative events. They consider, for example, the passage of comets as negative events or premonitory signs that warn people about possible dangers or disasters. Another common children misconception concerns the day/night cycle, which they attribute to the Sun anthropomorphic quality [6]. The authors claim that children have the idea that during the day the Sun "lives" like a normal being and then, in the night, it "goes to sleep" (as ancient people believed).

R. Servidio (✉)
Department of Linguistics, University of Calabria, Cosenza, Italy

M. G. Lorenzi
University of Calabria, Cosenza, Italy

M. Francaviglia
Department of Mathematics, University of Torino, Torino, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_41,

A deep understanding of scientific concepts requires that children significantly learn and share both theoretical and practical skills. Hence, to encourage this conceptual attainment, teachers should design educational settings able to stimulate the students' mental activity to analyse problems and then to apply "creative thinking" strategies. The main idea of this approach is to stimulate children to learn from direct experience without restricting the didactical potentialities of the scientific exploration. Furthermore, learning is more productive when children are exposed to the socials interactions, where language, group discussions, collaborative and cooperative work, and tools' use play an important cognitive function (Vygotskij 1986). There enhances reflection and helps learners to build well-grounded and shared scientific concepts.

Although the research on children's ideas remains an important field of inquiry into the astronomical education as a whole, many questions associated with the cognitive processes involved in knowledge acquisition and concept learning are still unanswered. We can summarize these aspects as follows: (1) is it possible to restructure the pre-existing children's misconceptions regarding Astronomy concepts? (2) What is necessary in order to improve the traditional educational settings to teach Astronomy concepts in a productive way? A common aspect of the aforementioned factors concerns teacher modalities to deliver scientific concepts.

In this study, made in a recent academic year, we tested the effectiveness of teaching intervention, which was based on "sociocultural learning" (Vygotskij 1986). We used a repeated measure design to compare teacher and investigator treatments. The last one emphasizes the role of "social learning theory" as a good didactical strategy to teach Astronomy topics. The instructional material was organized with the intent to motivate children learning interest, facilitating their mental elaboration of the information presented. Specifically, we argued that Astronomy concepts are not difficult to learn. Nevertheless, it is important to identify appropriate didactical methods that help children to better understand scientific ideas. Therefore, we also argued that children astronomical knowledge before treatment session was affected by naïve ideas. We expected that when applying the sociocultural learning approach to the treatment session, children would improve their astronomical understanding.

This chapter is organized as follows: in the next Sect. 41.2 we describe the aim of the current study; then, in the "Procedure" Section, we present the design and the procedure adopted in developing the current research. "Results" of the pre- and post-tests assessments follow. Finally in the Section "Conclusion" we present the outcomes of the study.

41.2 Aim of the Study

The purpose of the present investigation was to examine the effectiveness of a sociocultural didactical approach with the support of cognitive tools, aimed at teaching Astronomy concepts. Children and teachers were actively involved during the educational and didactical sessions. The main idea was to create a productive learning context, where teacher and children work cooperatively. It is widely agreed that "learning" is not merely an individual process, since a more complete cognitive result can be achieved when social and cultural aspects are taken into account [4].

We designed and used traditional didactical materials and multimedia contents oriented to attract students' attention and interest. By applying this didactical approach students were able to share ideas, make inferences, and they identified relationships among concepts, rather than exchanging information without a concrete direct experience.

41.3 Method

41.3.1 Participants

The participants in this study were twenty-two Italian children (10 males and 12 females), aged between 9 and 10 years (M = 9.8; SD = 0.35), took part in this study. They attended two fifth-grade classes of a Primary School in San Lucido (Cosenza, Southern Italy). For the study purposes, all the didactical activities took place in each single classroom, respectively. We did not use a preliminary test to measure children's cognitive abilities. Students participated in this study after having obtained a written consent of the School Director. The children sampling took into account the previous studies that showed different examples on how children learn Astronomy concepts ([2] Kikas 2003; Vosniadou and Brewer 1992).

41.3.2 Study Design

We used a repeated measure design to test the effectiveness of the sociocultural didactical approach (Vygotskij 1986). In this study, pre-test session represents the control condition and the post-test the experimental one. The proposed study design includes the Mathematics and Physics schoolteacher's suggestion, which wished that all children received the same astronomical concepts delivered through the proposed didactical approach. This educational approach was also coherent with the objectives and the guidelines of the Italian Primary School curriculum. In order to control the student's learning progress, we used different evaluation methods, such as graphical representations, questionnaire and social learning activity. Participants were individually tested except for the cooperative educational activity.

The twenty-two-grade students, attending the primary school, received the same basic teaching on the shape and rotational movement of the Earth and Sun, the day/night cycle, the seasonal changes, gravity, meridians, and parallels. We introduced these concepts in the fifth-grade primary class since during the research children have already completed the educational programme (according to the Italian Education Minister curriculum) that included the learning of astronomical concepts.

Before starting with the direct educational activities children were pre-tested by using a multiple-choice questionnaire, in order to evaluate their initial knowledge about specific astronomical phenomena. Schoolteachers examined the students' learning by using this type of test. Thus, we avoided to introduce new didactical constraints in the classroom. The investigator administered again the same questionnaire immediately after the treatment session, aimed at examining children learning progress, and again one month later, in order to evaluate how children maintained the acquired knowledge over time.

We assumed that after the treatment session children would better understand the astronomical concepts. In this study, the treatment included educational activities based on "social learning theory". All the didactical activities were presented to children with appropriate information along with conceptual tools and practical collaborative activities. This approach should help children to improve their scientific understanding, avoiding at the same time the learning of astronomical concepts in a naïve way. Specifically, it is rather effective when children learn scientific concepts in a non-monitored context (e.g., family environment, and so on).

41.3.3 Materials and Data Collection

A geographic map of Italy, two globes, a lamp with swing arm, equipment for drawing (a black marker), a blackboard, balloon and a flashlight, a thermometer, two manuals of Astronomy and a laptop were adopted during the whole period of the treatment activities. We also selected a series of multimedia lectures notes that the investigator used in the classroom during the lessons, to explain the Astronomy concepts to children. In addition, to support children's learning, we arranged simple hands-on Astronomy exercises in the classroom.

To assess children's Astronomy concepts and to easily introduce them we designed a questionnaire, taking into account the previous studies conducted by other researchers [12]. The questionnaire consisted of twelve items, and each one included two answer modalities: dichotomous (Yes or No) and open-ended. We used the open-ended modalities to collect a richer source of information about children's astronomical knowledge. The open-ended questions aimed at identifying the meaningful utterances used to explain some astronomical events and indicating the learning of new concepts. Therefore, we were able to understand as children explained scientific concepts they learned. The questionnaire items examined the Earth's shape, the day/night cycle, the Earth's rotation, the concept of gravity, meridians, parallels and others concepts regarding the Solar System.

41.3.4 Procedures

Two classrooms of students participated in this research carried out during in the 2009 academic year. According to the study design, the investigator delivered the same educational contents in each classroom, separately. Children were told that the proposed educational activities were not a test and that they could say if they did not wish to answer a question or did not know an answer. The same instructions were given for the graphical representations. In addition, schoolteacher of Mathematics and investigator gave no support to children to perform all the designed educational activities. One month after the end of the educational activities, we evaluated children's learning again.

Initially, students were individually tested administering them a pre-test questionnaire. Each participant was invited to read carefully the questions statement and then to pick out the correct answers, providing an explanation of it (open-ended modality).

During the second session (called treatment) we delivered target Astronomy concepts. Each educational activity consisted of two parts. In the first part of the lesson the investigator introduced theoretical concepts in Astronomy to children. Afterwards, the investigator designed some simple experiments in the classroom, in order to actively involve the children. The aim of this approach, based on "social learning theory", was to explain to the children the basic astronomical phenomena, making use of their own conceptual knowledge. The duration of each lesson lasted approximately 45 min. The research was carried for five and half weeks and the educational activities were delivered twice per week (Tuesday and Saturday). Finally, the third session included second and third observation, respectively.

The schoolteacher of Mathematics in one year earlier (April–May 2008) taught to the children basic astronomical concepts by adopting a normal textbook. These concepts included the shape of the Earth and the day/night cycle. The textbook included theoretical description of astronomical phenomena and practical exercises to improve the students' learning. For instance, the textbook displayed the position of the Planets by using images, chart and other materials. Finally, it provided a list of questions to evaluate the children learning. The duration of each lesson ranged from 35 to 45 min. It depended by the lesson organization and by the topics of the didactical activities scheduled, as well as by the students' attention.

According to the research purposes, we performed a classroom pre-test aimed at assessing the initial children's knowledge about Astronomy. We administered the questionnaire in the classroom and each student had to reply it individually. Children had approximately 50 min to fill it out. They knew that this was not a test, and that they could say if did not wish to answer a question or did not know one. In addition, subjects completed the task without receiving any support.

After this preliminary assessment, investigator started with treatment activities. The research design included both theoretical contents and simple hands-on Astronomy exercises. We designed the educational material to be flexible, taking into account children's ideas that emerged during the class discussions. It included multimedia resources (e.g., videos, animations, images, and sounds). During the educational sessions children were stimulated to apply a social-cooperative strategy aimed at enhancing the learning of astronomical concepts. Other activities included discussions in a whole-class session. For example, a child asked the following question: "*If I move, I go to the other side of the terrestrial globe, for example, in China, here will I walk to head down and therefore could I fall from the Earth?*" Hence, the investigator introduced the concept of gravity, explaining to children that it is an attractive force, which causes all bodies to move towards the centre of the Earth. In the next educational meetings the investigator introduced other astronomical concepts like: day/night phenomenon, Earth rotation, seasonal changes, meridians and parallels. Most of these concepts were delivered with the support of multimedia materials.

At the end of the treatment session, the investigator administered the post-test questionnaire. The setting was the same of the pre-test. One month later, we repeated again the questionnaire administration, without modifying the school setting.

41.3.5 Data Analysis of the Results

First, we assessed the internal consistence of the designed questionnaire performing a Cronbach alpha reliability test. We obtained a good reliability result. The Cronbach alpha value was 0.72. Nunnaly [8] has indicated 0.70 to be an acceptable reliability coefficient, but lower thresholds are sometimes used in the literature.

Preliminary pre-test analysis showed that all participants had similar prior astronomical concepts ($F_{\text{dichotomous}}(1, 21) = 1.53,\ p > 0.05$; $F_{\text{open-ended}}(1, 21) = 0.13,\ p > 0.05$). Children acquired these backgrounds on Astronomy during their previous educational activities carried out by schoolteacher. Vice versa, the differences between pre-test and post-test condition were significant ($F_{\text{dichotomous}}(2, 65) = 16.50,\ p < 0.01$; $F_{\text{open-ended}}(2, 65) = 35.73,\ p < 0.01$). Secondly, after pre-test 68 % of the 22 assessed children answers correct to the dichotomous items, whereas 32 % gave a wrong answer. Analysing the results of the post-test, we obtained that 98 % of the answers of 22 assessed children were correct and only 2 % of them answered wrongly. 1 month later, we have observed a short difference in comparison with post-test. In particular, 97 % of the assessed children answered correctly, while the remainder 3 % of them answered wrongly.

Results show that the mean number of correct answers to the questions after the treatment ($M = 11.79$, $SE = 0.10$) and 1 month later ($M = 11.63$, $SE = 0.10$) was greater compared with pre-test results ($M = 8.18$, $SE = 0.32$). Although all children had the same initial knowledge about Astronomy concepts, an active, social and collaborative participation improved their scientific understanding.

To examine the significance of the students' learning changes occurred over time and in connection with the treatment, we have performed a series of two-tailed paired *t*-test. The results indicated that there was an overall significant improvement from pre- to post-test scores for the dichotomous questions ($t21 = -10.62,\ p < 0.01$) regardless of treatment condition. However, the difference between the two outcome measures was not significant (t21 $= 0.70,\ p > 0.05$). We obtained the same result for the open-ended questions. A significant improvement from pre- to post-test scores indicated that the children language showed more details to explain Astronomy concepts (t21 $= -15.82,\ p < 0.01$). In contrast, we did not find significant differences between post-test and one month late measures (t21 $= 0.70,\ p > 0.05$).

41.4 Conclusion

The aim of this study was to experiment the social learning approaches to teach Astronomy concepts. After training treatment, children understood astronomical notions and in particular the behaviour of the Earth system, and they clarify their knowledge about day/night cycle, and other concepts introduced during the treatment session. For instance, children understood that day/night phenomenon is the result of the Earth rotation on its own axis. The pre-test results showed that few children answered correctly to the astronomical questions, but in many cases they gave an incorrect description of the phenomenon. On the contrary, post-test results showed that children increased their conceptual acquisition. As expected, the results of the current study provide meaningful evidence that children retrieved easily the Astronomy concepts learned during the classroom activities. In particular, this social learning astronomical experience shows that certain didactical strategies have a greater potential to improve scientific knowledge in children. After the treatment, children showed not only the ability to learn new astronomical knowledge in their memory, but also to retrieve it easily by using scientific language and avoiding naïve explanation.

These results support our expectation, indicating how social learning strategies and use of tools can improve children understanding of scientific events and, in particular, of astronomical concepts. It means that children's learning is more productive when teaching involves both social educational and interaction strategies among pairs with the support of tools that increase the scientific learning. Clearly, the level of acquisition of the new knowledge reflected the treatment carried out during the investigator activities. Children verbal descriptions and/or explanations represent important aspects to evaluate their scientific conceptual learning. For this latter aspect, examining children's open-ended statements, we found significant similarities in both theoretical and practical educational content delivered in the classroom. In addition, reading the children's descriptions it is possible to find also a narrative thinking that show the cognitive association between scientific concepts and quality of the learning [1].

The current results indicate also that the integration of social learning methodologies and use of multimedia contents create a didactical environment, in which children are stimulated to extend their own scientific mental skills. This learning setting encouraged children in an active exploration of ideas affording their opportunity to learn Astronomy concepts in a new and dynamic way. However, the role of the schoolteacher is crucial. Schoolteachers should be able to experiment new didactical approaches to reinforce children interest, designing classroom laboratory to teach astronomical phenomena. The social learning approach gives to schoolteachers the possibility to fascinate children, and capture their attention and stimulate the imagination.

Overall, this study demonstrated the successful use of the theory of social learning and tools for engendering playful astronomical learning. The didactical settings were designed to include theoretical and practical activities that mediated different astronomical concepts and phenomena. This enabled students to actively think on scientific concepts, thus providing their new way to explore scientific concepts and then to improve the learning outcomes. This combined use of social learning approach and tools means that the students themselves become a central part of the educational activity rather than just watching something in an inactive way.

References

1. Boulter CJ (2000) Language, models and modeling in the primary science classroom. In: Gilbert JK, Boulter CJ (eds) Developing models in science education, Kluwer Academic. Dordrecht, pp 289–305
2. Brewer WF, Chinn CA, Samarapungavan A (2000) Explanation in scientists and children, In: Keil FC, Wilson RA (eds) Explanation and cognition. The MIT Press, Cambridge, pp 279–298
3. Ehrlén K (2008) Children's understanding of globes as a model of the earth: a problem of contextualizing. Int J Sci Educ 30(2):221–238
4. Iiskala T, Vauras M, Lehtinen E, Salonen P (2010) Socially shared metacognition of dyads of pupils in collaborative mathematical problem-solving processes. Learn Instr 21(3):379–393
5. Kallery M (2010) Astronomical concepts and events awareness for young children. Int J Sci Educ doi:10.1080/09500690903469082
6. Kallery M, Psillos D (2004) Anthropomorphism and animism in early years science: why teachers use them, how they conceptualize them and what are their views on their use. Res Sci Educ 34(3):291–311
7. Miller BW, Brewer WF (2010) Research report misconceptions of astronomical distances. Int J Sci Educ 32(12):1549–1560
8. Nunnaly J (1978) Psychometric theory. McGraw-Hill, New York
9. Plummer JD (2009) Early elementary students' development of astronomy concepts in the planetarium. J Res Sci Teach 46(2):192–209
10. Skopeliti I, Vosniadou S (2007) Reasoning with external representations in elementary astronomy. In: Vosniadou S, Kayser D, Protopapas A (eds) Proceedings of EuroCogSci07, the European cognitive science conference, Delphi, Greece, pp 244–249
11. Stathopoulou C, Vosniadou S (2007) Exploring the relationship between physics related epistemological beliefs and physics understanding. J Educ Psychol 32(3):255–281
12. Vosniadou S, Skopeliti I, Ikospentaki K (2004) Modes of knowing and ways of reasoning in elementary astronomy. Cogn Dev 19(2):203–222
13. Vygotskij LS (1962) Thought and language. MIT Press, Chicago

Chapter 42
Dynamic Modelling with "MLE-Energy Dynamic" for Primary School

Enrico Giliberti and Federico Corni

Abstract During the recent years simulation and modelling are growing instances in science education. In primary school, however, the main use of software is the simulation, due to the lack of modelling software tools specially designed to fit/accomplish the needs of primary education. In particular primary school teachers need to use simulation in a framework that is both consistent and simple enough to be understandable by children [2]. One of the possible area to approach modelling is about the construction of the concept of energy, in particular for what concerns the relations among substance, potential, power [3]. Following the previous initial research results with this approach [2], and with the static version of the software MLE Energy [1], we suggest the design and the experimentation of a dynamic modelling software—*MLE dynamic*-capable to represent dynamically the relations occurring when two substance-like quantities exchange energy, modifying their potential. By means of this software the user can graphically choose the dependent and independent variables and leave the other parameters fixed. The software has been initially evaluated, during a course of science education with a group of primary school teachers-to-be, to test the ability of the software to improve teachers' way of thinking in terms of substance-like quantities and their effects (graphical representation of the extensive, intensive variables and their mutual relations); moreover, the software has been tested with a group of primary school teachers, asking their opinion about the software didactical relevance in the class work.

42.1 Introduction

During the recent years simulation and modelling are growing instances in science education. Modelling activity is useful to construct a language among students and teacher, to develop the ability to deal with variables, to design experiments, to

E. Giliberti (✉) · F. Corni
Università di Modena e Reggio Emilia, Modena, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_42,

interpret the results. These are abilities usually associated with secondary school sicence education curriculum, so the software are designed to accomplish the needs of students with a deeper mathematical grounding. In primary school, the main use of software is the simulation, which is an easier way to work with variables and their relations, without the need to deal with the definition of the model in which the varialbles are related. Often the simulation software for primary school are performing a single task and are useful as a substitution or an integration of the experimental activity. This approach can lead to a non-coherent design of activities, as the primary school teacher often doesn't have a theoretical framework in which the activities can be placed. For this reason we suggest that the teacher needs to develop the ability to build and test simple models of phenomena, which can be successively experimented in the class activity.

Today we observe the lack of modelling software tools specially designed to fit/accomplish the needs of primary education. In particular primary school teachers need to use modelling and simulation in a framework that is both consistent and simple enough to be understandable by children [2].

One of the possible approach to modelling is the construction of the concept of energy, in particular for what concerns the relations among substance, potential, power [3].

Following the previous initial research results with this approach [2], and with the static version of the software MLE Energy [1], we suggest the design and the experimentation of a dynamic modelling software—*MLE dynamic*-capable to represent dynamically the relations occurring when two substance-like quantities exchange energy, modifying their potential.

Modeling energy transfer processes is a possible activity to introduce primary school teachers and children. Energy transfer processes are a suitable starting point for simple modeling activities in primary school, as energy is an important subject in the primary school curriculum which is often approached in a simplified and descriptive way: energy in different forms, "transformation" of energy, renewable and disposable energy forms, etc.

Instead, our approach considers the construction of the concept of energy: the relations among substance, potential, power [3] are the foundation of energy concept. Starting with simple phenomena in which energy is transferred from an "energy carrier" to another. Every energy carrier has two aspects: a substance-like quantity (extensive quantity) and a conjugated quality, or potential (intensive quantity). Combining the two aspects of quantity and quality we have an effect which can be assimilated with power or "force".

As a starting point, static modeling of energy transfer processes is useful both for pupils and for teachers [1], as they can discuss and interpret the observed phenomena using the symbolic representation provided by the software as a language, helping to develop a more accurate use of common words.

Although useful, static modeling has a series of limitations in use: shows its limits in particular when the model needs to be verified (primary school teacher's competence is often not sufficient to predict the behavior of the model and this can lead to misunderstanding among children).

Another critical aspect connected with static modeling is the lack of quantitative information: to be effective and to guarantee a comprehension of the cause and effect relation between input and output, energy transfer needs to be quantified, at least showing the relation of direct and inverse proportionality between variables.

42.2 Dynamic Modelling Software for Primary School

At the present there are no modeling software designed for primary school, and the availability of dynamical modeling software although not specifically designed but at least suitable for primary school is very little.

We discarded software such as Stella (ISEE Systems) or Coach (CMA), in which the model has to be expressed by means of formulas and the relations are represented in graphical form, because in this case the poor mathematical grounding of the teacher represents a problem. Other software make use of graphical representation of simple mathematical relations, such as VnR (Variables and Relations, by Ian Lawrence). This software can be suitable for teachers but it can be difficult in the everyday use, as it reduces the relations between physical quantities to mathematical relations, leading teachers to think in terms of mathematical relations (with some critical aspects) rather than relation among physical quantities (Corni et al 2007). At last, software like Junior Simulation Insight (Logotron) or Modeling Space (…) present a graphical representation acts dynamically, driven by variables, which constitute the mathematical model.

Although Junior Simulation Insight is very good for simulation (it has a very comprehensive collection of simulations, from different contexts and well related with the primary school curriculum), the interface and the functions of the software are quite complex. Moreover, modeling is difficult for primary school teachers, due to their poor mathematical abilities.

Our proposal is to develop a simple modeling software, to let the teacher model simple energy transfer processes. Teachers will be able to predict the behavior of simple energy transfer processes (what happens when a variable changes); to verify the relation between cause and effect (causation); to verify the behavior of everyday life processes. In this way, teachers will be able to design experiments or exploration activities and to assist pupils during the design of the experimental activities.

42.3 Features of MLE Dynamic Software

According with the Force Dynamic Gestalt approach [3] and previous research results [1], we propose a software to model a single “device” and two energy carriers, one incoming and one outgoing. Each of the two carrier has two aspects: a substance-like quantity (extensive variable) and a conjugated quality or potential (intensive variable). According with Image schema theory [5] and Interface design theory [4]

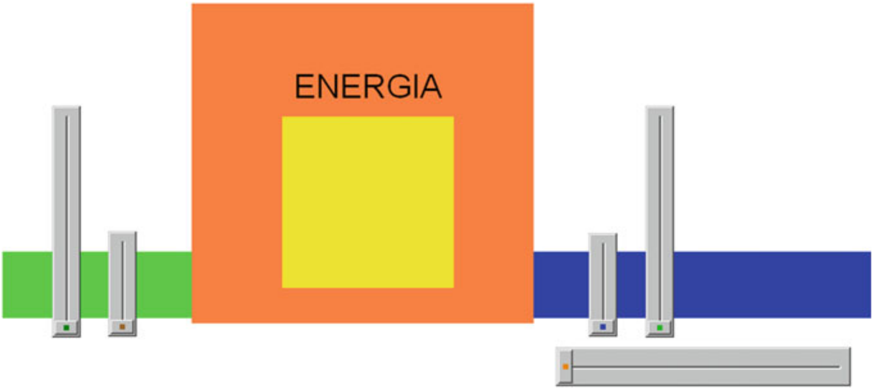


Fig. 42.1 MLE Dynamic

in the software each variable is represented with a slider: horizontal for quantity, vertical for intensity.

The interface of the software is abstract, requesting the user to assign to each of the four variables its name and its meaning. Our proposal is a sort of "abstract simulation". It has the characteristics of simulation, because it allows to predict the behavior and to verify hypotheses, but it does not present a graphical representation of the simulated phenomenon, so we can say that it is a simulation, although with abstract representation. The model on which the simulation relies on is a simple relation among four variables, two for the incoming carrier and two for the outgoing one, so it is Q_{in} x Δ V_{in} = Q_{out} x Δ $V2_{out}$, where Q is the quantitative variable and ΔV is the difference of potential.

Teachers are guided by a worksheet in the analysis of the process of energy transfer, following some steps: identify the relevant variables: substance-like/extensive quantity and conjugated intensity/intensive quantity both incoming and outgoing from a "device".

Teachers decide which is the dependent variable, choosing the relevant model among four (the dependent variable can be one of the four variables), then they decide two more variables (parameters) which have to be set and remain fixed, and finallyhey increase or decrease the independent variable and see the corresponding variation of the dependent one.

The area of the yellow square labeled "energy" represents the total amount of energy transferred during the process.

It is useful to focus on the role of energy in the transferring process: to change the amount of transferred energy it is necessary to change one of the two incoming variables. Changing one of the outgoing variables affects the other outgoing one, with no change in the total amount of energy transferred. Otherwhile we can change one of of the incoming variables, changing the total amount of energy transferred.

42.4 Conclusions

Abstract representation is useful to help the teachers to focus on the relevant variables, i.e. "building" the model. It helps to develop a general model, which is more suitable to interpret phenomena in different contexts. The relations among the variables are clear and unambiguous if the "choice algorithm" is followed carefully (by means of a guide worksheet). Abstract representation does not help to refer to the concrete and relevAnt aspects of the observed phenomena, so it would be useful a label or a iconic explanation of every variable. In the beginning it is important to understand the basic behavior of the device and the relevant variables (it's impossible to build a model without thinking about cause and effect process). After the initial comprehension of the observed phenomenon, the software is helpful to formalize the behavior under different circumstances and in different situations. The software shows its limits when representing currents expressed with discrete quantities (such as number of sharpened pencils with a mechanical pencil sharpener, or number of sacks of flour produced by a windmill) Further improvements in the software will let the user, when uses the simulation, to remember the choices made, including some iconic elements: labels for the four variables; an icon for "fixed" variables (parameters); a text field, in which the question(s) posed by the user, and to which the model they answer, can be written using natural language; iconic representation of the energy carriers (moving flames, charges, drops of water …) with different colors to indicate their different potentials. At the moment the MLE-Dynamic software graphical interface is context-independent, with no iconic reference to the real appearance or symbolic representation of the substance: water, electricity, motion, etc., have the same representation. This is useful to improve the teachers' ability to develop models which are independent from the context, but it is not suitable for children. In the future we are introducing the possibility to characterize every carrier with its iconic representation.

References

1. Corni F, Giliberti E, Mariani C (2009) The MLE-energy software for energy chains modelling. In: MPTL 14 proceedings, Udine 23–25 September 2009
2. Corni F, Giliberti E, Mariani C (2010) Quantity/potential-related elementary concepts in primary school teacher education. In: GIREP-ICPE-MPTL 2010 proceedings, Reims 22–27 August 2010
3. Fuchs HU (2007) From image schemas to dynamical models in fluids, electricity, heat, and motion. An essay on physics education research, https://home.zhaw.ch/~fuh/LITERATURE/Literature.html
4. Hurtienne J, Blessing L (2007) Metaphors as tools for intuitive interaction with technology. metaphorik.de 12:21–52
5. Johnson M (1987) The body in the mind: the bodily basis of meaning, imagination, and reason. University of Chicago Press, Chicago

Chapter 43
The Story Format and the Cycle of Meaning Construction for Physics Education in Primary Schools

Cristina Mariani and Federico Corni

Abstract The story format may provide a stimulating environment, including tasks, questions or problems, giving space for scientific experimentation and group discussions guided by the teacher. In this contribution we present the main advantages of the story format for physics teaching and learning and the features that a story should have in order to implement what we call the "cycle of meaning construction", which constitutes an attempt to integrate the attributes already accredited to the story format in science teaching with pedagogical, methodological and didactic approaches. Lastly, a story will be presented in brief as a possible example for primary school physics education.

43.1 Introduction

In primary schools, the teaching of sciences is a delicate issue, both from the point of view of the children learning and that of the teacher teaching. We suggest and illustrate that a story, appropriately planned and implemented with a series of strategies and activities and supported by a planned methodological approach, may offer a powerful tool as a context in which children and teachers are supported in their needs. From children side a story, in fact, may give value to the baggage of imagination and creativity that the child brings with him/her and which the teacher may draw on.

Moreover the story context gives to children the opportunity to find and to use ordinary words to express their ideas. Furthermore, the structure of the story in terms of experimental activities and didactic interventions based around disciplinary contents constitutes practical support for teachers and at the same time, it helps them to acquire skills in the planning and development of their own science curricula. In this article, we illustrate the features that a story should have in order to implement

C. Mariani · F. Corni (✉)
Department of Education and Humanities, University of Modena and Reggio Emilia,
Via A. Allegri 9, I-42121 Reggio Emilia, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_43,

what we call the “cycle of meaning construction”, i.e., to serve as a methodological support from the point of view both of the scientific disciplines and that of teaching. Lastly, the story “Rupert and the Dream of a Swimming Pool” will be presented as a possible example of application story features for primary schools.

43.2 The Story as Context to Improve the Cycle of Meaning Construction

Our interest here is limited to defining and systematising the characteristics that a story should have in proposing the story format as a medium for science education in primary schools.

The story turns around gaps, interruptions in which all children and teachers activities, related to the problem posed by the story, are inserted. There are gaps to allow for the imagination and reflection. “A story consists of narrated elements and of blanks and gaps” [1]. “Stories can tell the truth, but never the whole truth. (…) Gaps, on the other hand, must be left open to the imagination of the audience” [2].

Gaps may also be reinterpreted from a methodological and didactic point of view as interruptions to the story in which to insert children’s and teachers’ creativity and activities. The alternation between story and interruptions can help to intertwine the levels of logic and reality with those of thought and imagination, as well as encouraging the ongoing transition from one plane to the other. There are interruptions for experimental activities.

Experiments, according to the semiotic mediation framework [5], are carried out using an instrumental approach. The experimental apparatus embodies meanings that pupils have to discover, stimulated by the task. The story supplies the tasks and it may concern artefact or embodied questions [6]. The choose of the task is a critic and important point. Although children find the motivation in a story for studying a given topic, how they then formulate good questions that drive the investigation phase may by no means be taken for granted. Particularly in the early years of primary school, children are not generally able to come up with scientific questions by themselves, so the story must guide the investigation process, asking the right questions in the right order. As far as the scientific investigation method is concerned, H. Fuchs [3] highlights that the process of construction of more or less formal scientific models of natural reality exploits interconnections that oscillate cyclically between posing questions or problems, hypothesising solutions or foreseeing results, and planning experiments or experiencing: what he calls the four-cycle (4C) of hypotheses, questions, experiments and models. Suggesting the questions helps children to find their footing when faced with the other three stages of the four-cycle and, in the future, to learn how to formulate good questions independently. The child, stimulated by a contextualised question, construct situated texts [7] The “situated texts” produced by the children evolve to some form of “scientific texts” on the basis of the pupils’ age [9], fostered through the decontextualisation and recognition of analogies, thanks to teacher iterative didactic cycle (DC) that includes experimental activities, the individual production of output (oral and written texts, drawings, sketches, gestures, gazes and sounds...) and group discussions [8].

The story in science primary school teaching constitutes an attempt to integrate the attributes already accredited to the story with the 4C approach and the DC of semiotic mediation. This integration leads to a recursive sequential structure which we might call "cycle of meaning construction" (CMC). The recursive structure is the following:

- Telling a segment of the story that focuses on a problematic situation
- Presenting the task through a question
- Formulating individual hypotheses (written texts, drawings, etc.)
- Discussing in a group where pupils come together on a number of shared hypotheses
- Experimenting
- Observating and individually interpretating (written texts, drawings, etc.)
- Discussing in a group under the guidance of the teacher in order to reach a common text
- Restarting the story with the overcoming of the problematic situation.

The CMC, by virtue of its integrating three levels, constitutes an example of didactic transposition, and serves as a methodological mainframe for the teacher.

43.3 Features of Story Format and Examples from the Story "Rupert and the Dream of a Swimming Pool"

Having set forth these premises, we shall now propose a list of features to be included in a story for science education. We illustrate them (roman number) and their direct application (roman number-a) to the story "Rupert and the Dream of a Swimming Pool" intended for 4th-grade pupils but may also be suitable for children in the 3rd and 5th grades. This story has been experimented in numerous classroom [10–12].

(I) Each character has aspects with which children can easily identify. The children's identification with the characters is necessary for their emotional commitment and attention. (Ia) The frog Rupert and Pico are positive characters and represent two sides of the figure of the researcher: the former sums up the typical personality traits and behaviour of a child; the latter, the more serious and calculated behaviour of an adult. Students may feel emotionally more drawn to one or the other or to both, depending on their own personal inclination.

(II) The story consists of a series of problematic situations which the characters must deal with and which draw in children's attention and arouse their emotional involvement [13]. Children are thus highly stimulated to marshal their intellectual abilities in the acquisition of new knowledge and skills. (IIa)The swimming pool cannot be filled, or it is placed in the shade, or it overflows: circumstances that force Rupert to move the swimming pool to different heights and different distances from the aqueduct supplying the water, through tubes of various lengths and of different bores, from three different taps to be found at different heights.

(III) The circumstances created by the stories must be believable, yet the fantasy elements arouse children's imagination, attention and curiosity, serving to ensure that they may enter the imaginary world while exiting it just as easily. (IIIa) The story is set in a garden which could be real; the problematic situations are plausible and reproducible through real experiments using real apparatus; the funny or dramatic scenes refer to amusing situations that may occur in keeping with the laws of physics; the drawings are deliberately essential so as not to distract children and help them to focus on the elements to be considered in the study.

(IV) Pupils' involvement in the resolution of the problems is necessary for the story to progress. They are directly called upon from within the story by the characters, who explicitly consult them and ask what should be done. (IVa) Pico addresses directly to the children through the screen and serves as a link between the fantasy world of the story and the real world of the classroom. Pico brings about this interaction by sending the children of the class a case with all the materials needed to carry out the experiments, and during the interruptions in the story he assigns tasks to be carried out and asks them to then send the results back to him; he stimulates group discussion with questions and suggestions, and makes himself available for personal contact with the pupils by e-mail.

(V) The story features a great sense of aesthetics, harmony and order: all those qualities that come together to provide a familiar and serene atmosphere in order to place children in the best possible conditions for learning. (Va) The atmosphere of the story is created by both the narrative and the images.. The images are drawn by a children's illustrator (Arcadio Lobato) who, exploiting a form of beauty capable of arousing emotion, sets the atmosphere against a landscape with various shades of warm, intermingling colours.

(VI) Time is paced by logical events (problems and their solutions, causes and their effects, conceptual sequences, etc.), rather than by mere narrative events. For example (VIa) Rupert has to full fill the swimming pool connected to the aqueduct with difference in level between two containers and with or without inclination of the tube.

(VII) There are gaps to allow for the imagination, experimental activities and reflection (see paragraph 1.2) (VIIa) In "Rupert and the Dream of a Swimming Pool" Pico sent to the pupils the experimental apparatus shown in Fig. 43.1. Questions, related to the semiotic uses of this apparatus, are posed by Pico inside the story to help children to discover that water can be transferred; water flows from higher to lower pressure (represented by higher and lower water levels); the pressure difference provides the push to drive the water through the tube (represented by a difference of water level); the balance corresponds

Table 43.1 Example of pivot words

Narrative steps related to logical events	Children's pivot words	Meaning to be built
…Rupert wants to find a suitable home for his future family. It must be welcoming and have a swimming pool on the top of the hill...	Tube, pump, glass, bucket. Fill, bring	Water can be transferred
…Rupert thought he might link up the swimming pool to a nearby aqueduct with a tube. But only a little bit of water ends up in the swimming pool...	Aqueduct, tube. Upper level It has the pressure; it can go upwards	Water flows from higher to lower pressure (represented by higher and lower water levels)
…Loretta gives instructions to lower the position of the swimming pool compared to the aqueduct…	Up to a certain level; the right level; equal. Lower tap, lower, faster. Running, reaching, overtaking, passing	The pressure difference provides the push to drive the water through the tube (represented by a difference of water level). The balance corresponds to the condition in which the pressure levels are the same (same levels)
…Now the swimming pool is in the right place, it fills up perfectly and the water level is always the same, even when Rupert spills some of the water over the side…		
…Rupert decided to connect the swimming pool to the higher tap, thinking that the favourable inclination of the tube would allow the swimming pool to fill up faster…	Faster. One is emptied and the other is filled. To overflow	The pressure difference provides the push to drive the water through the tube (represented by a difference of water level). The driving force is not influenced by the inclination of the tube
…Rupert decided to connect the swimming pool to the lower tap...	From low down upwards. It doesn't fill up. Manage to pass. The push for it to get through anyway	

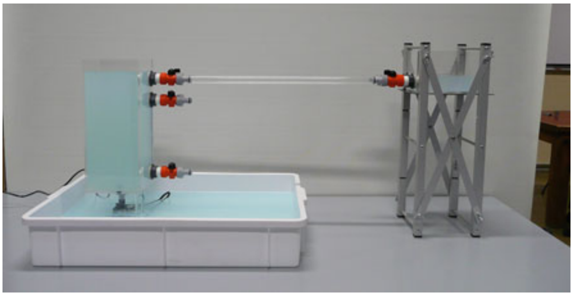


Fig. 43.1 shows the experimental apparatus sent by pico to the pupils, set up as per the problems presented in story presented in "Rupert and the Dream of a Swimming Pool"

to the condition in which the pressure levels are the same (same levels); the driving force is not influenced by the inclination of the tube.

(VIII) There are gaps to insert teacher didactical cycle for for teacher-led discussions (see paragraph 1.2). There is a knowledge to be pieced together by the child under the guidance of the teacher in light of the various experimental and didactic activities proposed during the interruptions (VIIa). In order to make the didactic aspect more explicit, In Table 43.1 we list a number of pivot words (second column) extrapolated from sentences uttered by children during the pilot experimentations. The pivot words may, in turn, be placed in relation to meanings to be constructed. (third column).

43.4 Conclusions

A story, put together with the features identified and presented here, may have a positive impact on the teaching/learning process. We wish to emphasise the importance of the story with its capacity to engage pupils both emotively and cognitively, interspersed with experimental activities and discussions led by the teacher, aimed at constructing scientific meanings.

The acquisition of the scientific method and of meaning construction skills are some of the key goals of science education. This is a very important aim considering that both the scientific method and meaning construction, in view of a future conceptualisation, are the fruit of a complex interconnection of specific cognitive tools, referred to as thinking and acting, as questions, experiments, hypotheses and modelling. This structure becomes even more complex if we consider that young pupils are unable to form scientific questions by themselves but must be taught to do so.

References

1. Sternberg M (1978) Expositional Modes and Temporal Ordering in Fiction. John Hopkins University Press, Baltimore
2. Kubli F (2001) Can the theory of narratives help science teachers be better storytellers? Sci Educ 10:595–599
3. Fuchs H U (2007) From Image Schemas to Dynamical Models in Fluids, Electricity, Heat, and Motion. An Essay on Physics Education Research https://home.zhaw.ch/~fuh/LITERATURE/Literature.html
4. Fuchs HU (2010) The dynamics of heat: a unified approach to thermodynamics and heat transfer. Springer, New York
5. Corni F, Giliberti E, Mariani C (2011) Semiotic mediation and didactical cycle as methodological reference for primary school teachers. The biannual ESERA conference 2011
6. Mariani C, Laurenti E, Corni F (2011a) Hands-on, minds-on activities to construct the concept of energy in primary school: Experiments, games and group discussions, International Conference on Physics Education ICPE 2011 proceedings
7. Bartolini Bussi MG, Mariotti MA et al (2008) Semiotic mediation in the mathematics classroom: artifacts and signs after a Vygotskian perspective. In: English L (ed) Handbook of international research in mathematics education, 2nd edn. Routledge, New York and London, pp 746–783
8. Bartolini Bussi MG (1996) Mathematical discussion and perspective drawing in primary schools. Educ Stud Math 31:11–41
9. Mariani C, Corni F, Giliberti E (2011b) A pr oposal for a didactical path for age 5–8 on the concepts of extensive quantity using stories as cognitive tool. The biannual ESERA conference 2011
10. Corni F, Giliberti E, Mariani C (2010) A story as innovative medium for science education in primary school, GIREP-ICPE-MPTL International Conference 2010 Teaching and Learning Physics today: Challenges? Benefits? August 22–27. Reims, France
11. Pagliaro M, Mariani C (2011) Come e perché l'acqua si muove: la storia della rana Rupert. Atti del Convegno Innovazione nella didattica delle scienze nella scuola primaria: al crocevia fra discipline scientifiche e umanistiche. Ed. Artestampa, Modena
12. Sedoni C, Mariani C (2011) La piscina di Rupert: un percorso sperimentale sullo spostamento dei fluidi per la scuola primaria. Atti del Convegno Innovazione nella didattica delle scienze nella scuola primaria: al crocevia fra discipline scientifiche e umanistiche. Ed. Artestampa, Modena
13. Casey B (2004) Kersh JE, Young JM (2004) Storytelling sagas: an effective medium for teaching early childhood mathematics. Early Childhood Res Q 19(1):167–172

Chapter 44
Teaching Modern Physics for Future Physics Teachers

E. F. Nobre, A. O. Feitosa, M. V. P. Lopes, D. B. Freitas, R. G. M. Oliveira, M. C. C. Filho, N. M. Barone, R. Monteiro and M. F. S. Souza

Abstract In this work we present a new way for teaching Modern Physics to students of course for teachers formation. One of the great difficulties faced by the teachers of the secondary schools of the State of Ceará, in Brazil, may have its origin in their formations during the undergraduate courses. We aim to offer a course in modern physics so that future physics teachers can transmit this knowledge to their future students in a clear, simple way and with good results, awakening in their future pupils the interest by the study of Sciences in general and in Physics, in particular.

44.1 Instructions

Since year 2000, when Modern Physics was included in the program of examination for entrance of students in the Federal University of Ceará (UFC), the secondary schools began to include Modern Physics into their programs. The big problem was that this inclusion has been made simply Since year 2000, when Modern Physics was included in the program of examination for entrance of students in the Federal University of Ceará (UFC), the secondary schools began to include Modern Physics into their programs. The big problem was that this inclusion has been made simply as an appendix to the contents of the third year of high school, almost in the end of the school year. In addition, many teachers were not familiar with this subject and thus the inclusion of Modern Physics was compromised.

Aiming to help solving this problem, we developed a discipline of Modern Physics to be applied to students of the Undergraduate Course in Physics from the Federal University of Ceará (UFC), sponsored by the Virtual UFC Institute. The course is specific for training teachers in pre-service and is applied to 5 cities in the State of

E. F. Nobre (✉) · A. O. Feitosa · M. V. P. Lopes · D. B. Freitas · R. G. M. Oliveira · M. C. C. Filho
Federal University of Ceará, Campus do Pici, Fortaleza, Ceará, Brazil

N. M. Barone · R. Monteiro · M. F. S. Souza
Virtual Institute-Federal University of Ceará, Campus do Pici, Fortaleza, Ceará, Brazil

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_44,

Ceará. The Virtual Institute of Federal University of Ceará (Virtual-UFC), besides creating, implementing and maintaining the technological core, organizes the physical structure, logic and the educational progress of each course. All over the world, the electronic learning (e-learning), a model of teaching based on the online environment, leveraging the ability of the Internet for communication has been an excellent way to open the University to the student's home, allowing a high degree of flexibility in any academic program. There is no doubt that the multimedia technology is a powerful tool in the development of e-learning contents, which is inserting deep changes at higher education [2, 6]. The use of e-learning mediated by the Information and Communication Technologies (ICTs) in the educational process, has favored the inclusion of a contingent of people in higher education, contributing to increase places in higher education. This work is part of project developed at the Learning Environment of the Virtual Institute of the Federal University of Ceará, in Brazil, called SOLAR. The most part of courses which are being developed in this e-learning environment are focusing on the teachers formation. The production of teaching materials for e-learning plays a more significant relevance role than those ones we use for presence courses. In this process the steps of planning, construction, mediation and evaluation inherent to the traditional teaching, plus the endless possibilities of combining the diversity of information available and the benefits of interactivity, flexibility and dynamism favored by the ICTs require the implementation of this material in a more safely way. We must consider that the students are not face to face with their teachers, so they need a material absolutely clear and easy of understanding. By developing the didactic material that is used at the distance courses, the didactic transition of the Virtual UFC, uses as a methodologic reference the socio-interactionism [5]. In this context, the education becomes a dynamic task that allows to the students, themselves, the active construction of their knowledge, according to their experiences in different situations where they live.

44.2 Methodology and Discussions

The material concerning of Modern Physics was developed by a team composed by a Professor, who was in charge of all the contents of Physics; Educators, experts at Pedagogy who are in charge of the Didactic Transition (DT) and technicians in ICT. This last team is composed by specialists in various areas of computing who use the multimedia technologies in order to transform the texts written by the Professor in e-materials. So, each class presents animations, retractable texts, vector animations, links to additional readings, chats, forums, and all kinds of resources currently used in digital technologies focusing on education. The use of these tools enables the integration of various resources in a single learning environment, and encourages the adoption and understanding of audiovisual language. Our main goal is develop e-lessons in order they are presented to the students in a very clear, simple and attractive way. Nevertheless our structure is developed based on the idea of the use of on line material, all materials of the virtual courses of the UFC, are available on line and in printed form for the students.

The course was developed in six lessons covering the following topics: Special Relativity; Thermal Radiation and the Origin of Quantum Theory; Atomic Models: Thomson, Rutherford, Bohr; Wave-Particle Duality and the Principles of Quantum Theory and Nuclear Physics Topics. The teaching material was structured in order to bring a bit of History of Physics in all contents. In some subjects, we developed animated lessons, with audio. In all classes the subjects were presented with simulations, animations, music, challenges that were intended to excite students to seek solutions to challenging problems, suggestions of sites for complementary reading and research and much supporting material, discussion forums, chats and an special virtual space called portfolio, where the students presented the solutions of the exercises. In nuclear physics lessons we also developed a special tool called Dynamic Book, which are virtual books to complement the covered topics. The structure of lessons is showed below:

Lesson 1: Special Relativity–8 Topics

- Topic 01: Galileo's Transformations Review
- Topic 02: The Michelson-Morley Experiment
- Topic 03: The Postulates of Relativity; Simultaneity
- Topic 04: Time dilatation
- Topic 05: Space Contraction
- Topic 06: Lorentz transformations
- Topic 07: Relativistic Dynamics-Relativity of mass
- Topic 08: Relativistic Dynamics: Energy

Lesson 02: Thermal Radiation and the Origin of Quantum Theory–5 Topics

- Topic 01: Black Body Radiation-Principles
- Topic 02: Black Body Radiation-Theory
- Topic 03: Photoelectric Effect, The Quantum Theory of Radiation
- Topic 04: Compton Effect
- Topic 05: X-Rays

Lesson 03: Atomic Models 1–4 Topics

- Topic 01: Thomson Model
- Topic 02: Successes and Failures of the Thomson model
- Topic 03: Model of Rutherford. Experiment Rutherford Planetary Model
- Topic 04: Successes and Failures of the Rutherford model

Lesson 04: Atomic Models II-Atom quantized Model–4 Topics

- Topic 01: Atomic Spectra, Spectrum Series, Bohr's Postulates
- Topic 02: Electron orbits, the Hydrogen Atom
- Topic 03: Energy Quantization
- Topic 04: Successes and Failures

Lesson 05: Wave-Particle Duality-4 Topics

- Topic 01: de Broglie Waves
- Topic 02: Diffraction Particle
- Topic 03: Principle of Uncertainty
- Topic 04: The Wave Function

Lesson 06: Topics in Nuclear Physics-4 Topics

- Topic 01: The Nucleus
- Topic 02: Nuclear Reactions
- Topic 03: Radioactivity
- Topic 04: Radiological Risks to Health

The lessons were developed in order the students could use the new multimedia technologies with the use of computers and the internet, but besides these virtual recourses, the students had six meetings at presence moments with their Teachers/Tutors during the application of discipline. There was one Teachers/Tutors for each city. These teachers/tutors, some of them are specialists in Physics Teaching, some with Master's degree in Physics and some finishing his doctoral degree, accompanied the students during the time that the lessons are applied. This team of Teachers/Tutors was coordinated by the same Professor who produced and wrote all the material and by the Coordinator of tutors, a Professor whose function is to monitor the progress of the work of Teachers/Tutors. In addition to face meetings, the students were oriented at a distance by their teachers/tutors through the discussion forums, chats, and all the resources currently available by the digital technologies.

The method of evaluation of the course in e-learning is different from what is done in traditional courses. In this course the students were evaluated personally and virtually. The students had several small conceptual evaluations at the end of each face meetings with their teachers/tutors. The goal of these evaluations was to provide students a better understanding of physical concepts. There was no math in these evaluations. Besides these small tests, the students did a big test which required physical concepts and also the mathematical tools. The virtual evaluations were distributed among the six discussion forums and the six lists of exercises solved by students and posted on the solar environment. The final average was formed with the results of all evaluations: in presence and virtual evaluations. For the student who did not achieve the minimum average, a kind of GPA (Grade Point Average) required for approval by the University, was offered a final exam: an written exam covering all topics studied. In addition to the material produced especially for this course, the students were guided to consult also some basic textbooks of Modern Physics (Acosta, Beiser, Eisberg-Resnick, Serway, Tipler). In addition to basic textbooks on the subject, the student also received a printed manual containing all lessons that were available on the website of the course. This greatly favored those students who still not have a computer in their homes.

The results of 31 students who participated of this project were excellent. The most of them was approved with A concept and there was a minimal dropout. These results are showed in figure bellow.

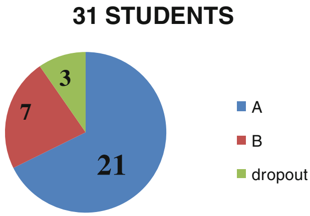


Fig. 44.1 General results of all students

44.3 Concluding Remarks

As a result of this way of teaching Modern Physics, we obtained one of the best results so far, since the beginning of the Physics Course in 2007. The students felt truly attracted by this subject. The abandonment was minimal, the frequency in the meetings was almost total and participation in activities was massive. The results showed approval ratings in the discipline between 80 % to 100 %. Next year, first half of 2012, we will apply the lessons of Modern Physics to all students, around 11 cities, involving about 100 students. We hope that the project will have again the same acceptance by the students and they are as successful as were the students of this first class. Another expected result is the involvement of students who attended the course acting as monitors in the next application of the project in 2012. Playing a role as monitors of the discipline, they will act under the guidance of teacher / tutor of the class, helping the new students, answering questions, forming study groups, which may bring greater benefits in learning. We also expect that helping to colleagues, may promote a consolidation in student learning that will act as a monitor.

References

1. Acosta V, Cowan CL, Grahan BJ (1975) Curso de fisica moderna. Mexico: Harla, p 512
2. Baggaley J, Hoon M, Ng L (2005) Pandora's box: distance learning technologies in Asia, learning, media and Technology v 30, (1)
3. Beiser Arthur (1981) Concepts of modern physics, 3rd edn. McGraw-Hill, New York
4. Eisberg, Robert, Resnick, Robert (1974) Quantum physics of atoms, molecules, solidas, nuclei and particles, John Wiley and Sons Inc., p 928
5. Fiorentini LE, Morais R (2000) Languages and interactivity in distance education, São Paulo, Brazil. P&D
6. Kirkup G, Kirkwood A (2005) Information and communications technologies (ICT) in higher education teaching: a tale of gradualism rather than revolution.Learning, Media and Technology
7. Serway A, Raymond MJ, Clement MA, Curt (1997) Modern Physics. Saunders College Publishing, 2*a*.ed. p 666
8. Tipler PA (1981) Modern physics. Rio de Janeiro: Guanabara Dois, p 422

Chapter 45
An Alternative Approach to Canonical Quantization for Introducing Quantum Field Theory: The Double-Slit Experiment Re-Examined

Eugenio Bertozzi and Olivia Levrini

It is in vain that we say what we see; what we see never resides in what we say. And it is in vain that we attempt to show, by the use of images, metaphors, or similes, what we are saying; the space where they achieve their splendour is not that deployed by our eyes but that defined by the sequential elements of syntax.
(Michel Foucault, The Order of Things)

Abstract In the last years a growing research concern, within physics education, has been addressed to the production of teaching proposals for introducing notions of quantum field theories (QFT) at the secondary school level. The proposals are usually the result of the effort of translating the most widespread approach to QFT in university textbooks, "canonical quantization", into natural language. After a discussion of the pros and cons of taking canonical quantization as reference, an alternative educational approach is presented. The approach is not yet a teaching proposal but a set of criteria needed for extracting the conceptual and cultural essence of QFT to be taught also without sophisticated formalisms.

45.1 Introduction to the Research Problem

Quantum Field Theory (QFT) is without doubt an advanced and specialized topic that only a part of university physics students is required to deal with. Nevertheless, a growing interest about teaching QFT can be observed in recent years and a certain number of studies exists aimed at producing teaching proposals for introducing notions of QFT at the secondary school level or within introductory physics courses at the university level (see for example [4]). The main motivations of such a growing interest stem from the requirements of secondary school physics curricula towards

E. Bertozzi (✉) · O. Levrini
Physics Department, University of Bologna, Bologna, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_45,

contemporary physics topics [9] and from the need of tuning school and extra-school activities: particle physics, the standard model, the last frontiers of physics are indeed object of popular science books and of important and successful exhibitions.

According to an architectural metaphor, QFT can be seen as an imposing building of linguistic engineering, built of very fine materials and assembled by refined craft-like experience; on the other hand, in its usual presentation in university textbooks, it looks like a tangle of threads and formal structures fit one into the other; a 'quarrel' between general-particular, new-old, physical-metaphysical, formal-linguistic aspects.

Entering such a tangle, pointing out its essential conceptual structure, analyzing it from the specific perspective of Physics Education Research (PER), so as to arrive at designing a teaching proposal for secondary school students, is the main goal of the research work we are carrying out.

The specific aims of the present paper are:

- To show that outreach activities and school teaching proposals are usually the results of the effort of translating the most widespread approach currently utilized in university textbooks (canonical quantization) into a familiar language (§2);
- To discuss the pros and cons of the choice of taking 'canonical quantization' as approach of reference for outreach and teaching proposals (§3);
- To present the main features of an alternative approach to QFT that can act as reference for designing outreach activities and school teaching proposals (§4).

45.2 Outreach Languages and School Teaching Proposal on Contemporary Physics

Steven Weinberg, in one of his popularization books on contemporary physics, introduces the notion of elementary particle as follows ([11], p. 25):

> "Furthermore, all these particles are bundle of the energy, or quanta, of various sorts of fields. […] There is one type of field for each species of elementary particles; there is an electron field in the standard model, whose quanta are electrons; there is an electromagnetic field (consisting in electric and magnetic field) whose quanta are photons; [...]"

Although no explicit formulas appear in the previous explanation, the opacity and technicality of some terms and expressions utilized by Weinberg suggest the idea that a certain formal apparatus underlies the whole argumentation. In particular, statements like these seem to be not only a 'direct translation' of QFT into natural language, but also, more specifically, a direct translation of the most widespread approach to QFT in the university textbooks: the canonical quantization procedure.

The procedure focuses on the electromagnetic field and on the process of quantizing it by replacing, with operators, the numbers representing the coefficients of the Fourier expansion of the solutions of the d'Alembert equation. Every other quantum field (for example, the quantum Dirac field for relativistic spin 1/2 and m mass particles or the quantum Klein–Gordon (KG) field for relativistic 0 spin and m mass particles) is indeed constructed extending such a formal procedure by analogy: take the generic solution of a wave equation expressed as the Fourier expansion on plane waves, focus on the coefficients of the expansion and elevate them from numbers to operators by defining their commutation (bosons) rules or anti-commutation (fermions) rules. As Mandle & Shaw write: *"From the quantization of the electromagnetic field one is naturally led to the quantization of any classical field, the quanta of the field being particles with well defined properties"* [8] .

Weinberg's expressions, like canonical quantization:

- Provide a paradigmatic role of the electromagnetic field;
- Seem to emphasize more the analogy between radiation (electromagnetic field) and matter fields than the differences among them.

Remarkable echoes of the approach described above can be individuated also in the context of PER. The teaching proposal of Art Hobson ([6]) is a significant example of that. Targeted to secondary school students, such a proposal introduces the notion of "particle as field quanta" by means of the comparison of the two interference patterns created by radiation and matter in the context of the double slit-experiments at low intensity. In particular, after observing the same behaviour between light and matter beam, the author provides the following interpretation:

> "Just as Fig 1.[the interference patter of light at low intensity] is evidence that light is a wave in a field, Fig. 3 [the interference patter of electrons at low intensity] is evidence that matter is a wave in a field—an extended real physical entity that comes through both slits and interferes with itself". As far as the relation between matter field and electron is concerned, the author remarks: "That is, when we say that 'an electron came through double slits,' we really mean that an extended single excited field (space-filling) came through the double-slit" (p. 631).

The key-linguistic expressions used by Hobson ("particle as field quanta", "matter is a wave in a field") are very close to Weinberg's words and, in our opinion, they share with Weinberg the same weaknesses:

i) What do such expressions mean for people or students lacking the formal background on which the reasoning is however grounded?
ii) If the language is taken literally, what does a sentence like "an extended single excited field (space-filling) came through the double-slit" say, when referred to an electron?
iii) Where and how a student is guided to grasp the differences between a classical field and a quantum one?

45.3 Pros and Cons of the Canonical Quantization Procedure

The pros of canonical quantization are very evident at university level and concern its technical effectiveness: the formal analogy is a short and direct way to construct the new theoretical entities.

The cons, at least in our opinion, are particularly evident when this approach is taken as a reference for outreach and in designing teaching proposals: in these cases, the lack of formalism on which the procedure is based leads the translation to be conceptually opaque and at risk of giving back a very simplified, semi-classical image of fields and particles.

In particular:

- A students is encouraged to associate to each particle a field whose image is almost the same of the classical electromagnetic field;
- Particle and fields are presented as two entities belonging to the same "level" of existence (particles are oscillations of a field coming through both the slits, Hobson).

These two issues are questionable, as well as from the PER perspective, also from the perspective of the Foundations of QFT:

i) on the field-particle duality, Haag remarks:

> "*Yet the belief in a field-particle wave duality as a general principle, the idea that to each particle there is a corresponding field and to each field a corresponding particle has also been misleading and served to veil essential aspects. The role of field is to implement the principle of locality*". (Haag 1996, p. 45–46)

ii) on the "different levels", Julian Schwinger, one of the fathers of QFT, underlines:

> "*Until now, everyone thought that the Dirac equation referred directly to physical particles. Now, in field theory, we recognize that the equations refer to a sublevel. Experimentally we are concerned with particles, yet the old equations describe fields.... When you begin with field equations, you operate on a level where the particles are not there from the start. It is when you solve the field equations that you see the emergence of particles*".

45.4 Alternative Approach to Quantum Field Theory

In order to overcome the cognitive, linguistic and philosophical weaknesses of canonical quantization, an alternative educational approach has been built. The approach is not yet a teaching proposal but a set of cultural and operative criteria needed for extracting the conceptual and cultural essence of QFT to be taught also without sophisticated formalisms. In this sense, the construction of the approach represents the preliminary stage needed for preparing the design of a teaching proposal for secondary school students.

45.4.1 Cultural Choices

The cultural choices that characterize the approach are the following:

1. The introduction/discussion of the notions of quantum, quantum field and particle is developed along a *multilevel structure*, i.e. a structure where the levels of phenomenology, formalism and interpretation are related but distinguishable, as suggested by the research on foundations of physics and PER;
2. In the wake of Bohr, the limits of natural (classical) language in translating quantum formalism are taken as a inner feature of contemporary physics and not as a problem to overcome. Familiar images, as well as good metaphors, take indeed their power in bridging the gap between the rich world of experience and the formal physical re-construction of such a world. The gap is in-principle bridgeable when the mathematical reconstruction of the world provides a "projection" of the world itself in a Euclidean space, i.e. when the properties conceptualized by the mathematical description can be re-synthesized in a representation that "lives" in a space somehow "isomorphic" to the space we experience. This game becomes more and more problematic in quantum physics, since the new mathematical description projects the real world in highly abstract, unfamiliar spaces, such as the Hilbert space. "Diagram space" cannot be simply related to "real-world space", and "diagram objects" do not have that set of properties, one is usually led to reassembled in something recognizable as an image of "real objects". As a consequence, the complex relationship between QFT formalism and natural language is assumed to contain a profound cultural value that ought to be exploited by showing the necessary partialness of familiar images or metaphors in giving back how the formalism models quantum objects and their interactions.
3. In order to introduce the notion of quantum field, a *basic* and *essential* conceptual structure is constructed: *basic* means that it does not have to refer to any specific quantum object but it must embody the *common* features of all the quantum objects; *essential* means that it must include *only* what is needed to justify phenomenological evidences. Only in a second stage, the basic and essential structure is *enriched* so as to take into account the "particles zoo". In the logics of searching, first, for what is common to all the quantum objects and, then, for what is specific to the various quantum objects, the phenomenological and conceptual analogy between radiation and matter is presented and where the analogy breaks downs is stressed.

45.4.2 Operative Criteria: Rough Structure of a Conceptual Path

The cultural choices of the approach have been implemented in a *rough* conceptual path that should operatively orient the future work aimed at designing a teaching proposal.

The conceptual path foresees two macro-steps.

The first step is based on the operative choice of referring to the double slit experiment at low intensity, as Hobson suggests, as an effective experimental basis for emphasising the analogy in behaviour of radiation and matter and for introducing, from such an analogy, the notions of quantum and quantum field.

The double slit experiment allows both phenomenological constraints to be selected (emission and detection of a quantum object, formation of an interference pattern of the screen, spot by spot), and the "basic and essential conceptual structure" to be constructed for introducing the notion quantum field. The conceptual structure includes: (a) time-like and space-like correlations (the quantum object propagates from where it is emitted up to where it is revealed and it is spatially extended) needed for justifying the formation of the interference pattern on the screen; (b) the discreetness of dynamical quantities needed for justifying the spots progressively recorded on the screen.

The emphasis on time-like and space-like correlations implies a selection of a few features of a wave-like formalism so as to stress that, when the wave metaphor is used, it cannot be freely extended up to where the common use of the metaphor could lead. For example, it cannot be extended for modelling the superposition of two quantum fields as a classical sum of amplitudes: amplitudes in the formalism of QFT become operators acting on a Fock space.

The focus on the discreetness of dynamical quantities, involved in the interaction with the screen, allows the metaphor of particle to be problematized and discussed within the model of interaction embodied into the theoretical construct of quantum field. This model of interaction is very different from the model expressed by a classical field, like the electromagnetic one. Roughly speaking, the mediation enacted by a classical field between two systems *can* be represented in ordinary three-dimensional space. The transformation of the coefficients of the Fourier expansion of a quantum field into creation and annihilation operators introduces a fundamental change: it introduces the need of referring to an abstract space (the Fock space), where the whole set of possible infinite interaction processes must be explored. In other words, the quantum model of interaction requires a preliminary construction of an abstract "space of possibilities" (represented by Feynman diagrams) that cannot be represented in ordinary space. The model implies hence a fine imaginative game needed for keeping together a not trivial space-time description of the quantum object propagation, the abstract space of all the possible interactions and the physical place where quantum systems are detected. According to our cultural perspective, it would be a pity to short-circuit such a game by using pseudo-familiar images and by "simply" telling that, in quantum interaction, *"the force at distance between two charges is nothing but an exchange of a virtual photon"* ([1], p.473).[1]

In order to stress both the counter-intuitive phenomenological behaviour of the quantum object and the need of taking the phenomenological level apart from the formal/conceptual one, we suggest a specific use of the words *quantum* and *quantum field*. In the wake of [3, 7], we suggest to use the word *quantum* just for fixing the point that the quantum object shows some phenomenological properties (propagation, diffraction/interference and discreteness in interaction) that cannot be all contained in a single image of wave or particle. The word *quantum field* is, instead, used for

[1] The book of Amaldi is one of the most widespread physics textbooks in Italy for upper secondary school.

indicating the formal/theoretical construct able to keep together the properties of *quantum*.

The second macro-step concerns the enrichment of the basic formal structure so as to show where each field gains those properties that allow it to be associated to a specific particle. In order to describe a *particle* it is indeed necessary to move from a model defined by the common properties of radiation and matter to the "particles zoo" constituted by electrons, photons, muons etc. From an operative and experimental point of view, this means to take into account what happens, for example, in the big-science experiments in order to justify the need of providing the notion of quantum with further specifications like mass, electric charge, spin and other quantum numbers (strangeness and so on). From a conceptual point of view, enrichment means to introduce new symmetries (conservation laws) needed to apply the model of interaction, embodied into the notion of quantum field, to the analysis of concrete processes ([10]).

From the more general perspective of the foundations of QFT, the process of enrichment is crucial for pointing out where the analogy between radiation and matter breaks down. Although radiation and matter present common behaviours in the double slit experiments, there is a fundamental difference between them: whilst a classical theory of the electromagnetism does exist, no classical theory of the matter fields exists, as it is formalized by the Gupta-Bleuer condition which is the quantum analogue of the Lorentz condition. Such a constraint makes photons (and the other interaction particles) very different from matter particles.

45.5 Final Remarks

In the paper, an approach to QFT alternative to canonical quantization has been presented. Three main elements remark, in our opinion, its peculiarity with respect to canonical quantization:

- Instead of playing a prototypical role, the electromagnetic field is addressed as the last step of the process of complexification of the basic structure and the specific feature of the photon of being "an interaction quantum bounded to a classical theory of fields" is stressed;
- The formal analogy, chosen by the canonical quantization as tool for building the different quantum fields on the basis of the electromagnetic one, is replaced by a multi-level structure where the notion of quantum, quantum field and particle are progressively built so as to stress what all the quantum objects have in common (both phenomenologically—the quantum—, and conceptually—the quantum field) and what makes the various objects different (particles);
- Instead of stressing mainly the continuity between natural language (or classical images) and QFT formalism, the limits of the former are assumed to contain one of the most important cultural messages to be passed to secondary students.

References

1. Amaldi U (2001) Le idee della fisica moderna. Zanichelli, Bologna
2. Bertozzi E (2010) Hunting the ghosts of a ‘strictly quantum field’: the Klein-Gordon equation. Eur J Phys 31:1499–1515
3. Dirac P (1967) (1930) The principles of Quantum Mechanics, 4th edn. Clarendon Press, Oxford
4. Giliberti M (2008) Campi e particelle: introduzione all’insegnamento della fisica quantistica. In: Guidoni P, Levrini O (eds) Approcci e proposte per l’insegnamento-apprendimento della fisica a livello preuniversitario. Forum Editrice, Udine, pp 161–209
5. Haag R (1992) Local Quantum Physics, Fields, particles, Algebra, 2nd Revised and Enlarged Edition, Springer - Verlag, Berlin, Heidelberg, New York
6. Hobson A (2005) Electrons as field quanta: a better way to teach quantum physics in introductory general physics courses. Am J Phys 73:630–634
7. Levy-Leblond JM (2003) On the nature of quantons. Sci Educ 12:495–502
8. Mandle F, Shaw G (1984) Quantum field theory. John Wiley & Sons Ltd., Chichester, New York, Brisbane, Toronto, Singapore
9. Miur (2010) Italian Ministry of Instruction, University and Research, National indications for upper secondary schools, http://nuovilicei.indire.it/content/index.php?action=lettura&id_m=7782&id_cnt=10497
10. Hoekzema D, Schooten G, van den Berg E, Lijnse P (2005) Conservation laws, symmetries, and elementary particles. Phys Teac 43:266–271
11. Weinberg S (1992) Dreams of a final theory. Vintage Books. Inc, New York

Chapter 46
Basic Concept of Superconductivity: A Path for High School

Marisa Michelini, Lorenzo Santi and Alberto Stefanel

Abstract The need for renewed physics curricula, it suggests to include topics of modern physics. The phenomenology of superconductivity is interesting, because: macroscopic evidence of quantum processes; presentable with simple and motivating experiments; involving interesting technological applications. An approach designed to introduce superconductivity in the high school, based on experiments, aimed to recognize the change in electric and magnetic properties of an YBCO sample at phase transition is presented with some results about experimentations with students.

46.1 Introduction

Teaching and learning Modern Physics is a challenge for Physics Education research involving multidimensional perspectives [1–5]. The results of different international surveys show a diffuse illiteracy in the scientific areas and from different perspectives [6], that motivated a diffuse reflection on how physics is taught, how can be proposed in a renewed ways [7, 8]. The curricular researches stress the need to reorganize classical physics as physics in context where the technological apparatuses used in everyday life become objects of study to develop concepts and theory, classical principles based techniques used in research offer to the pupils experiences of methodologies of physic ([9, 10], Cobai et al 2011). It seeks effective bridges from classical to quantum physics, to create continuity between what the students traditionally faced and the new topics [2, 9–11] also for education of Quantum Mechanics [1, 4]. Research experimentations on modern physics evidence the feasibility of the introduction of Quantum Physics in Upper Secondary, particularly when the focus is on the basic concepts [12, 13] or on the integration of physic and technology [10]. Empirical studies showed some typical learning difficulties of students in the acquisition of a a quantum mechanical way of thinking and also made clear the

M. Michelini (✉) · L. Santi · A. Stefanel
Research Unit in Physics Education (URDF-UNIUD), University of Udine, Udine, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_46,

effective learning produced by coherent research based teaching/learning sequences [13].

In this perspective, superconductivity offers many opportunities to explore phenomena very interesting for students because challenges that stimulate the construction of models, activate a critical re-analysis of their knowledge on magnetic and electrical properties of materials, stimulate links between science and technology, thanks to really important applications (i.e Maglev Trains, supermagnets for physics research and MNR test) [14, 15].

We developed, in the collaboration of the European projects MOSEM 1-2 [16], an educational path for high school to introduce superconductivity, integrating it in the courses of electromagnetism. The educational path developed implement an IBL approach using a set of hands-on/minds-on apparatuses designed with simple materials and High Technology, YBCO samples, USB probe to explore resistivity versus temperature of solids [17].

The present paper focus on the rational of the educational path, presenting its main steps aiming the following research questions: Which conceptual elements of superconductivity can be introduced from phenomenology? How to construct a coherent description of the magnetic and the electric properties of a superconductor? How to introduce the main aspects of the BCS theory?

Some general results are also reported concerning the students learning, referring to a future work for a more extensive discussion of it.

46.2 Superconductivity and Classical Physics

Superconduction is a quantum mechanical phenomenon to a large extent as implied in the Ginzburg-Landau theory [18], in the BCS theory [19], in the Josephson effect [20]. In spite of this, [21] pointed that "According to basic physics and a large number of independent investigators, the specific phenomenon of flux expulsion follows naturally from classical physics and the zero resistance property of the superconductor—they are just perfect conductors". In particular they showed, following [22] and their own work [23], that the Meissner effect ($B = 0$ inside a superconductor) can be derived just from Maxwell equation applied for a $R = 0$ conductor. At the same time they stressed on the possibility to use classical electromagnetism (without ad hoc hypothesis) the derivate the London equations, usually framed coherently just in the BCS theory [19]. [24], aiming to understand stable levitation when a magnet is posed on a superconductor, proposed a treatment of superconductors of I and II type to give into account Meissner effect as well as pinning where "the main concepts involved are electromagnetic energy, thermodynamic reversibility and irreversibility, Lenz-Faraday's law of induction, and the use of variational principles".

These theoretical results ensure that the phenomenology of superconductive state, almost for what concern the Meissner effect, can be described in the framework of classical electromagnetim. From an educational perspective, the concept of electromagnetism (magnetic field, flux, magnetization) can be used as tools to explore the phenomenology of superconductivity. In this perspective the Meissner effect ($B = 0$ inside a SC) can be explained as an em induction process occurring in an ideal conductor ($R = 0\ \Omega$). The London equations constitute the theoretical (implicit) reference for the exploration and a possible intermediate goal of the educational path. The attempt to explain how the $R = 0\ \Omega$ condition can be created inside a superconductor bridges toward the quantum mechanical frame for the creation of the superconductivity state.

46.3 The Rational of the Path on Meissner Effect for Upper Secondary School Students

The educational path approaches the Meissner effect through an experimental exploration of the magnetic properties of a superconductor sample (SC—a disc of YBCO with very weak pinning effect), aiming to the identification of the diamagnetic nature of a SC.

An YBCO disc at room temperature (T_e) does not present magnetic properties. When it reaches the temperature of LN (T_{NL}), the evident levitation of a magnet on it poses the problem to individuate what kind of magnetic property has or acquire a SC when T is close to T_{NL}. A systematic exploration of the interaction of the SC with different magnets and different objects (ferromagnetic objects in primis), with different configurations, put in evidence that a repulsive effect is ever observed when a magnet is posed close to the SC and no interaction is observed when a ferromagnetic object is putted close the SC alone. These exploration leads to the conclusion that: when $T = T_e$, a SC does not evidence magnetic properties; when $T \sim T_{NL}$, suddenly, magnetic properties emerge (there is a phase transition?). Moreover the SC is not ferromagnetic, it do not transmit the magnetic field inside, it is not a permanent magnet, it do not behave like a magnet, it do not become like the mirror image of the levitated magnet. It always shows repulsive effects close to a magnet, like all the diamagnetic objects placed near a magnet. For that a SC at $T \sim T_{NL}$ must be classified diamagnetic. The strength of the interaction between a SC and a magnet, several orders of magnitude greater than those observed with ordinary diamagnetic materials, suggest to give a more detailed characterization of the nature of the diamagnetism of the SC.

Starting from the evidence that the SC shows magnetic interaction only when a magnet is close to it and that the SC do not interact with a ferromagnetic object, it is possible recognize that the interaction with a magnet do not depend on the pole put close to the surface of the magnet, the equilibrium position is always the same. Changing magnet (inside a range of course), the equilibrium position changes but is always the same, for the same magnet. Moreover when $T_{NL} < T < T_e$ the B field can be different by 0 inside the YBCO sample (a magnet interact strongly with an iron ring also when a YBCO disc is putted in between), but when $T \sim T_{NL}$ the magnetic field inside the SC sample must be very little (in this condition the magnet and the iron ring do not interact when the YBCO disc is in between). The magnetization of the SC is always adjusted to react to the external magnetic field, tending to preserve the initial situation. In particular if $B = 0$ when the superconductivity state is created, the system tend to react to an external magnetic field creating a counter field that tend to maintain $B = 0$ inside the SC (Meissner effect). It is possible to find a similar behavior in the case of eddy current produced in the electromagnetic induction. For instance a magnet is evidently braked when falling down on a thick copper layer or inside a copper tube. It is evident however that never can be stopped, as in the case of the SC. In fact If the magnet were to stop, the induced emf immediately ceases and, due to the Joule effect, also the induced current ceases immediately. The magnet would be stopped just falling over a conductor with $R = 0\Omega$ (an ideal perfect conductor!). Supposing that an electromagnetic induction process will be at the base of the superconductivity levitation, a link between the magnetic and electric property of the SC is activated. This suggest that the resistivity of the YBCO sample could be suddenly change when $T \sim T_{NL}$ and the levitation phenomena is observed. The experimental measurement of the breakdown of the resistivity of a SC, give quantitative evidence to this link making explicit also that a phase transition occurs when the YBCO sample changes from the ordinary conductor state, to the superconductor state.

The phenomenological exploration leads to the conclusion that $R = 0$ as well as $B = 0$ inside the SC, aspects that characterize the Meissner effect. The levitation phenomena, that is a phenomenological evidence of this effect, can be described as a fall down of a magnet occurring at $v = 0$ velocity. When the superconductive state is create with the magnet away from the SC, the magnetic flux changes from 0 to the value corresponding to B*S (B* is the mean value of the magnetic field produced by the magnet on the surface of the SC of area S). When the superconductive state is create with the magnet over the YBCO sample, the flux variation is due to the change in the magnetic properties of the SC. So also in this case, where apparently there isn't a flux variation, an electromagnetic induction process is at the base of the levitation phenomenon.

To give into account how this superconductive state can be created, it is possible to start from the analysis of the energy of the electrons system inside of a crystal lattice. In the ordinary conduction state the electrons could be treated as not interacting particles, because of the negligible interaction occurring and due to two part: a repulsive part, due to the coulombian interaction essentially screened by the mean field produced by the lattice; an attractive part, due to an effective potential emerging as results of the interaction of the electron and the lattice. Only the second part is depending from the temperature, so when it is prevalent with respect to the coulombian part, the system of electrons become instable for creation on Cooper pairs, or in other words the formation of pair of correlated electrons is a process energetically favored. The collapse of these Cooper pairs on the same ground state is responsible of the $R = 0$ property of an SC, that is recognized to be at the base of the superconduction behavior.

The discussion of the creation of Cooper pairs ends the exploration of the Meissner effect that is the main focus of the present work. The educational path faces also the other effect evidenced usually by the II type of SC, as the persistent currents, the pinning effect and the correlated phenomenology. This effects are observed when the superconductivity state is created in presence of an external field. The pinning effect, due to the penetration of the magnetic field inside the SC sample inside the vortexes created by supercurrents, emerges as the anchoring of the magnet to the SC. This is at the base of the MAGLEV train.

46.4 From Electron Conduction to the Superconduction of the Cooper Pairs

The Meissner effect is described characterizing the SC state, on the phenomenological level, with the two conditions $B_{int} = 0$ and $R_{SC} = 0$, but do not giving into account how the phase transition can occurs and the processes that produce the change from the initial situation, to the SC state. This change can be described in an educational perspective as follow.

The energy of the electrons inside a solid, that can be written roughly as the following sum:

$$E_{el} = K_e + V_{int} = \Sigma_{(i)} K_i + 1/2[\Sigma_{(i \neq j)} V_{coul} - \Sigma_{(i)} V_{2ij}]$$

where:

- $K_e = \Sigma_{(i)} K_i = \Sigma_{(i)} p_i^2/2m^*$ is the kinetic energy of the electrons, p_i being the momentum of the i-th electron and m* is the effective mass of the electron inside the lattice (including in a phenomenological way the interaction pf the main field produced by the lattice)
- $V_{coul} = 1/2\Sigma_{(i \neq j)} V_{coul}(r_{ij})$ is the screened Coulomb energy (of the form: $V_{coul}(r_{ij}) = V_c \exp(-\beta r_{ij})(e^2/r_{ij})$
- $V_2 = 1/2\Sigma_{(ij)} V_{2ij}$ is an effective energy potential due to the interaction of electron and lattice (in the case of Sc of I type this can be written in a relative simple way as a phonon-electron interaction).

In an ordinary conductor the electrons can be treated as a system of non-interacting particles, or weakly repulsing particles due to the predominant term V_{coul} in the interaction part of energy V_{int}. When the temperature increases V_{coul} remains essentially the same,V_2 generally change. In the ordinary conductor V_{int} remains greater than zero. Their state is characterized by the values of: energy, momentum, spin. (es. E, $\mathbf{p}_1$, ↑). Many energy level are possible and are one very close in energy to the other. Of course many electrons can possess the same energy but have different values of the momentum. For each momentum two state are possible (one with spin ↑ and one with spin ↓). The electrons, being spin 1/2 h particles, due to the Pauli principle, (at $T = 0$ K) fill all the energy level till a maximum value named Fermi energy level. When T increases some electrons can usually acquire energy making a transition to excited state and leaving lower state partially empty. The presence of a band of energy level one very close to the other ensure that the system of electron can exchange energy with the lattice. This exchange is at the base of the presence of a resistivity greater than zero in the ordinary conductors.

In the SC this schema changes, because the attractive contribution V_2 can become greater than $V_{coul}$and the interaction energy V_{int} become negative ($V_{int} < 0$) and the electron are subjected to a net attractive energy. For anyhow smallI $V_{int} < 0$ the electrons system become instable for production of couple of electrons correlated by the long range interaction V_2. When one of this pair of electrons (Cooper pair) is created the energy of the system decrease of an amount of the order of the main value of $< V_2 >$. The pairs of electrons are (quasi) systems with integer spin. For energetic reasons (but not only) the coupling of electrons with opposite spin and opposite momentum is favored. Due to their integer spin, the electron pairs occupy all the same ground state, that is separated from the first excited state by an energy gap of an amount proportional to $N < V_2 >$, where N is the number of pairs created in the systems. The same process that produces the electron pairs, produces also the energy gap. The state created is not only relatively stable, but also ensures that the pairs of electrons cannot exchange energy with the crystal lattice. In other words the energy gap ensure that the resistivity of the system becomes zero. The creation of pairs of electrons is a crucial process in the creation of the superconductive state.

In the case of SC of I type the schema described is framed in the BCS theory of superconductivity [ ], being the V2 term of interaction expressed as a potential energy interaction between two electrons mediated by the exchange of phonons. In the case of the SC of II type it is not clear what kind of interaction is responsible of the V2 energy potential.

46.5 Experimentation with Students

Explorative activities with students are carried out in informal learning setting, in four contexts (Udine, Pordenone, Frascati) with 685 students. These activities was important to test microsteps of the path, in particular for what concern the way in which students discuss the distinction between the levitation due to the Meissner effect and the levitation due to the pinning effect. The research experimentations of the educational path on Meissenr effect of SC was performed with students in eight contexts in Italy (Udine, Pordenone, Cosenza, Bari, Crotone, Siena) with 335 students of Italian upper secondary schools (17–19 aged). In these experimentation tutorial worksheets are used to monitor the students learning, analyzing their answers, drawings and schemas after the construction of typical categories. Few results concerning three main conceptual steps are given, referring to further work a more thorough details.

The change in the properties of the YBCO disc when $T \sim T_{NL}$. When student answered to the question concerning the change of the properties of a YBCO sample when the temperature is $T = T_{NL}$, sentencing: (a) “the properties of the YBCO have changed”, “the diminution of the temperature produced a change in the behavior of the YBCO” (40 %); (b) “The disc of YBCO changes his properties. It repels the magnetic field" (35 %); "The YBCO disc is magnetized" (13 %); Re-arrangement of the atoms (7 %); other answers (5 %). The change in the property of the superconductor is the main focus of about 90 % of the students answers, explicitly referred to the magnetic properties in about half of the sample. When the students were requested to represents the magnetic field configuration that can give account of the stability of the levitation configuration, they drawn: field lines do not penetrating the YBCO if $T < T_{NL}$ (55 %); field lines around the SC and aroud the magnet (37 %); the lines of the magnet penetrate inside the SC at $T > T_{NL}$ (8 %). The first two categories include all the draw where the field do not entry in the SC.

For what concern the intermediate way to resume the Meissner effect indicate that: "It consists in canceling the magnetic field is part of a SC", "At a certain temperature of the transition to the SC the Meissner effect means that the SC ceases to oppose any resistance to the passage of electricity and being all or almost all internal magnetic fields" (22 %); The Meissner effect makes levitating a magnet over a superconductor without any constraint (57 %), "an effect which changes the physical properties of objects subjected to cooling" (14%); "magnetic fields are arranged in such a manner (such as jammed) that the magnetysme in air because it has no possibility to deviate) then receives only fields opposite to the magnet." (7 %). In all the students answers the Meissenr effect is identified phenomenologically. It Is also explicitly framed on the conceptual plane by the students of the first group.

46.6 Conclusion

An educational path on superconductivity for Upper Secondary School was designed. Students are engaged in an experimental exploration of the phenomenology of use the experiments developed in the context of the European projects MOSEM and MOSEM 2. The focus is on the analysis of the Meissner effect to characterize the perfect diamagnetism of the superconductors under the temperature of the phase transition. With an experimental exploration, carried out both with qualitative observations, both with measurements using on-line sensors, students gradually construct the conclusions that $B = 0$ and $r = 0$ inside a S under the critical temperature. The students use concepts of the electromagnetism (the field lines, the magnetization vector, the electromagnetic induction law) as tools to construct a link between magnetic and electric properties of a syperconductor, describing the phenomenology of the Meissner effect, according to the suggestion of many authors, that show the possibility to describe the Meissner effect in the framework of classical electromagnetism. In the phenomenological description of the superconduction the aim is the recognition of the role of the electromagnetic induction in the description of the Meissner effect. How this state is produced or the phase transition occurs, it is described for what concern the role of the creation of the Cooper pairs on the change in the energy of the electrons system of the superconductors. The focus on energy leave open the opportunity to go deep in the quantum description of the energy gap formation.

From research experimentations carried out in eight different contexts with 335 students emerge that the majority of students recognize the change in the magnetic properties of the superconductor under a critical temperature, the B = 0 condition, the different nature of the magnetic suspension and the levitation of a magnet on a superconductor sample.

References

1. AJP (2002) Spec Issues Am J Phys 70(3)
2. Johanssonm KE, Milstead D (2008) Phys Educ 43:173–179
3. Meijer F (2005) In: Fischer HE (ed) Developing standards in RSE. Taylor and Francis, London, pp 147–153
4. PE (2000) Spec Issues Phys Educ 35(6)
5. Steinberg RN, Oberem GE (2000) JCMST 19(2):115–136
6. PISA (2006), PISA results. http://www.pisa.oecd.org
7. Duit R (2006) Science Education Research—An Indispensable Prerequisite for Improving Instructional Practice, ESERA Summer School. http://www.naturfagsenteret.no/esera/summerschool2006.html
8. Euler M (2004) In: Redish EF, Vicentini M (eds), Proceedings of the international school of physics 'Enrico Fermi', varenna, course 'research on physics education' - July 2003, Italy, Amsterdam: IOS, pp 175–221; Euler (2004) Quality development: challenges to Physics Education, In: Michelini M (ed) Quality development in the teacher education and training. Girep, Forum, Udine
9. Battaglia RO et al (2011) In Rogers L et al (eds) Community and cooperation, vol II. Leicester University, Leicester, pp 97–136
10. Zollman DA et al (2002) AJP 70(3):252–259
11. Cobai D, Michelini M, Pugliese S (eds) (1996) Teaching the science of condensed matter and new materials. Forum, GIREP Book of selected papers, Udine
12. Fischler H, Lichtfeldt M (1992) IJSE 14(2):181–190
13. Zollmann D (eds) (1999) Research on teaching and learning quantum mechanics. http://www.phys.ksu.edu/perg/papers/narst/
14. Michelini M, Viola R (2011) Il Nuovo Cimento. doi:10.1393/ncc/i2011-10997-3
15. Taşar MF (2009) The international history, philosophy, and ST group Biennial meeting. University of Notre Dame, South Bend
16. Kedzierska E et al (2010) Il Nuovo Cimento 33(3):65–74
17. Gervasio M, Michelini M (2010), http://www.fisica.uniud.it/URDF/mptl14/contents.htm
18. Ginzburg VL, Landau LD (1950) On the theory of superconductivity. Zh Eksp Teor Fiz 20:1064–1082. English translation in Collected Papers by L. D. Landau edited by D. Ter Haar, (Pergamon, Oxford, 1965), pp 546–568
19. Bardeen J, Cooper LN, Schrieffer JR (1957) Theory of superconductivity. Phys Rev 108:1175–1204
20. Josephson BD (1964) Coupled superconductors. Rev Mod Phys 36:216–220
21. Essén H, Fiolhais N (2012) AJP 80(2):164–169
22. de Gennes PG (1999) Superconductivity of metals and alloys. Perseus Books, Reading, pp 4–7
23. Fiolhais MCN, Essén H, Providencia C, Nordmark AB (2011) Magnetic field and current are zero inside ideal conductors. Prog Electromagn Res B 27:187–212
24. Badìa-Majòs A (2006) Understanding stable levitation of superconductors from intermediate electromagnetics. Am J Phys 74:1136–1142
25. Farrell WE (1981) Classical derivation of the London equations. Phys Rev Lett 47:1863–1866

Chapter 47
An Experimental Approach of Nodes Towards the Electric Potential for Students and Teachers

Alessandra Mossenta and Marisa Michelini

Abstract Electricity is a common learning topic. Early research on students' difficulties on circuits pointed out that the concept of potential does not have role nor meaning and referred it to the link between electrostatics and circuits. Subsequent research highlighted how potential is not recognized as a significant quantity for charge transfer or in activities to bridge electrostatics and circuits, micro and macro view. Starting from these outcomes a teaching/learning proposal was designed on charge transfer with simple hands-on experiments and sensors. It aims at realizing the students' cognitive need for an interpretative quantity (potential) driving the charge transfer, with a meaning of energy.

$$S_{i,j}^{m,t} = (1-\alpha)S_{i,j}^{m,t-1} + \alpha S_{i,j}^{n,t,u,v}$$

47.1 Introduction

Static electricity forms part of our daily experience, concerning both natural phenomena (e.g. discharging by sparks) and technological devices (e.g. functioning of photocopiers).

The role of education, to provide students of tools to organize phenomenology and to build models for its interpretation and for a sound approach to everyday challenges, implies the need to take into account also this topic. From the educational point of view, electrostatics has a central role in electromagnetism as a fundamental context where charge, field and potential concepts are introduced and principles as the conservation of charge and the superposition principle organize the interpretation of phenomenology. On a methodological plane, electrostatics could provide a fertile

A. Mossenta (✉) · M. Michelini
University of Udine, Udine, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_47,

context where the process of building scientific knowledge as an activity of model development, starting from and based on phenomenology, could be exemplified.

A broad research on learning difficulties (Pfundt 2009), pointed out that for the students' perspective, the concept of potential (along with the related concepts of potential difference and fem) does not have a role in the operation of electric circuits nor a separate meaning justifying its use as an interpretative quantity [2, 3]. According to the first studies about the reasoning on circuits [2–4], students identify and/or not relate adequately potential difference with current and potential, analysing circuits with sequential and local reasoning. Potential is not associated with an electric field and/or forces. Students do not relate current, in its macroscopic manifestation, with the microscopic phenomenon of charge motion. A suggested solution was to connect electrostatics and electrodynamics, via microscopic models. Other researches relate the learning difficulties to a reuse of the previously acquired electrostatics knowledge in an unchanged form [5], potential difference corresponding to a difference in the local density of mobile charges at the circuits points considered. Subsequent research concerned both proposals of teaching/learning sequences designed to build links between microscopic view in electrostatics and macroscopic view of the electric circuits [6–8], and studies of the students' reasoning in interpreting electrostatics phenomena [9, 10]. As concerns the high school and university students' description in explaining electrostatics phenomena, four models of electricity were identified [9]. According to the most popular model, electricity is a fluid that passes from one body to another: through rubbing it passes only on to dielectrics; through contact, it passes only on to conductors. Researches on students' understanding of the transfer of charge between charged conductors [10] pointed out that students admit a transfer only between oppositely charged conductors until one of the conductors became neutral (according to the rule 'like charges repel'). Primary-school student teachers explained the charge transfer between two identical metal spheres with the different quantities of electrons on them, freely interchanged with the difference in potentials [6], and a process of electrons moving from the sphere with the greater number to that with the lesser, until both spheres obtain the same charge. Researches on university students' reasoning linked to the use of electrical potential and capacitance in the processes of electrical charging of a body [11–13] confirmed the previous model and pointed out that students attribute to the potential the meaning of "indicator of quantity of charge" which the body can contain, establishing an identification between charge and potential. The capacitance is identified with the charge, as if the body were a recipient that admits a certain quantity of charge depending on its size, or it is understood as the facility to drive the charge. When reasoning include Coulomb's force, a certain identification between electric force and electric potential is produced.

Some proposals focused in the domain of electrostatics for high school or university students were developed to overcome the learning difficulties, in the perspective of linking electrostatics and electrodynamics, a microscopic with a macroscopic view. Borghi et al. proposed a teaching sequence based on the use of microscopic models to link electrostatic phenomena with direct currents. Starting from experiments on charging objects by rubbing and by induction closely examines the Volta's ideas of tension and capacitance to make the students acknowledge that a charging process is due to the motion of electrons representing a current. Tension is considered as the particular condition of the charge configuration producing divergences of the foils of an electroscope and capacitance as a magnitude expressing the ability of the system to receive more charge. The sequence developed by Guisasola et al. [14], takes into account the historical concept of capacitance in phenomena of body charging, for a transition from electrostatics to electrokinetics, helping students establish connections between the movement of charges (microframe) and an energy analysis of the system (macroframe). As a result of the implementation of the proposal, authors found that many students seemed not to relate the processes of charging to other electrostatic concepts such as potential or field. Moreover, students did not develop model-building skills to the hoped level. The authors connect this problem to the previous research, suggesting that scientific reasoning skills are difficult to develop without a curriculum specifically aimed at giving students practice in engaging in scientific activity.

47.2 The Vertical Proposal on Electrostatics

Our activity focused on a vertical sequence, offering possibilities to develop scientific content and methodological skills to the pupils since their first explorations of phenomenology, to build scientific concepts and methods when their interpretative models are build first. This could provide pupils with scientific ideas avoiding the need of a change of their spontaneous ideas when they are deeply inserted in a spontaneous way of reasoning, during high school or university.

We developed a learning proposal starting from learning and conceptual knots emerged from literature, in the theoretic framework of the Model of Educational Reconstruction [15]. A first part of the proposal is focused on the macroscopic properties of the electric interactions, to build the first level of interpretation of the electrostatic phenomena yet at Primary School level. In this phase we aim to produce the need to introduce an interpretative entity, the charge, to give sense to phenomena observed during electrical interactions emerged in charging processes of several materials; the properties of dual nature and mobility also emerge from phenomena interpretation. Following this first part, a quantitative activity was proposed, focused on interpretation of the movement of charge and its transfer by rubbing and by contact. The measure of charge involved in the process aims to introduce the concept of potential, to give it a role of driving the charge transfer and a meaning of work to do it; the measurements aim to give evidence also of charge conservation. The proposal is organized as a sequence of explorative hands-on activities with explorative stimulus-worksheets [16, 17] to drive the development of a coherent building of macroscopic models then substituted by microscopic models; when implemented, a strategy of Conceptual Microsteps based up SPEA cycles (Situation, Prevision, Experiment, Analysis) is proposed, to give to the students methodological tools in training the scientific methods aspects. The sequence is designed also to be proposed to prospective Primary School teachers, during their education to develop their Pedagogical Content Knowledge [18]. We report the first outcomes of an implementation of the sequence with middle school and high school students.

47.3 The path

The first qualitative part of the proposal can briefly be summarized as follows:

Pulling and rubbing: Observed macroscopic interactions between couples of objects pulled off the same/different surfaces or rubbed with the same/different material aims to build the concept of electric state of objects linked to a macroscopic property responsible for the interactions. The two kinds of interaction are attributed to the relation between two kinds of property, each of which can be assumed by each object.

The couples: The attraction of two objects modified one with the other by pulling or rubbing gives a first appearing of the idea that charging is an effect of an interaction, in which both objects of the couple involved in pulling or rubbing are modified, resulting in different states.

An electric state detector: A simple detector made of strips of adhesive tape, pulled apart after being stuck together, allows to increase the kind of objects that can be electrified by rubbing (insulators and metals), gives evidence that the effects of pulling and rubbing are referred to the same origin (differently from magnetic effects) and brings out the conclusion that there are only two kinds of electric state. Observing many interactions leads naturally students to assume the dependence of the interaction from distance and level of modified state.

A can: Repulsion of aluminium strips from the bottom of a metal can after a contact with a charged object introduce charging by contact and movement of charge; when a charged object is only brought near to the can the observed effect is linked with the presence of (microscopic) entities into the neutral can; the process of induction offers an interpretation of the experience and is the starting point to explain the behaviour of a plastic bottle and of a foil electroscope. The activities involve insulators and conductors, to see the similarities in the process of charging, for a highlight of the global nature of electricity, and the different mobility in them, as responsible for the difference in time where electrical phenomena are involved.

Faraday cage: The change in distribution of charge on an flat aluminum sheet after its shaping as a cylinder and the induction produced by a charged bar inserted on a metal can leads students to acknowledge the distribution of charges on the outer surface of a conductor and the charge measurements made by a charge sensor (Pasco) and a Faraday cage.

The second part of our proposal aims to introduce potential as a new property needed to manage the transfer of charge, emerging from explorations designed to face the learning knots pointed out by the previous analysis of the subject content and of the learning difficulties. During the exploration of simple transfers of charge some features of the potential emerged, as its depending on variables specifically studied.

Moreover, the exploration leads students to observe the conservation of charge as a constraint of the process. A further step aims to give a meaning of potential as work needed for the transfer of charge. Then potential is linked to the electric field, that is explored.

The first exploration concerns the transfer of charge in rubbing: the objectives of measurements of charge on two objects rubbed together are to introduce students to the way of measurement proposed, and to develop the model of charge transferred from one object to the other while conserving it as explanation of the same quantity of charge of different kind, looking at the global system of both objects involved in rubbing.

The next step concerns the transfer of charge by contact. The measurements explore first the contact between a rubbed (charged) insulator and a neutral metallic sphere, showing a transfer by contact between insulator and metal; then, successive contacts of two metallic spheres, identical and not, in different starting situations are observed: (a) only one sphere charged, to study the transfer on a neutral metal, in equal or different quantities; (b) both spheres charged, to study the transfer between charged objects. The proposed sequence aims to create in the students a conflict situation when appears a transfer of charge from a little sphere to a big one, increasing the difference in charge on the spheres instead of decreasing it, as the usual model emerged from literature should predict. The measurements are easily explained if a new quantity is introduced that manages the transfer, “potential”, depending from the charge on the spheres and from their size. The proposed explanation is summarized in Fig. 47.1. It is organized in steps as follows:

Step 1: After a contact of two equal spheres, one of which charged, half charges transfer: assuming a state of the charged sphere as referent for the condition of the charge configuration producing a transfer, after the contact both spheres are in the same state; it can be labeled by the charge on them.

Steps 2 and 3: A transfer is carried out in two pairs of spheres, each of them composed of one of the previously charged spheres (having the same charge and size)

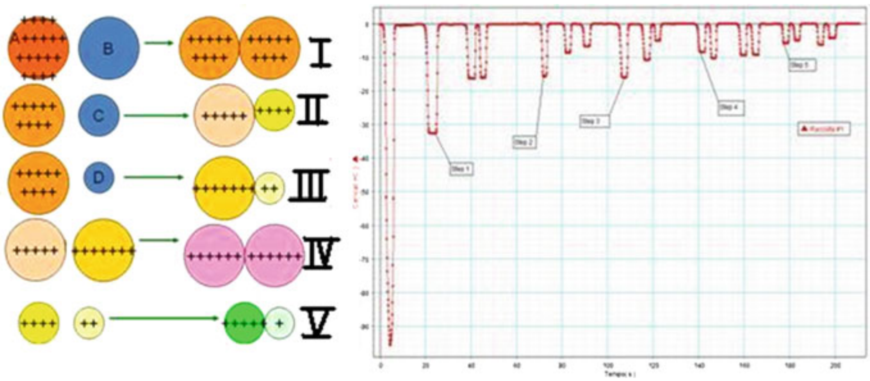


Fig. 47.1 Example of a measurement during the steps (*left*) for introducing the need of potential

and one smaller neutral sphere (of different size): the different amounts of charges on the smaller spheres are related to their different size, introducing a dependence of the state on the size of the objects; the amounts of charges on the two equal spheres are related to their state, as their size is the same. A hypothesis can be done that the greater amount of charge on them is related to a greater capability to transfer. The final state of each smaller sphere is the same of the big sphere it touched, so a comparison can be made also between the states of the two smaller spheres, showing in the smallest sphere a better capability for transfer than in the other little sphere.

Step 4: After a contact of the equal spheres with different amounts of charge, the amount of charges become the same on each sphere: the previous hypothesis is checked and the same final condition referred to the charge can be assumed also if both spheres are charged.

Step 5: After the contact of the two different spheres with different amounts of charge, ΔQ increases, according to the hypothesis that the sphere with a smaller amount of charge and a smaller size is in a state more favorable for the transfer than the other. The quantity related to the state of the spheres is named "potential" and it depends on both size and charge of the spheres. Each steps shows that in the transfer charge is conserved.

A meaning for the new quantity is provided with an exam of the charge transfer between two spheres linked by a paper strip, showing that charge is transferred also by an insulator, along a path, and then doing the same in different ways: by a person making a work on the charges, bringing them from one sphere to another by a little conductive disk attached to an insulated handle, or using a battery. This work is stored with the charge transferred and gives a measure of the capability of transfer of the charge. Bringing charge on a conductor (the sphere) requires increasing energy as the charge on it increases, and produces an increased capability of transfer in this charge. Measuring the charges on the involved objects (spheres and disk) during each step of the transfer, students can see that the disk does not discharge completely, and that the charge transferred diminishes with the increasing of the amount of charge (and its potential) on the sphere to charge. The charge transferred (and therefore the efficiency of the charging process) depends on the potential difference, as can be seen looking at the transfer between charged sphere and disk, according to the capability for the charge transfer, and the transfer stops when the potentials on the different conductors are the same.

A generator or battery is then introduced as a device able to charge a sphere at a fixed potential, 1, 2 or 3 kV with Pasco equipment; after charge measurements on the 3 different spheres used before, students observe that charges on each sphere are in the same ratio of its potential, establishing a relation between charge and potential, where capacitance, C, is introduced as the constant between the charge, Q, and the potential V, $Q = CV$; looking at the measurements on different spheres the dependence of capacitance from the size of the sphere (its radius, R) can be seen, $C = kR$; combining the two relations a formal expression for the previous qualitative observations can be stated, as $V = kQ/R$.

The process of charging a capacitor is examined in detail: it is introduced in the context of the previous experiences, and starts from their conclusions. It can be a starting point for exploring the space between its plates, to build the concept of field with one explorer moved and then filling the space in two or three dimensions with several little explorers as fragments of tea leaves in oil.

After the planning of the sequence we defined a main Research Question, asking if it is possible to build an educational proposal from middle/junior high school students and for teachers, starting from the macroscopic exploration of phenomena and using its power to interpret them, addressing subject and conceptual knots in electrostatics as regards potential and field too. We suppose that only a sequence coherent in the step by step conceptual outcomes could be a basis to link electrostatic and circuits. We designed a set of Research Questions to obtain an overall view. As concerns the transfer of charge by rubbing, and the objective of the measurements carried out, we state a first Research question:

RQ1: Do students recognize a transfer of charge and its conservation in exploring electrification by rubbing?
As concerns the transfer of charge by contact our Research Question takes into account the explanations of students:

RQ2: What are the spontaneous reasoning of students in explaining the transfer of charge? Do they take into account the conservation of charge?

RQ3: What changes produces the proposed sequence into the students reasoning about the transfer of charge? Did the reasoning involve only one object or take into account the system?

47.4 Data and Discussion

A one hour activity was carried out with 2 groups of middle school pupils, M1 and M2, focused on the measurements of charges on rubbed objects, after a brief discussion about the experiences of the first part of the path. The same focus was given in a path proposed in a last year class of high school, H1. In another class of the same level, H2, the whole proposal was carried out. We analyzed the worksheets related to the experience of measuring the charge on rubbed objects, and test in/out for the whole path in H2. The test proposed situations of contact of charged spheres, as in literature [10] asking for a prevision and an explanation about the final charge distribution.

After the experience of charging by rubbing, pupils were asked for an explanation for the unlike signs of the charges on the objects (QA) and for the equal magnitude of the two charges (QB). The answers take different directions in the two groups: a majority in M1 (20/30 pupils, 67 %) relates the systems condition with the measured values, tautologically, explaining QA: “because one (object) was positive and the other negative”; more than half of the group does not answer in QB, but 12/30 pupils (40 %) explain the same magnitude with the same process on the objects; “there was the same number of charges having suffered (the objects) the same process” or “because they were rubbed each other”. The other group (M2) related the observed outcomes to the features of the systems: 26/28 pupils (93 %) for QA explain the different sign of charges with the different color of the objects: "the two disks had a different colour", and for QB explain the same amount of charge with the same shape: "are both disks, the disks are equal". In H1 we observe for QA in majority the same kind of answers of middle school pupils: 5/12 students answer as M1, 2/12 as M2 (referring to the different material of the disks); 3/12 students claim a process (microscopic) that involves charges: a separation of charge, "in rubbing, the opposite charges displaced on different disks"; a transfer, "there was a charge passage from a material to the other", a polarization. "the charges take an opposite orientation". For QB 4/12 students relate the systems condition with the measured values, tautologically ("the charges after the rubbing are equal"), 2/12 explain the same magnitude with the same process on the objects ("They were charged in the same way for the same time").

The outcomes from the test in/out of the class that carried out the whole proposal can be summarized referring to situation (a), where 2 equal spheres were charged before the contact with $+8\mu$C and $+2\mu$C. In test in, 50 % of pupils explain the final situation in terms of evolution of systems in different states towards a common state of equilibrium (process); 33 % explain the impossibility of transfer with a reasoning based on the Coulombian force. Analyzing the other situations, reasoning based on the idea of force is expressed in different situations and produces several images about the processes and the final charge distributions. Two kinds of process emerge in answers admitting a charge flow for equilibrium: a charge transfer from the more charged sphere to the other with neutralization of opposite charge, or a charge flow in both directions to reach the same configuration of charge on both spheres without a need of neutralization. After the activity, in situation (a) 100 % of students admit a transfer; forces are not cited in explanations; 66 % of explanations referred to a process of transfer: "The two spheres have the same dimension, so the transfer of charge will happen from the more charged sphere and there will be a re-distribution between the spheres leading them to equilibrium"; the others cite the final equilibrium state: "The sphere will reach the electrostatic equilibrium"; 4/19 students introduce the idea of potential as reference for the transfer.

As concerns our Research Questions, we conclude that:

- RQ1: Pupils of middle and high school recognized that both objects charged by rubbing, but did not recognize a transfer and conservation of charge. They relate the same rubbing or the same shape of the objects with the same amount of charge, the difference in material with the difference in sign: pupils show a local reasoning, looking at the system as a set of separate systems; a process of transfer did not emerged. Moreover, the different material of the discs could be a cue of the need of a general rule managing the process, avoiding a different effect of the same rubbing on different materials. This did not emerge, also if the reasoning of high school students shows a implicit assumption of the charge conservation.
- RQ2: High school students' reasoning for charge transfer grounds on the concept of force (Coulombian), or on a model of fluid. The processes involving this model

are symmetrical (a transfer in both directions, a spreading of the common charge) or not: in this process the charge transfer happens according to the amount of charge. Conservation of charge is assumed implicitly, excluding the charge configurations where charge is not conserved.

RQ3: After the instruction, a new way to look at the situations is produced, where there is a referent for the final state of equilibrium, the potential, and it is taken as a criterion for the transfer. This new perspective implies to look at the system as a whole.

The different perspectives adopted by the students in observing a simple electrostatics phenomenology suggest the need to discuss these models also with teachers, during teacher education. The proposal could be further developed, to shift the concepts in the context of electrokinetics.

References

1. Pfund H, Duit R, STCSE. http://www.ipn.uni-kiel.de/aktuell/stcse.html. Accessed nov 2011
2. Mc Dermott L C, Shaffer P S (1992) Research as a guide for curriculum development: an example from introductory electricity. Part I: investigation of student understanding. Am J Phys 60(11):994–1003
3. Ganiel U (2000) Linking electrostatics to electrodynamics, macro to micro: student understanding of electric circuits. In: Xingkai L, Kaihua Z (eds) Proceedings of the '99 International conference of physics teachers and educators, Guangxi Normal University Press, Guilin, pp 69–76
4. Eylon B, Ganiel U (1990) Macro - micro relationship: the missing link between electrostatics and electrodynamics in students' reasoning. Int J Sci Educ 12(1):79–94
5. Benseghir A, Closset JL (1996) The electrostatics - electrokinetics transition. Historical and educational difficulties. Int J Sci Educ 18(2):179–191
6. Barbas A, Psillos D (1997) Causal reasoning as a base for advancing a systemic approach to simple electrical circuits. Res Sci Educ 27(3):445–459
7. Duit R, von Rhöneck C (1997) Learning and understanding key concepts of electricity. In: Tiberghien A, Jossem EL, Barojas J (Eds) Connecting research in physics education with teacher education, I.C.P.E
8. Borghi L, De Ambrosis A, Mascheretti P (2007) Microscopic models for bridging electrostatics and currents. Phys Educ 42(2):146–155
9. Furió C, Guisasola J, Almudì JM (2004) Elementary electrostatic phenomena: historical hindrances and students' difficulties. Can J Sci Math Tech Educ 4(3):291–313
10. Guruswamy C, Somers M D, Hussey R G, (1997) Students' understanding of the transfer of charge between conductors. Phys Educ 32(2):91–96
11. Zubimendi J L, Ceberio M, (2005) Los processos de carga eléctrica de cuerpos como instrumento de evaluación en el aprendizaje de la electricidad en estudiantes universitarios. Enseñanza de las Ciencias, 2005. Número extra. VII Congreso.
12. Guisasola J , Zubimendi J L, Almudì J M, Ceberio M, (2007) Using the processes of electrical charge of bodies as a tool in the assessment of University students' learning in electricity. Contrib Sci Educ Res 5:225–236
13. Guisasola J, Zubimendi J L, Almudì J M, Ceberio M (2008) Dificultades persistentes en el aprendizaje de la electricidad: estrategias de razonamiento de los estudiantes al explicar fenómenos de carga eléctrica Enseñanza de las Ciencias 26(2):173–188

14. Guisasola J, Zubimendi JL, Zuza K (2010) How much have students learned? Research-based teaching on electrical capacitance. Phys Rev ST Phys Educ Res 6(2)
15. Duit R (2006) science education research internationally: conceptions, research methods, domains of research. Eurasia J Math Sci Technol Educ 3(1):3–15
16. Martongelli R, Michelini M, Santi L, Stefanel A (2001) Educational proposals using new technologies and telematic net for physics. In: Pinto R, Surinach S (eds) Physics teacher education beyond 2000, Girep book, Elsevier, 615
17. Michelini M, (2004) Physics in context for elementary teacher training in Quality Development in Teacher Education and Training, Girep Book of selected papers PT_F8, FORUM, Udine 389–394
18. Shulman L (1986) Those who understand: knowledge growth in teaching. Educ Researcher 15(2):4–14

Chapter 48
From Heuristics to Humble Theories in Physics Education: The Case of Modelling Personal Appropriation of Thermodynamics in Naturalistic Settings

Olivia Levrini, Giulia Tasquier, Barbara Pecori and Paola Fantini

Abstract The paper concerns the analysis of a successful classroom experience where secondary school students made evident progress in appropriating thermodynamics according to personal approaches. The main result of the analysis is the construction of a definition of appropriation which is *operational* in two senses: (i) it includes the indication of what *observable features* must be searched in students' discourses for recognizing appropriation; (ii) it is *effective* for recognizing appropriation also in cases where it is not evident. The study is the first step of an iterative process aimed at developing a "humble theory" for explaining *when*, *how* and *why* appropriation is triggered and supported in real classrooms.

48.1 Introduction

The study concerns an analysis of data collected during the implementation of a teaching proposal on thermodynamics (about 25 school-periods) in a class of 20 students (17 years-old) of a scientifically-oriented secondary school in Italy (teacher: P. Fantini).

During the activities we observed that something important happened: the students appeared to make evident progress in *appropriating* thermodynamics according to personal approaches, by taking part in the teaching/learning dynamics of the classroom. In other words, most of the students seemed to be able to find a way, that they *feel comfortable with*, for understanding and re-organizing knowledge.

In the light of this evidence we decided to carry out a study in order to answer the following research questions:

O. Levrini (✉) · G. Tasquier · B. Pecori
Department of Physics, University of Bologna, Bologna, Italy

P. Fantini
Liceo Scientifico "A. Einstein", Rimini, Italy, PhD School, Anthropology and Epistemology of Complexity, University of Bergamo, Bergamo, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_48,

RQ1 - How can the word "appropriation" be defined so as to become an operative tool for recognising and evaluating the effectiveness of a teaching/learning experience (*What "appropriation" means*)?

RQ2 - What factors (features of the contents knowledge reconstruction, collective activities, learning environment, mediation action,..) trigger and support individual processes of appropriation *(When, why and how appropriation occurs)*?

48.2 Methodological Framework

Both the selection and formulation of the research questions mirror our choice of referring to Design Studies as methodological framework [1, 2]. Of the Design Studies we considered, in particular:

– the theoretical orientation: *"Design studies are conducted to develop theories, not merely to empirically tune 'what works"'* [1].
– the "local" (or "humble") character of the theories [2, 3]: *"Unlike grand theories or orienting frameworks, they [the theories] aim at being specific enough [local] to do real work for improving instructional design"* [1];
– the iterative design.

A sentence particularly stimulating for the work is: *In order to pursue this intimate relationship between developing theory and improving instruction, the Design Studies indicates that, in collecting and analyzing data, "what works" must be underpinned by a concern for "how, when and why it works, and by a detailed specification of what, exactly, 'it' is"* [1].

According to the Design Study, we organized the research as follow.

In order to answer RQ1 ("*What appropriation is*"*)*, we:

(a) overviewed the whole corpus of data in order to select a manageable and significant set of data for providing the first operational definition of appropriation;
(b) tested the definition against more complex data [*by enlarging the empirical base up to "theoretical saturation",* [6]] in order both to refine it and to evaluate its operative power in making invisible visible.

As far as the RQ2 (*"How, when and why appropriation occurs"*) is concerned, we are now working for abstracting, from the specific context, the first hypotheses so as to point out new foci of attention for designing a second level of the experiment within an iterative design.

The last point is in progress and the paper will focus mainly on the work we did for answering RQ1.[1]

[1] A brief description of the work we did for answering RQ2 is reported in Levrini et al. [5].

48.3 The Educational Reconstruction of Thermodynamics

The teaching proposal was designed so as to shape the learning environment as *a properly complex territory* [4]. At the basis of the notion of learning environment as properly complex territory there is the idea that some forms of complexity in physics contents can be transformed into *productive complexities*, i.e. complexities that allow learning environments to be, besides intelligible, rich enough to enable each student to reach deep understanding and to exploit his/her cognitive potential.

The forms of productive complexity, chosen and implemented in the thermodynamics path, are:

(1) *Multi-perspectiveness:* the same contents (phenomenologies) have been analyzed from different approaches - the macroscopic and microscopic ones - treated consistently as two different models;
(2) *Multi-dimensionality:* a critical-philosophical reflection has been developed on the peculiarities of the two approaches through specific activities (questionnaire, discussions);
(3) *Longitudinality:* according to our idea that learning physics is a continuous process of widening, refining, revising knowledge already acquired, the thermodynamics ways of modeling systems, processes, interaction have been systematically compared with the models used in the theories previously studied (classical mechanics and special relativity).

48.4 Data Analysis and Results

48.4.1 Data Sources

During the activities we collected data from different sources: written tasks (quantitative and qualitative problems on key concepts, including problems known as puzzling from the research in physics education); audio-recording of lessons; students' notebooks; notes of a Master student (GT, an Author) observing the classroom activities; questionnaires stimulating critical-epistemological reflection; video and audio recording of classroom discussions about the issues raised by the questionnaires; audio-recorded individual interviews to 10 students conducted at the end of the whole work; audio-recorded interview to the teacher and discussions with her. Since we could not complete 2 interviews because of time constraints, only 8 out of 10 have been taken into account for the analysis.

48.4.2 Working out an Operational Definition of Appropriation

In order to work out an operational definition of appropriation, we focused the analysis on the data related to 5 students: their semi-structured individual interviews; their

Table 48.1 Central ideas around which students' discourse runs

Michele	The Curiosity for "how things work"	"I like Physics because it explains how reality works, so to say, I'm very curious about how objects work and natural events"
Matteo	The pleasure of speculating, in particular on the philosophical distinction between "being and becoming"	"I believe that […] it is fundamental to build a basis and to speculate on how theories are found, how concepts are elaborated. These concepts will certainly last longer than formulas"
Chiara	The pleasure of understanding by exploring and testing different points of view or perspectives	"It may be more complete to try to analyze a phenomenon, or whatever is around us, from two different points of view rather than from one […]. It widened my view"
Lorenzo	The need of searching for a unified consistent framework for deep understanding	"One can see that everything is integrated, it is not divided in topics each with its own laws, instead everything can be connected, unified; the argument becomes wider, more uniform"
Caterina	The fascinating search for "not obvious details that are usually taken for granted"	[I like physics since] "obvious things are not taken for granted".[Of thermodynamics I like] "that it made me notice something I had not noticed before"

answers to questionnaires on a critical-epistemological reflection; their contributions to classroom discussions.

The 5 students were chosen on the basis of three criteria: our impression they could represent interesting cases for studying appropriation; the richness of their interviews; their being representative of the class (3 boys, 2 girls, with different levels of performance in physics and different roles within the classroom dynamics).

The data concerning each student were selected and re-arranged so as to draw "the student's profile". In order to make the 5 profiles comparable, all of them were constructed following the same steps: (i) to pick up the *central idea*, if there is one, around which the student, during the interview, develops her/his discourse; (ii) to see if and how such an idea is developed along the interview; (iii) to reconsider the classroom discussions and the answers to the questionnaire in order to check if there is a relationship between the student's position during the interviews and what appeared in the previous activities.

The profiles' comparability allowed us to get the definition of appropriation out of observable features of students' discourses.

The first main evidence which comes out from the construction and the comparison of the profiles is the variety of the central ideas around which students developed their discourse (see Table 48.1).[2]

[2] The names of the students are invented.

The other very important evidence is the students' personal way of talking about temperature, since it is evident they focused their attention on those features of temperature more meaningful to them.

Michele, for example, according to his general perspective, focused his attention on the temperature gradient because this is what makes engines work: "*Different temperatures are necessary…, only with different bodies with different temperatures we can have a cycle and work".*

Matteo, instead, focused his attention on the distinction between ΔT and T, because he saw, in this distinction, the philosophical difference between "*being* and *becoming*": *[I think that] Q = mc ΔT is 'becoming' […]… there is a change. […] the first relation [PV = nRT] is 'being' because […] [there is] absolute temperature T, that doesn't change."*

Thanks to the construction and the comparison of the profiles, the following *features* of students' discourses can be recognised as signals they did appropriate thermodynamics:

(a) the discourse is developed, around one key-idea, *consistently* along the whole interview;
(b) the key-idea allows the formulation of arguments *on-task*, i.e. arguments which allow aspects of thermodynamics to be selected and meaningfully linked one to the other;
(c) the keywords, around which each student constructs her/his arguments, reveal personal engagement, by being *genuine* and *emotion-laden*;
(d) the argumentation is *thick*, i.e. it includes elements belonging to a meta-cognitive and/or epistemological dimension;
(e) the specific approach to disciplinary knowledge is *not occasional* (it can be traced back to the student reactions in different classroom activities).

The acknowledgment of these features allowed us to formulate an operational definition of appropriation: *A personal process of content knowledge transformation that leads disciplinary knowledge itself* (feature b) *to be a consistent* (feature a) *and personal reconstruction of physics "signed" by the students*, where "signed" means that:

– the voice of the teacher is not present (feature c);
– physics content knowledge is assumed within a personal broad path to knowledge that goes far beyond the mere aim of learning thermodynamics (features d, e).

48.4.3 Testing the Definition Against Complex Data

For completing our answer to RQ1 we have, then, tested our definition against 3 interviews that appeared, at a first level of analysis, cases of non-appropriation, because of:

– the lack of an explicit and evident personal idea around which student's discourse runs;
– the common "flat" sound of their ways of talking about physical concepts (*"In classroom, we have done…", "we have introduced the concept of temperature by…", "I remember that we had…", or "I don't remember if…");*
– the same strategy of talking about physics by going over, chronologically, the classroom activities.

Instead, the finer analysis we could carry out by applying the definition of appropriation shows that only one interview is an evident case of non-appropriation. Another interview shows that the student resonated with the global sense of the work, but she seemed to be still inside the process of searching for her own approach to learning. The last interview reveals not only that the student appropriated thermodynamics, but also that appropriation was particularly difficult to be recognized because of a "natural", spontaneous cognitive resonance of the student with the inner physical language and its formal structure.

As example, we report the results of the analysis of the interview to Paolo, the last one, we carried out by applying the criteria we had pointed out for recognising appropriation.

In the first 30 min (out of 37) of the interview, Paolo goes over the whole path for describing where the concept of temperature has been addressed and in what formal relations it appeared.

The language used in the description sounds as an attempt to remember what has been done: *"We have given the initial definition to the temperature… ", "Then, we have done the macroscopic approach…", "In between we have seen the laws of Gay-Lussac and Boyle…"*

In his discourse, however, Paolo shows to move within the disciplinary dimension very easy and with self-confidence. In particular, he appears to be very sure both in moving along a formal dimension and in connecting formulas to phenomena or in recognizing the modelling dimension that stays behind formalism.

He also shows a marginal interest for reflections *about* physics or *about* his personal view as the few minutes devoted to the second and third parts of the interview prove (7 min).

Paolo may, hence, seem a student who deeply and efficiently understood the contents, without searching for that personal approach needed for transforming disciplinary knowledge into a "signed" knowledge.

Instead, in his discourse all the features we pointed out for recognizing appropriation can be found. In particular, as soon as we re-analysed the interview by applying our criterion of *consistency*, the repetition of a specific point became visible: he stresses several times that, for him, what matters is that physics has the power of reaching the "same result" ("the same formula") in different ways and that some routes (like the microscopic one) are based on already acquired knowledge (like mechanics). These two aspects make, in his opinion, physics knowledge reliable: *reliability of knowledge* emerges as that key-idea around which his discourse is consistently developed.

This idea can be consistently related to a specific image of physics to which he, implicitly, refers when he says that the boundaries of physical knowledge are clearly demarcated with respect to other forms of knowledge, such as philosophy:

"The questionnaire has been something interesting but, sincerely, it seems to me that it did not have many implications on thermodynamics… perhaps more on philosophy."

His *consistent* and *not occasional* position in looking at physics as reliable and well-demarked knowledge makes us to think that, along the whole interview, he is also implicitly supporting the *objectivity* of physics as another feature that makes physical knowledge reliable. From this perspective, the impersonal way of describing the path in the first part of the interview, seen at a first level of analysis as sign of lack of appropriation, can be acknowledged as the expression of a well-formed epistemological position: *physics is physics*. He was in fact re-constructing the path by selecting those pieces of knowledge he considered reliable and by cutting off what could sound subjective.

In this sense, we can also infer that Paolo's discourse is *thick*: a precise epistemological dimension can be recognised behind his pure and competent disciplinary (mono-dimensional) discourse.

To sum up, the application of the specific criteria we pointed out for recognising appropriation leads us to conclude that also Paolo's discourse can be acknowledged as a clearly *signed* content knowledge reconstruction. The relative poorness of his philosophical and meta-cognitive language, which made rather complicate to acknowledge signals of appropriation in his words, can be, in our opinion, ascribed to his "natural", spontaneous and authentic cognitive resonance with physics. Because of this natural resonance with physics, a fine and specific analysis has been needed for recognising that the impersonal language used by Paolo was, in fact, *his* personal language.

48.5 Final Remarks… Toward a Humble Theory

Following the Design Studies methodology, data collected during an experiment of thermodynamics have been analysed so as to work out a definition of appropriation. The definition we formulated is *operational* in two senses: (i) it includes the indication of what *observable features* must be searched in students' discourses for recognizing appropriation; (ii) it is *effective* for recognizing appropriation also in cases where it is not evident.

In the light of the obtained results, the next important move of our research is the construction of a humble theory able to explain *when*, *how* and *why* collective activities trigger and support personal appropriation. Such a concern is, in our opinion, a fundamental step for disseminating good practices, as well as for making the research results more and more powerful to promote innovation in schools.

In analogy with what happens in the physics of complex systems, the kind of humble theory we have, now, in mind ought to be shaped as a dynamical model comprised of: (i) boundary conditions, i.e. those context conditions which make the learning environment suitable and fruitful for triggering and supporting appropriation; (ii) a dynamical mechanism lying at the back of the process, i.e. a mechanism enacted and managed by the teacher in order to make *individual cognition, collective dynamics* and the *disciplinary knowledge* resonant with one other [5].

From an operative point of view, we are analysing data coming from 3 further experiences with other 2 teachers (one of them indeed implemented the same proposal, during the same period, in two different classes). What makes these data particularly interesting is that one experience, out of three, was unsuccessful: not only appropriation did not occur (with only few exceptions), but also the forms of complexities seemed to play a counterproductive role in fostering basic understanding. A further element of interest is that the unsuccessful experience is one of the two experiences realised by the same teacher.

This specific situation will allow us to treat the analysis of the boundary conditions somehow independently of the dynamical mechanism of mediation: being, indeed, the teaching proposal and the teacher the same in a successful and in the unsuccessful experience, we expect to find different boundary conditions responsible for the different reaction of the two classes.

References

1. Cobb P, Confrey J, diSessa A, Lehrer R, Schauble L (2003) Design experiments in educational research. Educ Res 32(1):9–13
2. diSessa AA, Cobb P (2004) Ontological innovation and the role of theory in design experiments. J Learn Sci 13(1):77–103
3. Levrini O, diSessa AA (2008) How students learn from multiple contexts and definitions: proper time as a coordination class. Phys. Rev. STPER, 4(010107):1–18
4. Levrini O, Fantini P, Gagliardi M, Tasquier G, Pecori B (2010) A longitudinal approach to appropriation of science ideas. In: A study of students' trajectories in thermodynamics. Proceeding 9th ICLS conference, Chicago (IL)
5. Levrini O, Fantini P, Gagliardi M, Tasquier G, Pecori B (2011) Toward a theoretical explanation of the interplay between the collective and the individual dynamics in physics learning. Paper presented at the ESERA 2011 conference, September 5th-9th 2011, Lyon
6. Strauss AL, Corbin J (1990) Basics of qualitative research. Grounded theory Procedures and techniques, Sage, Newbury Park, Ca

Chapter 49
Theory Versus Experiment: The Case of the Positron

Matteo Leone

Abstract The history of positron discovery is an interesting case-study of complex relationship between theory and experiment, and therefore could promote understanding of a key issue on the nature of science (NoS) within a learning environment. As it is well known we had indeed a theory, P.A.M. Dirac's theory of the anti-electron (1931), before the beginning of the experiments leading to the experimental discovery of the positive electron (Anderson 1932). Yet, this case is not merely an instance of successful corroboration of a theoretical prediction since, as it will be shown, the man who made the discovery, Anderson, actually did not know from the start what to look for.

49.1 Introduction

Over the years many researchers argued for the usefulness of history and philosophy of science (HPS) in science teaching. Among the main reasons for using HPS are its power to promote understanding the nature of science (NoS) by making it concrete and meaningful (e.g. [18, 25, 27, 30]); to provide scientific clarification of the concept to be taught [11] to overcome conceptual difficulties by drawing on the similarity between philo- and onto-genesis of knowledge (e.g. [13, 15, 26]). Despite the intensive support for using the HPS in science teaching, however, "the issue continues to be complex and controversial" [12] see also [14, 29].

This paper is essentially a case study in the history of particle physics that could likely promote understanding a NoS key issue, namely therelationship between

M. Leone (✉)
Dipartimento di Fisica, Università di Genova, Genoa, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_49,

theory and experiment.[1] The starting point will be an interesting remark originally put forward by Kuhn, of *Structure of Scientific Revolutions* fame. According to [20] there are two classes of scientific discoveries, namely those discoveries which "could not be predicted from accepted theory in advance and which therefore caught the assembled profession by surprise" (e.g. the oxygen, the electric current, x-rays, and the electron), and those that had been "predicted from theory before they were discovered, and the man who made the discoveries therefore knew from the start what to look for" (e.g. the neutrino, the radio waves, and the elements which filled empty spaces in the periodic table). As emphasized by Kuhn, however, "not all discoveries fall neatly into one of the two classes," and one notable example of such an occurrence had been the discovery of positron by Carl Anderson in 1932. The positron is therefore especially suitable to show the complexity of theory-experiment relationship within a learning environment.

49.2 Theory Versus Experiment in the Positron Discovery

The standard history of positron discovery is well known [9, 16, 17, 23, 31]. In May 1931 P.A.M. Dirac brought up the hypothesis of the anti-electron from his relativistic quantum theory of the electron. According to [10] , "an encounter between two hard γ-rays (of energy at least half a million volts) could lead to the creation simultaneously of an electron and anti-electron". In October 1931, during a Princeton lecture, the soon to be appointed Lucasian Professor of Mathematics stated that "this idea of the anti-electrons doesn't seem to be capable of experimental test at the moment; it could be settled by experiment if we could obtain two beams of high frequency radiation of a sufficiently great intensity and let them interact" (Dirac, as excerpted by [19]).

Within a completely different context, on September 1, 1932, Carl Anderson reported about a discovery obtained during a cosmic radiation research program at the Caltech Laboratory in Pasadena, under the directorship of Robert Millikan. By means of a vertical cloud chamber operating in a strong magnetic field, Anderson photographed indeed the passage through the cloud chamber volume of "a positively charged particle having a mass comparable with that of an electron" [2] eventually named "positron".

The connection between Dirac's anti-electron and Anderson's positron occurred in February 1933 at the Cavendish Laboratory in Cambridge thanks to P.M.S. Blackett and G. Occhialini. By means of their recently developed triggered cloud chamber [7], the Cavendish researchers succeeded indeed in collecting many photographs showing

[1] The U.S. National Research Council committee for K-12 education recently advocated the view that using HPS materials might promote understanding NoS. In the section "Practice 7: Engaging in Argument from Evidence" of their recent recommendation, it is indeed emphasized that the "Exploration of historical episodes in science can provide opportunities for students to identify the ideas, evidence, and arguments of professional scientists. In so doing, they should be encouraged to recognize the criteria used to judge claims for new knowledge and the formal means by which scientific ideas are evaluated today" ([30] pp 3–19).

positron tracks that they interpreted through the theoretical framework provided by Dirac [8].

In fact, (Dirac's) theory preceded (Anderson's) experiment. In this respect, during the 1933 Solvay Conference, Ernest Rutherford expressed his regret at the way the history of positron discovery occurred, since "we had a theory of the positive electron before the beginning of the experiments. [...] I would have liked it better if the theory had arrived after the experimental facts had been established" [21].

49.3 Anderson's Experiment

According to Anderson, his celebrated photograph (Fig. 49.1) showing a positron traversing a 6 mm lead plate upwards was obtained on August 2, 1932 [3]. The change of curvature below and above the plate showed that the particle went upwards and lost energy while crossing the lead shield. Since the curvature indicated that the particle had a positive charge, while the length of path and the specific ionization were electron-like, Anderson concluded that the particle behaved as a positive electron.

Anderson's original papers devoted to the positron [2, 3] do not report about the actual circumstances leading to the discovery. The first details date back to his 1936 Nobel Lecture, where Anderson explained that the plate was inserted "to determine without ambiguity [the particles'] direction of motion", due to "the lower energy and therefore the smaller radius of curvature of the particles in the magnetic field after they had traversed the plate and suffered a loss in energy" [4]. As regards the theoretical framework, many years elapsed before Anderson explained that the Dirac's theory "played no part whatsoever in the discovery of the positron" [5].

The primary historical sources, however, tell quite different a history. As for the experimental setup, in all likelihood the 6 mm lead plate had a different goal than later recollected by Anderson. Two months before his discovery, he had indeed reported that the goal of distinguishing between positive and negative particles was pursued

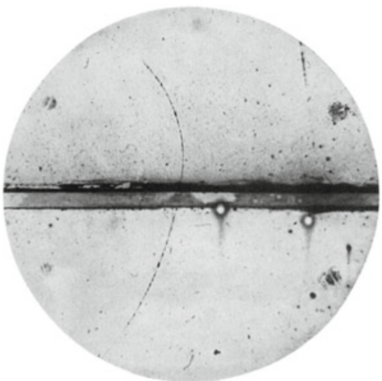


Fig. 49.1 Anderson's cloud chamber photograph of a positron traversing the 6 mm lead plate upwards, discussed in September 1932 in Science [2] and submitted in February 1933 to Physical Review [3]

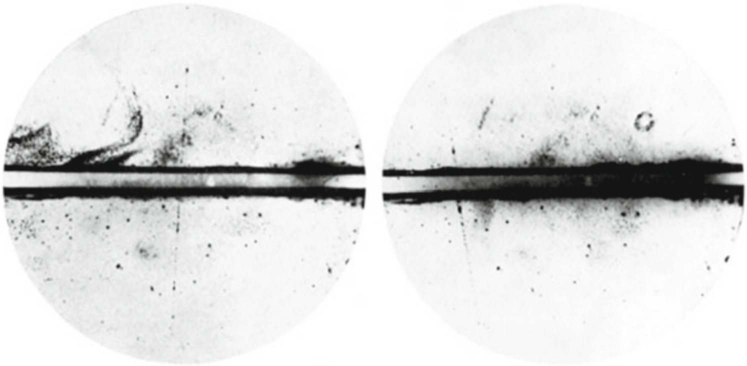


Fig. 49.2 Anderson's cloud chamber photographs, submitted in June 1932 to physical review [1], showing a particle of uncertain sign of charge that suffers a deflection of 0.5 degrees in traversing the 6 mm lead plate

by collecting "precise data on the specific ionization of the low-energy positives" rather than with a lead plate. Since at low energy protons and electrons ionize very differently, measures of specific ionization will indeed "distinguish [...] between downward positives and upward negatives" [1]. But what is most significant here is that plates of lead were used by Anderson two months before his discovery (Fig. 49.2), with the reported goal of pursuing the study of the "scattering of the cosmic particles" ([1] p. 410). For some reasons, aims and methods change as of the later recollections. What if the discovery of positron was an entirely accidental and unplanned issue?

The primary sources offer a new perspective also for what concerns the theoretical framework. While no grounds exist to support the view that Dirac's theory influenced Anderson's experiment, another theory actually played a relevant part. It is worth to recollect that Anderson worked under Millikan's directorship, and that according to Millikan's "atom-building" theory, cosmic-rays are γ-rays. Central to this theory is the idea that cosmic rays band spectrum is due to the absorption of photons emitted in the atom-building, "in the depths of space", of abundant elements like helium, oxygen, silicon and iron, out of hydrogen. The appearance of positive charges in the cloud chamber photographs, detected since late 1931, could be explained within Millikan's framework by suggesting the ejection of protons following the "disintegration of the nucleus" [28]. Thus, Anderson was clearly thinking in terms of Millikan's theory when he wrote in his first paper devoted to the discovery of the positron that one of the possible ways to explain the photographs was that a negative and positive electrons "are simultaneously *ejected* from the lead [emphasis added]" (Anderson 1932, 238) according to a process similar to that formerly suggested by Millikan to explain the alleged proton tracks. In 1934, Anderson was still moving within Millikan's theoretical framework when he wrote that "the simplest interpretation of the nature of the interaction of cosmic rays with the nuclei of atoms, lies in the assumption that when a cosmic-ray photon impinges upon a heavy nucleus, electrons of both sign are *ejected* from the nucleus [...]. [The photographs] point strongly to the existence of nuclear reactions of a type in which the nucleus plays a more active role than merely that of the catalyst [emphasis added]" [6].

49.4 Blackett and Occhialini' Synthesis

By their discovery of plenty of electron-positron pairs within the new phenomenon of cosmic-ray "showers" (Fig. 49.3), made possible by their efficient triggered cloud chamber [22], Blackett & Occhialini's "constructed a forceful case" [20] for the existence of positron. Furthermore, they grasped that Anderson's positive electron and Dirac's anti-electron were the same particle, a view supported by the fact that Dirac's theory predicted the successful detection of a positron by the cloud chamber method. As reported by Blackett and Occhialini, Dirac's calculation of probability of electron/positron annihilation process led to a positron mean life in water close to 3.6×10^{-10} s. Thus, the Cavendish researchers concluded, "in [Dirac's theory] favour is the fact that it predicts a time of life for the positive electron that is long enough for it to be observed in the cloud chamber but short enough to explain why it had not been discovered by other methods" ([8], p. 716).

Notwithstanding their support to Dirac's theory, Blackett and Occhialini's original paper provides reasons to believe that, in some respects, they departed from Dirac's pair production mechanism. In one instance, they suggest indeed that electron and positron "may be born in pairs during the *disintegration of light nuclei* [emphasis added]" since the showers had been observed to arise in air, glass, aluminum and copper" ([8], p. 713). Within another context, namely F. Joliot and I. Curie's April 1932 observation of electrons moving towards a polonium-beryllium neutron source [24], they concluded that Joliot and Curie's electrons were actually positrons arising from the action of neutrons (as opposed to γ rays). Both instances reveal a production mechanism dissimilar of the one originally put forward by Dirac.

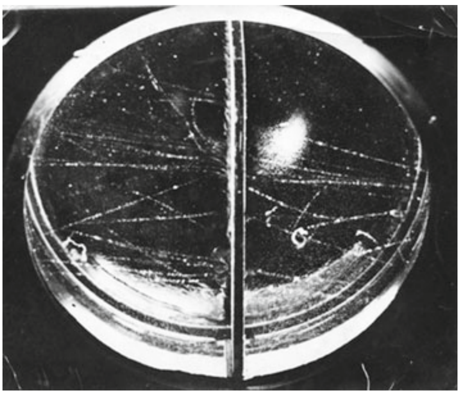


Fig. 49.3 Photographs of electron-positron showers captured by Blackett and Occhialini via their triggered cloud chamber technique [8]

49.5 Concluding Remarks

As pointed out by Kuhn, it might be safely concluded that Dirac's theory preceded positron discovery (to Rutherford's regret); that Anderson's experiment was done in complete ignorance of Dirac's theory; and that Blackett and Occhialini made use of Dirac's theory to corroborate the positron existence. However, as we have evidenced above, on the one hand a "wrong" theory (Millikan's one) guided Anderson's discovery of the positron and interpretation of the tracks. And on the other hand, a "correct" theory (Dirac's one) was not fully exploited by Blackett and Occhialini in interpreting the positron production mechanism.

This case study raises therefore a number of relevant issues about NoS. Firstly, that the theory versus experiment relationship is not always a two-poles one (often in actual science, as in the actual discovery of positron, more theories and more experiments are involved). Secondly, that sometimes "wrong" theories led to "correct" discoveries. And, finally, that original historical sources tell histories in some respects at odds with textbooks histories or with later recollections by the protagonists themselves.

References

1. Anderson CD (1932a) Energies of cosmic-ray particles. Phys Rev 41:405–421
2. Anderson CD (1932b) The apparent existence of easily detectable positives. Science 76:238
3. Anderson CD (1933) The positive electron. Phys Rev 43:491–494
4. Anderson CD (1965) 1936 The production and properties of positrons, Nobel Lectures, Physics 1922–1941. Elsevier Publishing Company, Amsterdam
5. Anderson CD (1961) Early work on the positron and muon. Am J Phys 29:825–830
6. Anderson CD, Millikan RA, Neddermeyer S, Pickering W (1934) The mechanism of cosmic-ray counter action. Phys Rev 45:352–363
7. Blackett PMS, Occhialini G (1932) Photography of penetrating corpuscular radiation. Nature 130:363
8. Blackett PMS, Occhialini G (1933) Some photographs of the tracks of penetrating radiation. Proc R Soc Lond A139:699–727
9. De Maria M, Russo A (1985) The discovery of the positron. Rivista di Storia della Scienza 2:237–286
10. Dirac PAM (1931) Quantised singularities in the electromagnetic field. Proc R Soc Lond A133:61
11. Duit R, Gropengießer H, Kattmann U (2005) Towards science education research that is relevant for improving practice: the model of educational reconstruction. In: Fischer H E (ed) Developing standards in research on science education (1–9), Taylor & Francis, London
12. Galili I (2011) Promotion of Cultural Content Knowledge through the use of the history and the philosophy of science. Sci Educ 21(9):1283-1316
13. Galili I, Hazan A (2000) The influence of a historically oriented course on students' content knowledge in optics evaluated by means of facets-schemes analysis. Am J Phys 68(7):S3–S15
14. Galili I, Hazan A (2001) Experts' views on using history and philosophy of science in the practice of physics instruction. Sci Educ 10(4):345–367
15. Halloun IA, Hestenes D (1985) Common sense concepts about motion. Am J Phys 53:1056–1065

16. Hanson NR (1961) Discovering the positron. Brit J Philos Sci 12:194–214, 299–313
17. Hanson NR (1963) The concept of the positron. A philosophical analysis. Cambridge University Press, New York
18. Kipnis N (1998) A history of science approach to the nature of science: learning science by rediscovering it. In: McComas W F (ed) The nature of science in science education (177–196), Kluwer, Dordrecht
19. Kragh H (1990) Dirac: a scientific biography. Cambridge University Press, Cambridge, p 107
20. Kuhn TS (1962) Historical structure of scientific discovery. Science 136:760–764
21. Institut International de Physique Solvay (1934) Structure et Propriétés des Noyaux Atomiques. Gauthier-Villars, Paris, pp 177–178
22. Leone M (2011) Particles that take photographs of themselves: the emergence of the triggered cloud chamber technique in early 1930s cosmic-ray physics. Am J Phys 79:454–460
23. Leone M, Robotti N (2008) P.M.S. Blackett, G. Occhialini and the invention of the counter-controlled cloud chamber (1931–32). Eur J Phys 29:177–189
24. Leone M, Robotti N (2010) Frédéric Joliot, Irène Curie and the early history of the positron (1932–33). Eur J Phys 31:975–987
25. Matthews M (2000) Time for science education: How teaching the history and philosophy of pendulum motion can contribute to science literacy. Plenum Press, New York
26. McCloskey M (1983) Intuitive physics. Sci Am 248(4):122–130
27. McComas WF (2008) Seeking historical examples to illustrate key aspects of the nature of science. Sci Educ 17(2–3):249–263
28. Millikan RA, Anderson CD (1932) Cosmic-ray energies and their bearing on the photon and neutron hypotheses. Phys Rev 40:325–328
29. Monk M, Osborne J (1997) Placing the history and philosophy of science on the curriculum: A model for the development of pedagogy. Sci Educ 81(4):405–424
30. National Research Council (2011) A Framework for K-12 Science Education: practices, cross-cutting concepts, and core ideas. Committee on a conceptual framework for new K-12 science education standards. Board on science education, division of behavioral and social sciences and education, The National Academies Press, Washington, DC
31. Roqué X (1997) The manufacture of positron. Studies in History and Philosophy of Modern Physics 28:73–129

Chapter 50
Mass from Classical to Relativistic Context: A Proposal of Conceptual Unification Experimented in the IDIFO3 Summer School

Emanuele Pugliese and Lorenzo Santi

Abstract The IDIFO3 summer school proposed a basic concept construction in modern physics to young talented students. Inertial and gravitational mass concepts were examined in Newtonian paradigm, the former according to Mach as well. Temporal component of 4-momentum was interpreted as total relativistic energy, through series expansion at low velocities; we then built the relativistic conceptual extension of mass starting from *rest energy*. We integrated an analysis on how students consider mass with an inquiry on mass meanings produced by our path. The results allowed us to recognize five student profiles.

50.1 Mass

"The concept of mass is one of the most fundamental notions in physics, comparable in importance only to the concepts of space and time" [6]. The historian of science Max Jammer puts mass on a level with the "backstage" of all physical phenomena: space-time. Mass plays such a founding role because massless macroscopic objects would be made of massless particles, without any form of interaction energy among them: we would have a Universe of free particles travelling at light speed.[1] Mass is to be admitted to endow matter with "something" that characterizes it and that determines its motion in the absence of electromagnetic fields.[2]

This quantity has a manifold character. In Newtonian physics, it is essentially the metaphysical "quantity of matter", deeply connected with inertia; the conception is explained in the opening paragraph of *Principia*:

[1] We take into account Special Relativity here.
[2] Gamow considers indeed gravity as the force that rules Universe.

E. Pugliese (✉) · L. Santi
Chemistry and Environment Department – Division of Physics, University of Udine — Physics, Udine, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_50,

> The quantity of matter is the measurement of the latter obtained from the product of density and volume. […] Air of double density, in a double space, is quadruple in quantity […] And the *norma* [*Latin text*] of all bodies, which be differently condensed for any causes, is identical […]. I will mean hereafter, everywhere, this quantity under the name of body or mass.

Inertial mass acquires its gravitational meaning with the development of the universal gravitation law. In 1883, Mach [9] criticized Newton's conception of mass, because it brought to a vicious circle, and proposed an operative definition through the measure of accelerations. He used a meaningful result derived from Newton's second and third laws, where m_1 can be put as unit of mass:

$$\frac{m_2}{m_1} = \frac{a_1}{a_2}. \tag{50.1}$$

Eventually, in Special Relativity (SR) a variation of mass corresponds to a variation in the internal energy of a body[3]: the idea of mass becomes more and more subdivided.

50.2 Learning Problems

The concept of mass is very difficult to understand because of this multifaceted character and of the persistence of the ontological vision above as well. A further facet: in the theory of General Relativity, the momentum-energy density modifies the geometry of space-time, so matter exerts an active action on the last, differently from what happened with Newtonian space.

In high-school textbooks, the following learning problems have been found: identification of weight with gravitational mass; belief that a beam balance measures weight; inertial mass defined operationally through *F/a*, where, as a literature example, F is the measurable force impressed by a compressed spring. On the subject matter ground, the conception of *quantitas materiae* has also generated misconceptions concerning the mass-energy relation: mass is converted into energy, in a generic sense (6 high-school textbooks); $E = mc^2$ represents conversion of mass into energy (1 high-school textbook); confusion/mistakes between energy conservation law and mass conservation law (3 textbooks) [8]. In matter of students' learning difficulties, an experiment on 16 to 18-aged Spanish pupils was performed. They tend to prefer a teleological-qualitative (pre-theoretical) view of mass to an operative-quantitative-formal (scientific) vision. For instance, most of them identified it with other quantities: volume or density on one side and weight on the other; in addition, *quantitas materiae* prevailed on inertia [2].

[3] This quantity is usually called *rest energy* E_0*:* the energy of a body measured in its rest reference frame.

50.3 Relativistic Mass Debate

In several treatises on relativity a quantity called "relativistic mass" appeared, whose expression is

$$m_r = \frac{m_0}{\sqrt{1 - \frac{u^2}{c^2}}}. \tag{50.2}$$

Using the latter, one can rewrite relativistic momentum in the form

$$\mathbf{p} = m_r \mathbf{u} \tag{50.3}$$

Is it enough to justify the use of (50.2) as the expression for a proper physical quantity?

Taylor and Wheeler underline that relativistic mass is the first component of relativistic quadrimomentum, while the invariant mass is the magnitude of the latter: a scalar. Actually, why should we interpret *m* in the following master equation as the inertial classical mass?

$$E^2 - p^2c^2 = m^2c^4. \tag{50.4}$$

In general, there are good reasons both to consider mass in continuity with Newtonian mechanics and to interpret it as a completely new quantity[4] in a new "paradigm" [7]: ≪the expression $\frac{m_0}{\sqrt{1-\frac{u^2}{c^2}}}$ is best suited for THE mass of a moving body≫, wrote [11]. From the mathematical point of view, the two quantities are completely symmetric: the classical mass can be generalized to two quantities with different tensorial characters [1].[5] One strong objection to relativistic mass is that

$$m_r = \frac{E}{c^2} \tag{50.5}$$

varies with the reference frame: it should not define a physical property of a real particle or body. In spite of the last considerations, however, [5] claimed the importance of relativistic mass:

> Newton's Second Law […] was stated with the tacit assumption that *m* is a constant, but we now know that this is not true, and that the mass of a body increases with velocity […]. For those who want to learn just enough about it so they can solve problems, that is all there is to the theory of relativity – it just changes Newton's laws by introducing a correction factor to the mass.

[4] The philosopher of science Feyerabend considered it a ≪*relation*, involving relative velocities, between an object and a coordinate system≫, instead of a ≪*property* of the object itself≫ [4].

[5] Similarly, you can generalize the classical quantity "time" either to the proper time or to the zero component of (*ct, x, y, z*).

50.4 The Activity

We took a sample of 42 talented high-school students, aged 17 to 19, selected for attending a modern physics summer school. Our activity consisted in an interactive tutorial with proposals for individual reflections and group discussions. We gave each student worksheets containing open and closed questions; for the discussions, we asked for conceptual differences among the (classical) notions of mass examined first, then what classical meanings of mass underwent a change passing to relativity.

We began the rationale of our path with the quotation in Sect. 50.1, together with another one containing an operative definition of mass through weight, from *Principia*. An analysis of inertial and gravitational mass concepts in Newtonian theory followed, as well as an overview of Mach's definition of the former.[6] We then introduced proper time using light clock, built 4-displacement through analysis of the world lines, and came to quadrimomentum dividing by proper time, in analogy with classic momentum. We finally performed a Taylor series expansion of the temporal component in the Newtonian limit and defined this quantity as relativistic kinetic energy, apart from an additive constant. The identification

$$E_0 = mc^2, \tag{50.6}$$

as well as its interpretation, were straightforward.

We asked the students some questions after the path:

1) "When does mass plays a role in your everyday life and which phenomena is it involved in?"
2) "What theories of physics do study these phenomena?"
3) "What do you mean when you talk about the *quantity of matter*?"
4) "What connotations and definitions of mass do you know?"
5) "Does the inertial mass of a body change in function of its energy, apart from the kinetic energy?"
6) "In many textbooks *relativistic mass* is mentioned. Explain what it is".

50.5 Data

Histograms from the collected data follow. For the analysis, we used "vertical" and "horizontal" modalities. The former consists in a separate examination of every answer: we aimed at categorizing them and finding their distribution in the student sample. The latter was a search for each student's own way of "looking at" mass – examining the correlations among answers – to recognize the five profiles, or levels of "physical representation", pointed out by Doménech et al.

[6] We also inserted a reference to mass as the result of interaction with all other masses in the Universe.

They are

- ontological (pre-theoretical[7]): mass is either a general property of matter or completely identified with bodies/matter/particles. In this case *quantitas materiae* becomes the best example;
- functional: the quantity is identified with a property, behaviour or trend of a physical system, for instance *inertia* or *heaviness*. An implicit theoretical framework is already present here;
- translational: the quantity is identified with another one whose definition is assumed as well stated, for instance mass as either *density/ volume,* or *weight*;
- relational: the quantity is defined through precise conceptual relationships inside a formal theory made by a number of mathematical laws.[8] In our analysis this representation was reduced to a clear and well-defined outline of *conceptual* relations;
- operational: the quantity is a number that you obtain through a series of explicit and achievable operations; gravitational mass comes out of a measure with an equal arm balance.

Moreover, answers to questions 4–6 allowed us to recognize several scientific meanings of mass: inertial, gravitational, Mach's empirical meanings, mass in SR (referred as "energy", "rest energy", ≪rest mass≫ by students), "relativistic mass.",[9] together with *quantitas materie.*

50.5.1 Analysis and Results

For questions 1–3, we divided described phenomena in classes; then physical theories that students join to phenomena were grouped, as well as things described as quantity of matter. The categories are not mutually exclusive.

We found out that in free fall mass is taken into account by a high percentage (17 %) of students, showing they haven't understood the physical meaning of the Equivalence Principle Fig. 50.1.

We also found that 24 % of the sample considers kinematics as a theory in which mass is a prominent quantity and that 8 students don't distinguish areas from theories as interpretative models for phenomena. From the third question emerges that 32 % of students identify correctly "quantity of matter" with mole, both in a general sense (ontological profile, 10 %), and as *numbers of moles* (relational profile, 22 %). However, identification with mass is present in 15 % of the sample.

---

[7] It does not refer to a theoretical framework. The definition is then concrete, in spite of appearance: it's implied by a concrete view of physical world.

[8] Mass is given by F/a in the inertial sense and by P/g, or through universal gravitation law as well, in the gravitational sense.

[9] We took these meanings from [2] and extended it to Special Relativity.

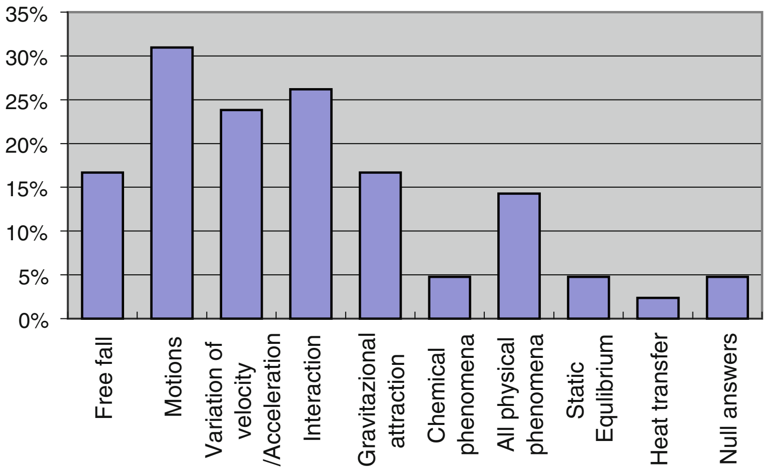


Fig. 50.1 Recalled familiar phenomena, "null answers" stands for answers not concerning the first question

A meaningful quotation from an answer classified in Fig. 50.2 as*quantitas materiae*is "Mass [...] can be derived from a formula m = ϱ/V". 45 % of students only listed mass meanings in these answers.

Only 35/42 students answered to the fifth question. We found no conceptual reference at all to the mass-rest energy equality in 40 % of cases. This lack is associated in several cases (20 % of the total) to the presence of the idea of "relativistic mass" in students, in an explicit or implicit form, although we stressed the role of relativistic energy in our rationale. In contrast, the conceptual reference is present in 43 % of cases, in implicit form or explained in words. There are relatively many uncertain

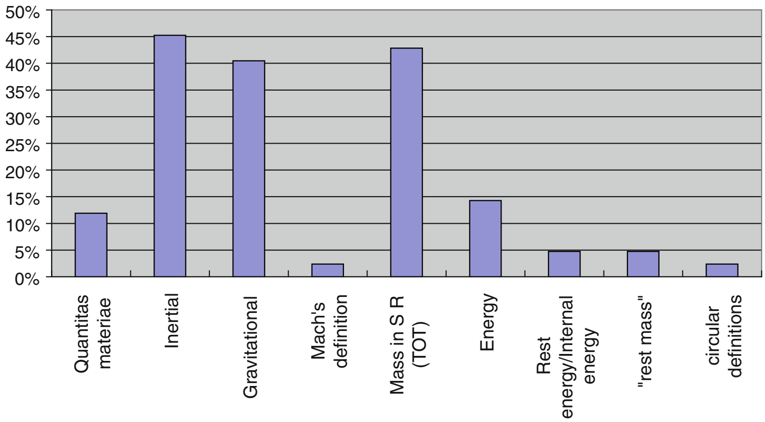


Fig. 50.2 Meanings of the idea of mass (IV question)

answers (14 %), that consists in enunciations, invocation of a generic mass-energy relation or not understandable sentences: 3 students answered "No, because kinetic energy doesn't affect the rest energy", showing that very likely in their conception rest energy is completely unrelated to (inertial) mass.

Thirty-nine students answered to the last question. The concept of "relativistic mass" appeared integrated in Einsteinian paradigm in 36 % of the sample. We could not find, by contrast, this integration in 31 % of the students. A meaningful example for the category "mass at relativistic speed" is the following: ≪That means that mass in motion at very high speed can become energy and vice versa≫. According to this conception, only at high velocities we enter in the realm of relativity, where mass in motion could turn energy. The correct definition – mass depending on speed – was given instead by more than 15 % of students, also using formalism; one among them has directly deduced the formula, with the aim of reducing relativistic expression for momentum to the classical one.[10] These students learned the right concept of "relativistic mass" in the proper theoretical framework.

As regards students' profiles, the relational one is clearly dominant: it affects about 60 % of the sample. Notice that there is no operational profile, that only one among students uses a partially translational representation and four students the ontological one Fig. 50.3.

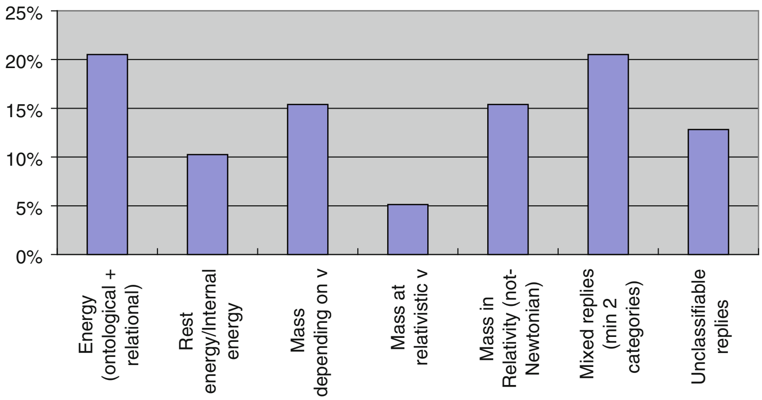


Fig. 50.3 "Relativistic mass" (VI question). The first category includes both a generic relation with energy (ontological level) and a formula connecting E, m, $E_0 m_0$ in every possible combination (relational level)

[10] This is, as we have seen, the crucial issue in the debate on this topic.

50.6 Conclusions

Our first research question was "How and in which contexts did our students relate themselves to the term 'mass' and make use of it?" Some interesting elements came out. Phenomena – as well as physical quantities – recalled in familiar semantic areas are in large part mechanical ones, except for quantity of matter, "force fields", "gravitational field". The pre-theoretical conception *quantitas materiae* is rooted in most of students' minds; however, students that answered the first group question never made use of it, as verified in the analysis of video recordings. A few students are aware of the importance of mass in electromagnetism and no one can contextualize it in familiar phenomena. "Ubiquity" is a character of mass expounded by 6 students, but it's not been rationalized by the half of them. In the end, mental representations of mass seem to be strongly affected by learning areas, so it is important to design integrated teaching [3]. We noted in particular a local view of the mass in SR in a context defined by speed and a grasping of the concept of mass in SR as limited to a "chapter" of physics.

Our second research question was "How did students interpret the extension of the concept 'mass' to 'mass as rest energy' in the relativistic context (under the influence of our path)?" It came out that the young talents show very good ability to formalize, the relationship (50.6) being an important exception in this regard. Besides, *terminology* plays an important role in the proper understanding of mass in relativity and in framing its conceptual relations with total energy, rest energy, "relativistic mass". Wrong answers about the latter – reporting learning difficulties concerning relation (50.2) – have been found in 26 % of the students indeed. Moreover, misconceptions found about mole were probably generated by the identity between the word used for the physical quantity and for the unit.

References

1. Bickerstaff RP, Patsakos G (1995) Relativistic generalization of mass. Eur J Phys 16:63–68
2. Doménech A, Casasús E, Doménech MT (1993) The classical concept of mass: theoretical difficulties and students' definitions. Int J Sci Educ 15(2):163–173
3. Fabri E (2007) Insegnare relatività nel XXI secolo, selected contributions from A.I.F. 2000 summer school
4. Feyerabend PK (1965) Problem of empiricism. In: Colodny RG (ed) Beyond the edge of certainty, Prentice-Hall, Englewood Cliffs
5. Feynman RP, Leighton RB, Sands M (1969) Feynman lectures on physics, vol.I, Addison Wesley, Reading
6. Jammer M (2000) Concept of mass in contemporary physics and philosophy. Princeton University Press, Princeton
7. Kuhn TS (1970) The structure of scientific revolutions. University of Chicago Press, Chicago
8. Lehrman RL (1982) Confused physics: a tutorial critique. Phys Teach 20:519–523
9. Mach E (1883) La meccanica nel suo sviluppo storico-critico, trad. it. di Die Mechanik in ihrer Entwicklung historisch-kritisch dargestellt

10. Okun L (2001) Photons, clocks, gravity and the concept of mass. In: Transparencies of the 18th Henry Primakoff lecture at the University of Pennsylvania, Philadelphia, USA, April 11, 2001 and of the special lecture given at the 7th international workshop on topics in astroparticle and underground physics, Laboratori Nazionali del Gran Sasso, Italy, 9 Sept 2001
11. Tolman RC (1934) Relativity, thermodynamics, and cosmology. Clarendon Press, Oxford

Chapter 51
Theories as Crucial Aspects in Quantum Physics Education

Marco Giliberti

Abstract A lot of difficulties aroused in interpreting quantum physics from its very beginnings up today are still at the core of most educational presentations. In fact, usually people try to grasp theoretical concepts by referring to a blend of ideas taken from scientific, pre-scientific and common sense schemes. In this paper a pre-condition for every educational approach will be proposed: to rigorously keep to quantum mathematical formalism in order to understand the meaning and the "reality" of quantum physics. It will be argued that the addition to quantum theories of most extraneous concept or common sense scheme comes from an ambiguous idea of nature, scope and aims of science itself.

51.1 Introduction

Many difficulties faced by most educational reconstruction of quantum theories have epistemological reasons, mainly embedded in the so called paradoxes of the quantum world. Even the educational paths that do not follow a historical (more often a pseudo-historical) approach, often lack a crucial point: that the meaning and the "reality" of quantum physics, just as the meaning of every physics topics, must be read from the theory, i.e.: the meaning and the concept of force should to be stressed from newtonian mechanics and the "reality" of the electric field should, in physics, be understood by Maxwell equations.

In this paper the lack of consciousness about the proper and exclusive nature of physical theories is related to thess well known difficulties and it is proposed that, in order to avoid many misinterpretations, which especially come by connecting quantum theories with classical physics concepts in wrong ways and by trying to reduce it to common sense schemes, a pre-condition is crucial: in quantum physics education please rigorously keep to quantum mathematical formalism (that is that of Quantum Mechanics or that Quantum Field Theory) and its interpretation.

M. Giliberti (✉)
Physics Department, University of Milan, Milan, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_51,

51.2 Nature of Difficulties: Theories

First of all our language describes an image. [...] It's clear that, if we want to understand the meaning of what we say, we must explore the image. But the image seems to save us this effort; it already hints at a determined use. So it mocks at us [24].

The undetermined "mess" to which we give the name "reality" is subjected to continuous changes because the status of the supposed entities that should form this "reality" is very flexible. Important and very known examples of these changes are the idea of absolute time, that of a luminiferous ether as a medium for the electromagnetic waves and the very idea and structure of atoms [3].

In fact, it is not possible to separate the object of knowledge from the instrument of knowledge: they must be considered as a whole. And this aspect is one of the main teaching of quantum physics from its very beginning. Theories are but mental constructions that help us find and define reality and utilize its resources. Physics inquiry begins with schemas and concepts that need not to be explained by theories; they come from common notions and language. Then proceeds introducing other concepts by means of what we can call "pretheories", that is already known physics theories stated as granted. Pretheories are unavoidable features and, for instance they are at the basis of our understanding and designing of measuring devices. Finally, our inquiry arrives at physics theories, that are determined by the basic concepts introduced before (common schemas plus pre-theories) plus formalized well defined new disciplinary concepts [19]. Actually we can imagine facts like icebergs, submerged under the surface of the sea of immediate experience that is perceived through common schemas. The submerged part of these facts can only be hypothesized, or, in a sense, imagined. A coherent and formalized imagination of this under-the-sea level of reality gives rise to physics theories. We can even think of these imagined realities as fairy tales; in this sense a scientific explanation is a story about how some entities, that are imagined but considered as real, would, by their very hypothesized nature, have worked together to generate the phenomenon to be explained [20]. A very simple and clear example of what is stated above about the construction of useful entities in theories as fairy tales can be found in star constellations. They are useful in finding the way on open sea and they are obviously man made constructions. But are they real? In what sense of the word real? Do they form natural entities? [21]. A way to answer to the previous questions is thinking that the truth value of a theory is given by the complete set of its mathematical formalism, by its field of applicability and by the rules of correspondence between these two [5].

A more physical example can be found in classical mechanics. In the mechanical description of the world, to the word force a reality in sé is often associated, as if forces were independent and external elements of reality. On the contrary, in physics, they find their meaning in the context of the Newtonian theory with its three principles. Newtonian mechanics is not a way to describe forces, but a conceptual schema into which forces, by means of their formal connections with other elements of the theory, become a part of reality [5].

Difficulties clearly emerge when we start to believe in the very real objective existence of constellations and of forces as they were natural real entities independent on human scientific schemas. But after all, it could be said that these questions regarding the objectivity of the world have not prevented physicist from regarding their results as objective facts. However it should be clearly stated that when this attitude is put forward even to quantum physics, our understanding of the world explodes into paradoxes.

51.3 Nature of Difficulties: Quantum Physics

Initially science does not develop by reflecting on its foundations, but with the accumulation of facts and the assimilation of new knowledge. But sometime one meets great contradictions and the greater the contradictions, the greater is the success in overcoming them. One important example of contradictions can be found in the divergence between the classical model of the atom and the studies on the emission of light by substances that eventually brought to the birth of quantum mechanics [19]. We can say that when we face contradictions we are like a paramecium that meets an obstacle: at first he goes backward and then starts again to go forward in a direction chosen at random. One could advise him of a better direction, nonetheless what he knows is correct: He cannot go in that direction! [18].

In the developing of quantum physics at least three steps can be singled out.

(1) Old quantum physics: that is facts and interpretations from 1900 till about 1925. Examples are the problem of black body radiation and the model by Planck; the photoelectric effect with its explanation by Einstein and the various model of the atom, mainly Bohr's atom. It's a set of facts and interpretations with the background idea of the existence of the so called quanta. Old quantum physics is the first response given by physicists to faced contradictions, but by no means constitute a theory.

(2) Quantum mechanics: with its formalism given by Heisenberg, Jordan and Born or that given by Schrödinger and Dirac, is a non-relativistic theory with well defined axioms that describes the behavior of a finite number of interacting particles. There are many different formulations of Quantum mechanics i. e. are matrix formulation, wave function formulation, path integral and second quantization formulation and even non orthodox formulations like Bohm's one [22].

(3) Quantum field theory: it is a relativistic quantum theory and every relativistic quantum theory will look, at sufficiently low energies, like a quantum field theory [23]. The most known example of a quantum field theory is quantum electrodynamics, but even the standard model of particle physics is a quantum field theory.

(4) And here come the problems: popular and even didactic interpretation of quantum physics very often mix these three parts together with great ingenuity. Moreover they often focuse on old quantum physics that, as stated above, is not even a theory and therefore it cannot be a reference for understanding.

51.4 A Pre-step Solution for Quantum Physics Education

Instead of working in the general mess of the old quantum physics, before educational reconstruction of the topic, as a pre-step, we should choose a reference theory in one of its formulation, identify the concepts of the theory in this formulation and understand their meanings inside the theory [5].

In my opinion the framework of quantum field theory is best suited than quantum mechanics for quantum physics education in general and, more specifically at high school. One of the reason is that only in quantum electrodynamics the concept of photon can be well defined. At the University of Milano we have been working on this subject since 1995 and many encouraging results have come [2]; [4]; [6–8]; [10–17].

Anyway in general, whatever reference theory is chosen, in physics education one has to try to avoid misunderstanding of ideas and words that are used by the theory, but come from preceding conceptions (pre-theories, or even common sense) rooted in the biased idea that physics reality can be identified even out of a formal theory. A clear example of this can be seen in the idea of particle in quantum physics. When we speak of particles we should clearly state what we mean, what the theory allows us to think of, and what it does not.

We must be careful that with that word we do not implement the idea that particles indicate physics entities in the sense of an ingenuous realism, and that these (ingenuous) entities coincide with the quanta of the theory.

As it has been previously pointed out, in scientific construction, reality comes out of a set of coherent interpretations of the formalism of the reference theory. In this sense the common word "particle" is nothing but a useful metaphor of what is meant by the theory. In quantum mechanics particles of the same "kind" are identical, not only because they have the same charge, the same mass, the same spin,... but also because they are indistinguishable even through their position. They are identical because they have the same physical properties.

We could think of a system with two electrons of different energy. In this situation we can say that one electron has a certain energy while another one has another energy, but it could be impossible to answer to the question "Which electron has which energy?". In more formal terms, the wave function obtained by the exchange of this two particle would yield the same previsions for the measurements of every observable. It is thus clear that the intuitive semantic content of the word "particle" given to the quantum mechanics quanta is, in general not adequate.

From an educational point of view I believe it could be much clearer if we spoke of quanta as linked to the excitations of the normal mode, as it is done in quantum field theory, in this way it would be evident that they are identical and indistinguishable... and have little to do with the "usual" particles. It is the event of revelation of a quantum in a device that drives us to use the word "particle", giving a metaphorical sense to a word coming from classical physics (and from common language).

If one strictly follows the guide of a theory most of the paradoxes (in quantum physics more or less all coming from wave/particle dualism) become not so central, with a great help for teaching.

51.5 A More Deep Difficulty

I'm not claiming that keeping in close touch with a quantum theory all difficulties run away. In fact an objective problem remains and it comes from the theory itself. In formulating quantum mechanics (and even quantum theory of fields) the world must be split in two parts; in this way we have a dichotomy: a microscopic quantum physics description for the system we are studying and a macroscopic classical description for measuring devices; experimental context and results must be described in classical terms. Quantum mechanics cannot even be formulated without this distinction. The problem is that the theory gives no indication on how "to cut the world". Quite obviously to get information from a microscopic experiment we must produce an amplification process that leads to a macroscopic change; but, as even the apparatus are in principle describable in terms of quantum mechanics, this leads us to an aporia. Is there a macroscopicity parameter? Not in the theory. Nor it seems in experiments (see for instance experiments of diffraction of macromolecules that show their wave-like proprieties [1] or the quantum macroscopic proprieties of a superconductor). So some deep difficulties are still rooted in quantum mechanics itself: there's really no need of making educational path that instead of presenting difficulties where they are, generate confusion mixing aspects, ideas and words coming from a too ingenuous vision of reality.

51.6 Conclusions

In conclusion this paper is a call to realism. I would like to stress that, as we already (some time unconsciously) do for classical physics, when dealing with quantum physics education we keep closely to a specific formulation of the theory. We can choose among one of the many formulation of quantum mechanics or one of the formulation of quantum filed theory (we in Milan suggest the latter, for his more easily grasped epistemology, linked to specific space time field instead of the configuration space wave function of quantum mechanics).The path to follow and the results obtained in experimentations are but a secondary problem in this perspective: they come after.

Historical or conceptual presentations and educational reconstructions of the topic cannot in our opinion skip this point, as instead many times they do keeping an eye closer to the path than to the goal.

References

1. Arndt M, Olaf Nairz O, Gerbrand van der Zouw G (1999) Wave-particle duality of C60. Nature 401(680–682):14
2. Bartesaghi P, Camera F, Cazzaniga L, Costigliolo M, Giliberti M, Labanca I (2004), Alpha natural radioactivity in the air: an experiment for high-school students. In: Mechlovà E (ed) Proceedings GIREP conference teaching and learning physics in new contexts", University of Ostrava Ostrava, 215–216 ISBN 80-7042-378-1
3. Bellone E (2006) L'origine delle teorie, Codice, 118 ISBN 8875780579
4. Bergomi N, Giliberti M (2001) Modelling rutherford scattering: a hypertext for high-school teachers net-training. In: Pinto R, and Surinach S (eds.) International conference physics teacher education beyond 2000. Selected contributions, international ICPE-GIREP conference, Barcellona 28 Agosto-1 Settembre 2000, Elsevier Editions, Parigi, 145–146 ISBN 2-84299-312-8
5. Cavallini G, Giliberti M (2008) La lezione della fisica quantistica, Epistemologia (XXXI), 219–240.
6. Giliberti M (1997a) Teaching about Heisenberg's relations. In: Proceedings of the G.I.R.E.P. I.C.P.E. international conference: new ways of teaching physics Ljubljana 21–27 agosto 1996; board of education of Slovenia, Lubiana (1997), 533–534 ISBN 961-234-051-X
7. Giliberti M (1997c): Popularisation or teaching? How much math in physics courses? In: Proceedings of the G.I.R.E.P. I.C.P.E. international conference: new ways of teaching physics Ljubljana 21–27 August 1996; board of education of slovenia, Lubliana (1997), 408-411 ISBN 961-234-051-X
8. Giliberti M (1998), Il ruolo della matematica nell'insegnamento della fisica moderna, La fisica nella scuola, XXXI, 1 Supplemento 162–166
9. Giliberti M (2002), A modern teaching for modern physics in pre-service teachers training First international GIREP seminar developing formal thinking in physics; selected contributions Michelini M, Cobal M, Udine, Forum, Editrice Universitaria Udinese, 403–408 ISBN 88-8420-148-9
10. Giliberti M (2002b): Alcuni elementi di didattica della fisica moderna nel corso di "Teorie Quantistiche" della S.I.L.S.I.S.-MI; La Fisica Nella Scuola, XXXV, 2 Supplemento (2002), 56–75
11. Giliberti M (2007) Elementi per una didattica della fisica quantistica, Milano, CUSL, Materiali didattici dell'indirizzo Fisico Informatico e Matematico SILSIS-MI, 2. ISBN 9788881324736
12. Giliberti M (2008) Campi e particelle: introduzione all'insegnamento della fisica quantistica", in Approcci e proposte per l'insegnamento apprendimento della fisica a livello preuniversitario, a cura di P. FORUM Editrice Universitaria Udinese, Guidoni e O. Levrini, pp 161–181
13. Giliberti M, Barbieri S (2004a) Teorie Quantistiche. CUSL, Milano. ISBN 88-8132-305-2
14. Giliberti M, Cazzaniga L, Prosperi GM, Rinaudo G, Borello L, Cuppari A, Rovero G, Michelini M, Santi L, Sciarratta A, Sonego S, Ragazzon R, Stefanel A, Ghirardi G (2002a) Proposte per l'insegnamento della Meccanica Quantistica in SeCiF; Atti del Congresso Didamatica 2002 Informatica per la Didattica; Liguori Editore, Napoli, luglio 2002, 15–28 ISBN 88-207-3455-9
15. Giliberti M, Lanz L, Cazzaniga L (2004b) Teaching quantum physics to student teachers of S.I.L.S.I.S.-MI. In: Michelini M (ed) Second international GIREP seminar "quality development in teacher education and training", selected contributions FORUM Editrice Universitaria Udinese, Udine, 425–429 ISBN 88-8420-225-6
16. Giliberti M, Marioni C (1996), Revisiting the H atom and electron configurations in atoms: a case study at high-school level. In: Proceedings of the G.I.R.E.P.-I.C. P.E. international conference: teaching the science of condensed matter and new materials, Udine, August 1995, 24–30 Forum Editrice Universitaria Udinese, Udine, 168–171 ISBN-88-85756-11-9
17. Giliberti M, Marioni C (1997b), Introduzione di alcuni elementi della fisica dei quanti nella scuola secondaria superiore, La fisica nella scuola, XXX (3) Supplemento, Q7, 23–45
18. Lorenz K (1973) L'altra faccia dello specchio, Adelphi, Torino, 25–26 translated by the author
19. Ludwig G, Thurler G (2008) new foundation of physical theories. Springer, Heidelberg, 4

20. Ogborn J (2010) Science and commonsense. In: Vicentini M. and Sassi E. (eds) Connecting research in physics education with teacher education,New Dehli, Gautam Rachmandani7
21. Stenholm S (2002) Are there measurements?. In: Bertelmann RA, Zeilingfer A (eds) Quantum unspeakables: from bell to quantum information, Springer, Berlin Heidelberg, p 188
22. Styer D et al (2002) Nine formulation of quantum mechanics. A. J. Phys 70(3):288–297
23. Weinberg S (1995) The quantum theory of fields, I. Cambridge university press, Cambridge 2
24. Wittgenstein L (1967), Ricerche filosofiche, Einaudi, Torino, 244, translated by the author

Chapter 52
An Interference/Diffraction Experiment for Undergraduates

Milena D'Angelo, Augusto Garuccio, Fabio Deelan Cunden, Francesco Fracchiolla and Nicola Minafra

Abstract We present an educational double-slit experiment aimed at strengthening students' ability in physical reasoning, both from a theoretical and an experimental perspective, while improving their understanding of interference/diffraction phenomena. In particular, students are lead to focus their attention on the hypotheses employed in the reference theoretical model (i.e., Fraunhofer or far-field diffraction), and are guided to build an experimental setup that satisfies such conditions. For completeness, we also present some quantitative results based on two different measurement techniques: The direct measurement of intensity versus position, and digital photography.

52.1 Introduction

Young double-slit experiment is one of the most common educational experiments performed in undergraduates optics labs. Beside representing a simple tool to familiarize with interference/diffraction phenomena, the double-slit experiment is a quite useful preliminary step to the experimental discovery/verification of Fresnel-Arago law, which naturally leads to address the role of polarization in the context of coherence and may help to guide students toward the intriguing classical-to-quantum transition [1].

In this paper, we propose to exploit the double-slit experiment to strengthen students' ability in physical reasoning, while guiding them to develop some essential lab skills. In fact, one of the leading ideas is that the experimental study of the double-

M. D'Angelo (✉) · A. Garuccio
Dipartimento Interateneo di Fisica, "M. Merlin" dell'Universita degli Studi e del Politecnico di Bari and INFN sez. di Bari, Bari, Italy

F. D. Cunden · F. Fracchiolla · N. Minafra
Dipartimento Interateneo di Fisica, "M. Merlin" dell'Universita degli Studi e del Politecnico, Bari, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_52,

slit interference/diffraction phenomenon offers students the precious opportunity to recognize the physical meaning of the mathematical approximations at the heart of their reference theoretical model (i.e., Fraunhofer diffraction), and to develop strategies to experimentally satisfy them. Our approach may become particularly useful in view of the introduction of Michelson stellar interferometer, as well as spatial coherence through measurement of fringe visibility (see, e.g., Refs. [2–4]).

In this respect, we aim at focusing students attention on the hypotheses on which the reference theoretical model is based; in particular, we guide students to build the experimental setup in such a way that it satisfies the "plane wave approximation." Such steps naturally lead students to face the strong setup-dependence of this approximation. Once students are confident with the phenomenon under consideration, they are guided to build their own setup while taking into account some practical issues such as the resolution limit of their measurement device, thus getting ready to perform quantitative measurements on the observed interference/diffraction pattern.

52.2 Theoretical Background

Interference/diffraction phenomena are generally introduced by considering a monochromatic plane wave impinging on a double-slit mask followed by an opaque screen, parallel to the plane of the mask, and placed far enough from it that the impinging radiation can be assumed to be made of plane waves ("Fraunhofer or far-field approximation"); the angular intensity distribution in the interference/diffraction pattern observed on the screen is well known to be given by [2]:

$$I(\theta) = 4\, I_0 sinc^2\left(\frac{\pi a}{\lambda}\sin\theta\right)\cos^2\left(\frac{\pi d}{\lambda}\sin\theta\right) \tag{52.1}$$

where λ is the wavelength of the incident radiation, a is the slit width, d the center-to-center slit separation, θ is the observation angle with respect to the direction orthogonal to the mask (for an incoming plane-wave at normal incidence: $\theta_0 = 0$), and I_0 is the intensity of the radiation diffracted by a single slit in the direction of incidence of the incoming plane-wave (i.e., at $\theta = \theta_0 = 0$). A schematic setup at normal incidence is reported in Fig. 52.1, together with an (out-of-scale) sketch of the intesity distribution given by Eq. (52.1).

At the undergraduate level, Eq. (52.1) is seldom demonstrated. Most often the positions of maxima and minima predicted by Eq. (52.1) are obtained by applying the superposition principle and the Huygens principle, so that the slits are seen as made up of point sources emitting in phase (or with a fixed phase delay defined by θ_0) thanks to planar wavefronts assumed for the incoming light. Moreover, the far-field hypothesis for the position of the observation screen is employed to simplify the expression of the optical path differences from the mask to a given point of the screen.

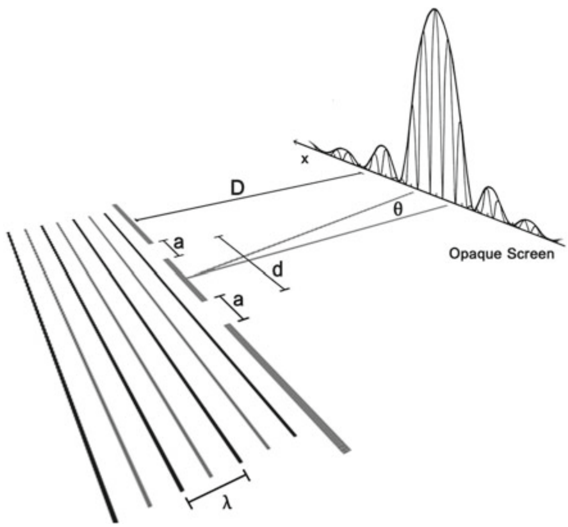


Fig. 52.1 Typical simplified version of a standard double-slit experiment

In the context of diffraction thoery, the plane-wave approximation at the heart of Eq. (52.1) is introduced in terms of the Fraunhofer approximation, and is said to be satisfied when the smaller between the source-to-mask and the mask-to-screen distances (B and D, respectively) satisfies the condition: $min(B, D) >> (a + d)^2/\lambda$. This condition can also be expressed in terms of angles:

$$max(\Delta\theta_{in}, \Delta\theta_{out}) << \Delta\phi, \quad (52.2)$$

where $\Delta\theta_{in} = \arctan[(a + d)/B]$ is the angle under which the aperture is seen by the point source, $\Delta\theta_{out} = \arctan[(a + d)/D]$ is the angle under which the aperture is seen by an observation point, and $\Delta\phi = \arctan[\lambda/(a + d)]$ is the angle subtended between the aperture and the distance between two consecutive wavefronts (λ), as indicated in Fig. 52.2. Such a pictorial representation could be drawn and discussed with students for them to realize the interplay between the involved parameters for satisfying the plane-wave approximation. For instance, as the aperture transverse dimension increases, while keeping the wavelength fixed, $\Delta\phi$ decreases and the input and output divergence angles subtended by the aperture increase, thus requiring larger distances B and D for satisfying the plane-wave approximation (Eq. (52.2)).

The importance played by the plane-wave approximation for obtaining either Eq. (52.1) or the positions of the intensity maxima/minima, is generally not explicitly emphasized when dealing with interference/diffraction phenomena, neither it is its physical meaning. This may be adequate when first introducing the phenomenon at the undergraduate level, however, a lab experience offers students the opportunity to realize that this approximation is not valid in an absolute way, but it rather depends on the setup parameters. A wavefront appearing planar to a given

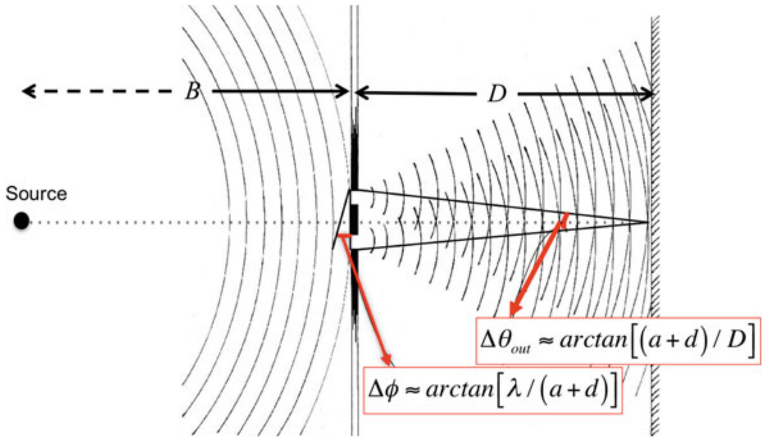


Fig. 52.2 Schematic representation of the "plane wave" or "Fraunhofer" or "far-field approximation": The slit width a and slit-to-slit distance b are such that the portion of wavefront incident on the double slit can be taken as planar. At fixed λ, B, and D, greater values of a and/or d may not satisfy this approximation

double slit, will not be so for a double slit with larger transverse dimensions; the same can be seen in terms of different wavelengths. By analyzing these kinds of situations, students begin to acquire an operational tool to better interpret the physical meaning of the mathematical approximations that are always at the heart of reference theoretical models; in particular, in the present situation, students get a more correct feeling for the concept of plane wave, a better understanding of the laser beam as its best approximation (a pretty common misconception students have until they happen to study Gaussian beams), and, most important, learn do develope experimental procedures for verifying whether or not their setup verifies the theoretical hypotesis.

52.3 Building the Setup: Role of the Experimental Parameters

The relevant parameters in a double-slit experiment are: Source wavelength (λ), mask transverse dimension (in our case, center-to-center slit distance d and slit width a), source-to-mask distance B, and mask-to-screen distance D (as compared to the first two parameters).

In order to qualitatively understand the role played by the source spectrum and central wavelength, one may compare the interference/diffraction patterns observed by looking at LEDs of different colors (such as blue, green, yellow, and red) through a double slit ($a \sim 0.1\,mm$, $d \leq 1\,mm$) placed as closed as possible to the observer eye. Standing $2-3\,m$ away from the LED is generally sufficient for the light waves reaching the double slit to approximate plane waves, so that high visibility interference/diffraction patterns are observed. Besides observing the expected shrinking of the interference fringes with decreasing wavelengths, students have the opportunity to get an idea of how sensitive the double slit interference/diffraction pattern is to wavelength variations. In fact, the difficulty in comparing the width of interference fringes for yellow and red, as well as yellow-green and green-blue, suggests that double-slit interference would not be a very precise tool for measuring wavelength and/or wavelength differences, or, equivalently, that the condition for monochromatic light is not a particularly stringent one in the present experiment. This result may of course be employed to remind/introduce the role played, in diffraction gratings, by the number of slits for improving the spectral resolution.

The next step is to start building the experimental setup for performing more quantitative observations and measurements. Our experiment is genrally performed with either an He-Ne laser (JDSU) or red laser diodes (Futura Electronics), and by employing different sets of double slits from 3B Scientific.

As discussed in Sect. 52.2, although the source-to-mask and mask-to-screen distances do not appear in the intensity distribution of Eq. (52.1) (or, equivalently, in the positions of intensity maxima and minima) the plane wave approximation is essential to obtain this expression. In order to build a setup satisfying this approximation (Eq. 52.2), it is necessary to implement a strategy which guarantees observable differences in the interference/diffraction pattern as B and D go above and beyond the Fraunhofer condition.

Concerning the distance B, such a strategy may be based on the fact that Fraunhofer diffraction patterns are expected to be insensitive to transverse shifts of the diffracting mask: In the far field limit, the diffraction patterns produced by each single slit, separately, perfectly overlap. Hence, by observing on a distant screen ($D \sim 2\,m$) the pattern produced by each one of the two slits separately (i.e., when the other one is covered), one can easily verify whether or not the incident light beam approximates a plane wave. By signing on the observation screen the center and width of each single-slit diffraction pattern, students will notice that the larger the distance B (i.e., the smaller the divergence $\Delta\theta_{in}$ of the incident wavefronts as seen from the diffracting aperture), the more the two single-slit diffraction patterns get closer and tend to overlap. Students should be brought to qualitatively analyze and justify the variation, with both a and d, of the minimum distance B_{min} giving an acceptable plane wave approximation. Two examples of such observations are reported in Fig. 52.3. Beside focusing their attention on the operational meaning of the plane wave approximation, students have a chance to avoid, once and for all, one of the most common mistakes related to far-field interference/diffraction phenomena: The idea that each slit, separately, would give a far-field diffraction pattern centered on the slit center (idea which is in clear contrast with the interference fringes appearing when both slits are open). This preliminary qualitative analysis thus allows students to face some of the well known difficulties related to their understanding of the wave model for light (e.g., the tendency to inappropriately apply and combine elements of geometrical and physical optics to a given experimental situation [5]).

Once the distance B is adequately chosen so that most of the available double-slits see an incoming plane wave, the variation of the interference/diffraction pattern

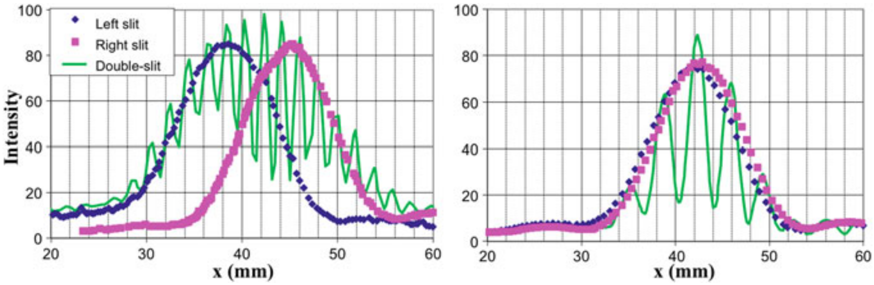


Fig. 52.3 Experimental data obtained by scanning the distant observation plane when either one of the slits is coverd and when both slits are open. Experimental parameter are: $B = 8\,cm$, $D = 264\,cm$, $a = 0.15\,mm$, $d = 1\,mm$ (*left plot*) and $d = 0.5\,mm$ (*right plot*). In the left plot the plane-wave approximation is clearly not satisfied, and the distance between the two diffraction patterns is comparable with their width. In the right plot the plane-wave approximation is pretty well satisfied, as the distance between the two diffraction patterns is reduced to 0.03 of their width

with the mask-to-screen distance D can easily be observed by simply displacing the observation screen starting from the smallest possible distance from the mask.

Now that B and D have been chosen so that most of the available double-slits satisfy the far-field condition, students can take note of the positions of a few interference and diffraction maxima and minima, for different double-slits. Beside checking the role played by a and d on the interference/diffraction fringes, students should be brought to choose both the mask and the mask-to-screen distance D in order to guarantee the better possible precision for such position measurements, as required by any standard quantitative analysis.

In Fig. 52.4, for completeness, we show the results obtained by students who autnomally chose to compare two different measurement devices: A *Lucegrafo* [6] and a digital camera (Canon Digital Rebel with Canon EF-S 18–55 mm lens). Both instruments, relatively cheap and easy to use, allow precise measurements, speed up the data acquisition, and give direct information about the intensity distribution in the observed interference/diffraction pattern. The results obtained by employing the *Lucegrafo* are more precise than those obtained with the more noisy digital camera: The error goes from 7 % in the first case, to about 14 % in the latter. Still, the achieved accuracy is better than the one generally obtained with graph paper (of the order of 20 %). The precision of the first method could be further improved by increasing the statistics: By increasing the photodiode gain diffraction peaks over the 10th order have been measured. In general, both methods are relatively simple and powerful, and offer students the chance to develop basic lab skills and familiarize with measurement techniques and instruments, thus enabling them to develop knowledge and competences that will be useful in a broad variety of lab experiences.

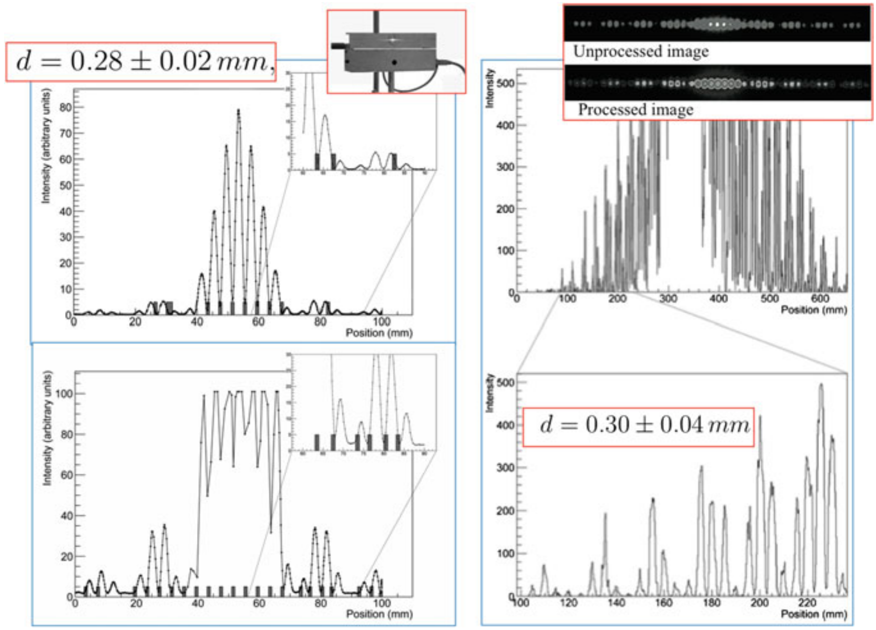


Fig. 52.4 Interference diffraction pattern of a double-slit with nominal values: $d = 0.3\,mm$ and $a = 0.07\,mm$, measured at a distance $D = 180 \pm 2\,cm$ from the mask. Left column: same pattern as mesured at two different gain settings of the lucegrafo (shown in the *rigth top* corner); the grey strokes are the minima found by the program written by students to automatically find the intensity minima. Rigth column: histogram of the average intensity recorded in all vertical lines of the processed image shown in the right top corner (where the comparison between the raw image taken by the camera and its contrast-enhanced version is shown)

52.4 Conclusion

We have presented a basic double-slit experiment that, thanks to its simple implementation, is adequate for any undergraduate optics lab. The essential point of the proposed experiment is to exploit one of the simplest optics lab experiments for strengthening students ability in physical reasoning and developing essential lab skills. In particular, we have indicated strategies for leading students to recognize the role played by certain basic hypotheses (such as monochromatic source and plane wave approximation), and to realize that the corresponding mathematical approximations, although never perfectly verified in practice, can be sufficiently well satisfied in well thought experimental setups. In fact, lab experiments are certainly the most immediate way for students to understand the operational meaning of many mathematical approximations, thus gaining physical insight into the phenomenon under consideration and the basic concepts connected to it.

The Authors sincerely thank M. Dabbicco and M. Ciminale for useful discussions and suggestions.

References

1. Hillmer R, Kwiat P (2007) A do-it-yourself quantum eraser. Sci Am 296:90
2. Hecht E (1998) Optics. Addison-Wesley, Reading
3. Bloor D (1964) Coherence and correlation. Am J Phys 32:936
4. Basano L, Pontiggia C, Piano E (1996) Simple demonstrations for introducing spatial coherence. Am J Phys 64:1257
5. Ambrose BS, Shaffer PS, Steinberg RN, McDermott LC (1999) An investigation of student understanding of single-slit diffraction and double-slit interference. Am J Phys 67:146
6. Gervasio M, Michelini M (2009) Lucegrafo. A simple USB data acquisition system for diffraction experiments. online at http://www.fisica.uniud.it/URDF/mptl14/ftp/full_text/T6_124_Lucegrafo_MPTL_MM.pdf

Chapter 53
Disciplinary Knots and Learning Problems in Waves Physics

Simone Di Renzone, Serena Frati and Vera Montalbano

Abstract An investigation on student understanding of waves is performed during an optional laboratory realized in informal extracurricular way with few, interested and talented pupils. The background and smart intuitions of students rendered the learning path very dynamic and ambitious. The activities started by investigating the basic properties of waves by means of a Shive wave machine. In order to make quantitative observed phenomena, the students used a camcorder and series of measures were obtained from the captured images. By checking the resulting data, it arose some learning difficulties especially in activities related to the laboratory. This experience was the starting point for a further analysis on disciplinary knots and learning problems in the physics of waves in order to elaborate a teaching-learning proposal on this topic.

53.1 Introduction

Wave phenomena are everywhere. Everything can vibrate. There are oscillations and waves in water, ropes and springs. There are sound waves and electromagnetic waves. Even more important in physics is the wave phenomenon of quantum mechanics. When and how can it make sense to use the same word, wave, for all these disparate phenomena? What is it that they all have in common?

A first answer lies in the mathematics of wave phenomena. Periodic behaviour of any kind, one might argue, leads to similar mathematics. There is a more physical answer to the questions. If it is possible to recognize deep similarities in different physical phenomenology, then it is likely that we can describe them by mean of the same mathematical tools.

S. Di Renzone (✉) · S. Frati
Educational Innovation in Physics and Orienting, University of Udine, Udine, Italy

V. Montalbano
Physics Department, University of Siena, Siena, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_53,

In order to introduce interested and talented students on this intriguing field in which wave phenomena can be described most insightfully, we designed an optional laboratory within National Plan for Science Degree.[1]

In the following, we describe methodological choices made in planning an extracurricular learning path on waves and oscillations and some examples of activities. In particular, we outline the advantages of introducing mechanical waves by using the Shive wave machine [6]. Many laboratory activities can be proposed in which students explore waves behaviour in qualitative way, guess what can happen and suddenly test their hypothesis. Furthermore, we present same disciplinary knots that arise usually in empirical investigation, according to the Model of Educational Reconstruction [1, 2].

Finally, we describe learning problems on fundamental topics which arose in the laboratory. Further analysis on disciplinary knots is needed for a new teaching-learning proposal.

53.2 An Extracurricular Learning Path on Waves

In the last years, many Italian Universities are involved in National Plan for Science Degree.

The PLS guidelines are the following:

- orienting to Science Degree by means of training
- laboratory as a method not as a place
- student must become the main character of learning
- joint planning by teachers and university
- definition and focus on PLS laboratories.

There are several types of PLS laboratories. Laboratories which approach the discipline and develop vocations, self-assessment laboratories for improving the standard required by graduate courses and deepening laboratory for motivated and talented students.

We proposed two deepening laboratories for selected students[2] titled *Waves and energy* and *Sound and surroundings*. The laboratories are optional and the activities take place in Physics Department. We are meeting students for 3 hours almost every month and planning to continue for last 3 years of high school. We decided that for the first year both laboratories had the same introductory activities on waves physics and all students works together.

We planned activities by focusing on

- conceptual issues such as characterization of the oscillatory motion and energy aspects versus characterization of wave energy and energy transport

[1] Piano nazionale Lauree Scientifiche, i.e. PLS.

[2] Coming from third class of *Liceo Scientifico Aldi* – Grosseto, followed by their teacher G. Gargani.

- methodological issues in order to propose a complementary experience compared to what was done in class.

Students ended their learning path on waves and sounds in class before the laboratory starts.

Moreover, their class made an instruction trip to our department and perform a standard laboratory experience on diffraction and interference.

53.3 Shive Wave Machine

We chose to begin with a series of activities performed by student by mean of a Shive wave machine, showed in Fig. 53.1a.

This wave machine, developed by Dr John Shive at Bell Labs in '50, consists of a set of equally-spaced horizontal rods attached to a square wire spine. Displacing a rod on one of the ends will cause a wave to propagate across the machine. Torsion waves of the core wire translate into transverse waves.

Table 53.1 shows the main advantages and disadvantages of Shive wave machine versus wave tank, another educational device very common in school.

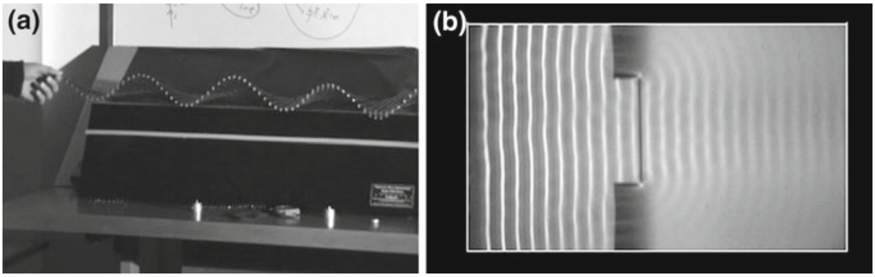


Fig. 53.1 **a** Shive wave machine, **b** Wave tank

Table 53.1 Shive wave machine versus wave tank

Shive wave machine	Wave tank
Easy student interaction	
Measure λ, T, v	Surface waves
Superposition principle	Measure λ, T, v
Reflection study	Reflection and refraction study
Energy considerations	Interference and diffraction
Easy study of stationary waves and resonance	
Limit of application	
One-dimensional wave	Lack of information on energy
Lack of study of refraction, diffraction and interference	Poor interaction with student

Measures were obtained by using a camcorder and extracted from the captured images.

Students used Shive wave machine for studying the following wave aspects:

- dependence on space-time of waves
- impulsive and periodic waves
- wavelength and frequency
- energy transfer
- speed of propagation
- superposition principle
- reflection and transmission.

For example, by coupling core wires of two Shive wave machine with different rod lengths, an impulsive wave can be reflected and transmitted through the discontinuity (two examples are given in Fig. 53.2). Students can measure from captured images all wave amplitudes and speeds and verify that energy is conserved.

53.4 Further Activities in Laboratory

After using Shive wave machine, sound waves were studied by using a microphone and an oscilloscope (Fig. 53.3a). Further verification of the principle of superimposition was made. Also beats and patterns of periodic beats (Moiré fringes) were studied.

Interference was studied for sound waves in order to stimulate reflection around similarities and differences between different kinds of waves (longitudinal versus transverse, three-dimensional versus two-dimensional, and so on). Figure 53.3b shows a schematic layout of interference experiment.

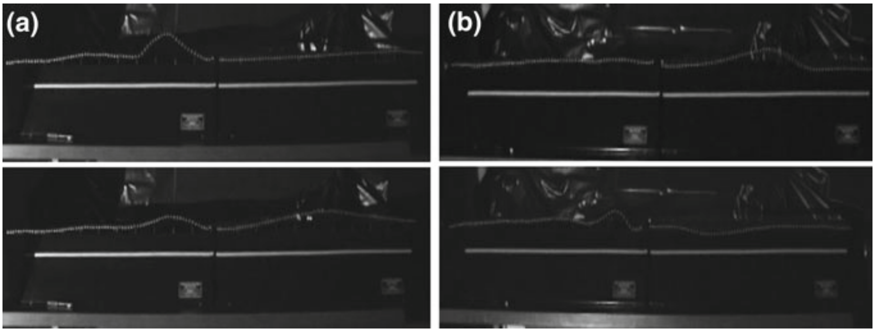


Fig. 53.2 **a** A pulse from the left (*top*) and transmitted and reflected pulses (*below*), **b** A pulse from the right (*top*) and transmitted and reflected pulses

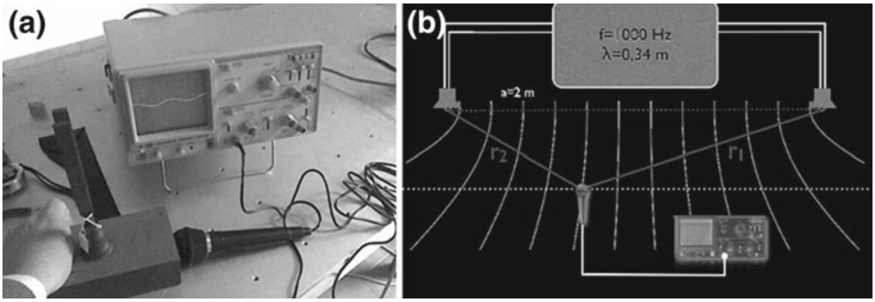


Fig. 53.3 **a** Characterizing a sound waves by using an oscilloscope , **b** Acoustic interference

This experiment is complementary to interference with laser light and allows measuring the speed of sound and discussing problems and connected applications of interference in acoustics.

In order to study resonance, we started from a mechanical system with proper modes which can be forced by induction. A ceramic magnet is hanging to a string suspended over an electromagnet. Students can vary the frequency of a square wave that power the electromagnet, they can measure the amplitude value varying the frequency of the forcing term.

The next steps will be to study resonance with the Shive wave machine, in acoustic and optical resonant cavities.

53.5 Some Disciplinary Knots in Wave Physics

Starting from our previous empirical investigation in teaching waves, we designed the learning path on waves described in previous sections in order to use it as a pilot study in an optional laboratory performed with few, interested and talented students, i.e. in the best conditions for teaching and deepening topics in physics.

We focused our attention on topics in which the main difficulties in learning usually appear.

We have considered the following conceptual knots:

- Waves as function of several variables; this usually is an hidden trouble. Even brilliant students can use for long times functions of one and several variables without any real understanding of deep difference.
- Energy transport is essential in order to distinguish waves from other periodic phenomena and comprehend many applications.
- Superposition principle is a fundamental concept. In wave physics, many phenomena can be clarified and some unexpected behavior can be explained by applying it.

- Analogy in waves phenomena; difficulties in this area are very common and reported [4, 5], especially in recognizing the same behaviour in different context such as total reflection, diffraction, beats, interference and so on.
- Resonance; despite it is a relevant phenomenon which runs through almost every branch of physics, many students have never studied it. Yet, resonance is one of the most striking and unexpected phenomenon in all physics and it easy to observe but difficult to understand.

53.6 Learning Problems

Despite excellent boundary conditions (motivated and talented students), despite optimal behavior and relationship between instructors and pupils (we both have really fun in making this laboratory), we have encountered learning problems on fundamental topics.

The first learning problem arose when we requested graphics for Shive wave machine position of one rod versus time (position fixed) and position of rods versus space (time fixed), in order to outline the dependences from space and time variables and put them in relation with measurable physical quantities.

Home students' task was to extract all data from video captured by themselves. We had outlined the importance of calibrating the video for time and space measures by giving examples and hints.

Students prepared two graphics with 3 experimental points with no errors, the points were 0, max, 0, i.e. they measured only the maximum wave amplitude and the period or the wavelength, guessed the nodal point must be 0 and make graphics by drawing one half period, like shown in Fig. 53.4.

The next time students presented graphics with about seven points and uncertainties of about 15 or 20 %

We were astonished. With camcorder, the data can have uncertainties of few percent.

After a brief discussion, we discovered that videos captured in the previous lab are useless for time calibration because digits on clock display are too small and they had forgot to measure a length in the devices that can be used as calibration.

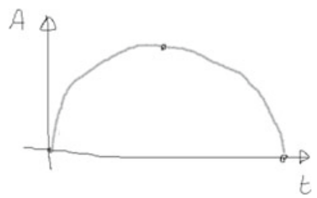


Fig. 53.4 A drawing like that shown was done on a large sheet of graph paper

Thus, they adopted the following strategy: measuring lengths on the computer display with a ruler and measuring time by a chronometer. You can imagine what precision could have measuring lengths on a computer display.....

On the other side, these students are very clever in executing tasks in which they have received instruction as in a laboratory at school where usually a worksheet is given.

In our opinion, they are still missing the main purpose of a measure: obtain the maximum of information from nature with the given devices.

Furthermore, how can they understand deeply the difference between a one dimensional wave and a function encountered in a math lesson, or between one dimensional wave and two-dimensional or space waves or surface waves, if they are not able to construct graphics that represent completely a physical situation starting from data?

53.7 Conclusions

We proposed and are testing a learning path in wave physics.

Some disciplinary knots were identified and laboratory activities were developed in order to help understanding in learning processes. Testing these activities in an optional laboratory with high school student can be considered the first step in order to develop a designed-based research learning path [3].

These laboratories seemed to be very successful but in the meantime students showed serious lacking in fundamental topics, especially in collecting properly and in using correctly the experimental data in order to describe physical system.

These lacks are the basis of observed learning problems and remains a real trouble for any further progress in knowledge.

It is crucial to analyze deeply the teaching-learning processes in undergraduate school in order to avoid, or at least, minimize these kind of learning problems.

References

1. Duit R (2007) Science education research internationally: conceptions, research methods, domain of research. EJMSTE 3:3–15
2. Duit R, Komerek M, Wilbers J (1997) Studies on educational reconstruction of Chaos theory. Res Sci Educ 27(3):339–357
3. Hake RR (2008) Design-based research in physics education research: a review. In: Kelly AE, Lesh RA, Baek JY (eds) Handbook of design research methods in mathematics, science, and technology, pp 493–508
4. Podolefsky NS, Finkelstein ND (2007a) Salience of representations and analogies in physics. AIP conference proceedings 951, (2007) physics education research conference, pp 164–167
5. Podolefsky NS, Finkelstein ND (2007b) Analogical scaffolding and the learning of abstract ideas in physics: An example from eletromagnetic waves. Phys Rev ST Phys Educ Res 3:010109
6. Shive JN (1959) Similarities of wave behavior, http://techchannel.att.com/play-video.cfm/2011/3/7/AT&T-Archives-Similarities-of-Wave-Behavior,. Accessed Oct 2011

Chapter 54
Lorentz' Force as a Tool for Physics Inquiry: Studying Particle Tracks in Cloud and Streamer Chambers

Pasquale Onorato and Anna De Ambrosis

Abstract A sequence of experiments aimed at exploring magnetic force on moving charged particles is presented. At first students work with equipment that reproduces Thomson's experiment of 1932. Afterwards they analyze images of sub-nuclear particle tracks in cloud and streamer chambers in historical experiments. The sequence has been tested with high school students. Our results compared with the ones reported in the literature indicate that students' understanding of the direction and magnitude of the magnetic force markedly improved and some typical difficulties were overcome.

54.1 Introduction

Many researches have investigated students' difficulties in understanding main features of magnetic force on moving charged particles [3, 4, 6, 7]. Results show how students easily confuse electric and magnetic fields and how they are inclined to think that the force experienced by a moving charged particle is directed toward a magnetic pole or has the same direction as the magnetic field lines independently of the velocity of the charged particle [8]. To help students overcome these difficulties our idea is to stress the peculiarity and uniqueness of this force as it appears in real experiments involving elementary particles. In this way Lorentz force and its characteristics are introduced as tools of physics inquiry rather than a formula to be memorized.

In this paper we present three of such experiments which are included in a wider sequence of activities on electromagnetic interaction. They have been designed in cooperation with a group of high school teachers involved in a program funded by the Ministry of Education and aimed at preparing students for science studies at university. Two main choices were shared by the group:

P. Onorato (✉) · A. De Ambrosis
Department of Physics "A. Volta", University of Pavia, via Bassi 6, 27100 Pavia, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_54,

- emphasizing the experimental approach;
- making students carry out quantitative measurements by themselves while working in groups.

In order to keep to these choices we resorted to an extensive use of digital camera and image processing software that made it possible for the students work directly not only on tracks obtained by the experimental apparatus available in the laboratory, but also on images acquired in historical experiments. The experiments included in the activity sequence are:

- The study of the magnetic force on electrons, emitted by a cathode, moving through a homogeneous magnetic field.
- The experiment made in 1932 by C. Anderson, while studying cloud chamber tracks left by cosmic rays, that allowed the "discovery" of the positron.
- The pion-muon-electron (pi-mu-e) decay chain resulting from antiproton annihilation $\pi \rightarrow \mu \rightarrow$ e, observed at Experiment PS 179 at CERN, Geneva in 1983.

The activity sequence has been planned for students in high school or in introductory physics courses and tested with about 100 high school students. The students worked in groups of four and completed the sequence in two sessions of three hours. They used tutorials and worksheets designed to guide their work and to collect data on their ideas, predictions, experimental results, and interpretations. In the following the activities are briefly described and examples of results obtained by the students are reported.

54.2 The Thomson Experiment and the Lorentz Force on Electrons

The first experiment refers to the study of the magnetic force on electrons moving through a homogeneous magnetic field. At this purpose an apparatus designed to reproduce Thomson's experiment for measuring the electron charge-to-mass ratio is used.[1]

Actually with this kind of equipment the magnetic force

$$\vec{F} = \frac{q}{m}\vec{p} \times \vec{B}, \tag{54.1}$$

was first measured by J.J. Thompson in 1897.

In order to do a quantitative analysis students acquire, by means of a digital CCD camera (see Fig. 54.1), the images of the circular electron orbits and then use the photos to find the quantitative relationship between force, cyclotron radius and electron momentum. A sequence obtained by varying the magnetic field and the

[1] In particular, we used the PASCO e/m apparatus available in the students' laboratory of our Physics Department.

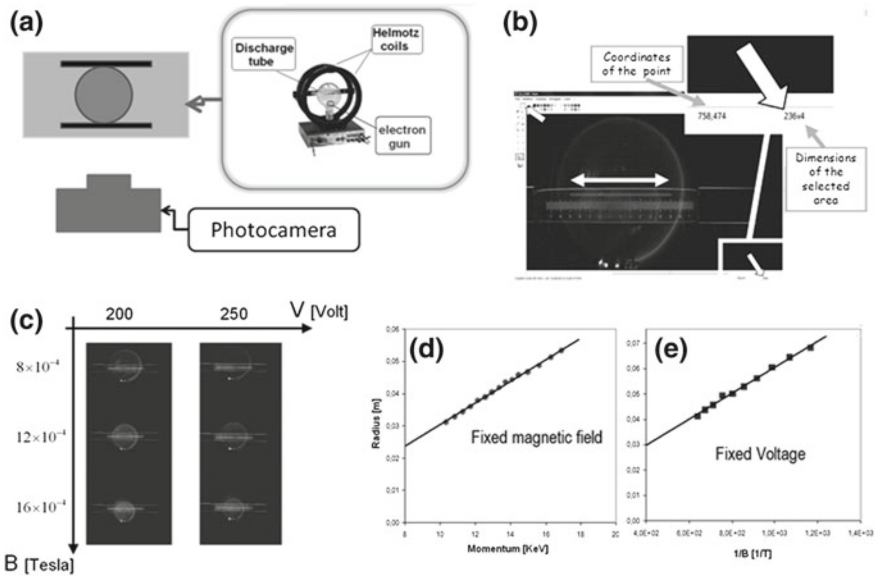


Fig. 54.1 **a** The experimental apparatus for measuring the Lorentz force. A discharge tube is placed within the uniform magnetic field produced by two Helmholtz coils. Photos are acquired by the Web Cam. **b** Students can find the coordinates of any point of the electron track by moving the cursor in the picture. They can measure straightforwardly the diameter of the circular electron track, or use a formula to compute the radius of a circle going through three points. **c** Photos acquired at different values of the accelerating voltages (200 and 250 V) and of the magnetic field (8–16 Gauss). **d** and **e** Experimental results show a linear relation between the momentum of the particle charge and the radius of the trajectory and an inverse relationship of proportionality between the magnetic field and the radius

accelerating voltage is reported in Fig. 54.1c. The momentum, p, of the electrons can be obtained from the electric potential V applied between anode and cathode and expressed in keV/c (c is the velocity of light in a vacuum) by using the formula $cp = \sqrt{m_e c^2}\sqrt{2eV}$. The photo analysis develops in the following steps:

- From measurements of the radius R at fixed magnetic field, and with increasing accelerating voltage students obtain that the radius of the cyclotron orbit depends on the momentum of the electron according to a direct proportionality (Fig. 54.1d) $r \propto p \Leftrightarrow r = \eta(B) \cdot p$.
- From the measurements of R at fixed accelerating voltage and with increasing magnetic field students find that the radius of the circular trajectory and the magnetic field are inversely proportional (Fig. 54.1e), $r \propto 1B \Leftrightarrow r = \frac{\beta(p)}{B}$. Then $r = \lambda \frac{p}{B}$, where λ does not depend on p, nor on B. Since $\eta(B) \cdot p = \beta(p)/B = \lambda \frac{p}{B}$ we can write$\lambda = \eta(B)B$. The value of λ can be determined by finding $\eta$for a given value of B.

- From data in Fig. 54.3 we can find, for $B = 1,01 \cdot 10^{-3}T$, $\eta \approx 0.337c\left[\frac{m \cdot T}{KeV}\right]$. By converting in SI units we obtain $\eta \approx 0.337 \cdot 3 \cdot 10^{8}\left[\frac{10^{-2} \cdot T}{10^{3}e}\right] = \frac{0.11}{e}10^{3}\left[\frac{T}{C}\right]$, then: $\lambda = \eta(B)B = \frac{1.02}{e} \approx \frac{1}{e}$, and $r = \frac{p}{eB}$. The value of r is connected to the centripetal force needed to produce the observed circular motion: $|\vec{F}_c| = p^2 m_e r$, then $|\vec{F}_c| = \frac{e}{m_e}pB$.
- Finally, the direction of the magnetic force is considered. Students are guided to recognize that the force on the moving charges results always perpendicular both to the field and to the velocity (momentum). Data collected in testing this activity with high school students show that most of the student groups obtained graphs as the ones reported in Fig. 54.1d and e and were motivated to find the relation among $|\vec{F}_L|$, $|\vec{B}|$, and $|\vec{v}|$, following the suggestions included in the worksheets.

The experience acquired by the students in working on digital photos, taken while observing directly a phenomenon, was helpful to prepare them to the following activities, based on the analysis of sub-nuclear particle tracks in cloud and streamer chambers as they appear in images taken by researchers in their original experiments. In these experiments the momentum is generally expressed in MeV/c, then it is useful to transform previous results in a compact formula with conventional units. We obtain: $r[cm] \approx 0.33\frac{p[MeV/c]}{B[T]Z}$, then $p[MeV/c] \approx 300r[m]B[T]Z$, where r is the radius of the orbits measured in cm, p is the momentum in MeV/c, B is the magnetic field given in Tesla and Z is the charge in electron units. This formula, which holds also in relativistic mechanics, allows students evaluate the momentum of a particle from the measurement of the radius of the particle track.

54.3 Particle Tracks in Wilson Chamber: The Anderson Experiment

Students analyze particle tracks in cloud chambers by using images of real historical experiments and retrace Anderson's experiment of 1932 that led to the observation of the "positron", the first evidence of antimatter [2]. At first they observe the tracks of different known particles in a cloud chamber (Fig. 54.2a) and answer to some questions about the masses and the charges of the particles "photographed" in the figure. Then they analyze Anderson's photo (Fig. 54.2b) and are guided by worksheets to follow the procedure devised by Anderson to discover the charge of the unknown particle. Students measure the radii of the particle tracks in pixels and transform it in centimeters once known the width of the lead plate. Typically students measured a larger radius of 13.7 cm and a shorter of 5.5 cm in an external field of 15,000 Gauss, corresponding to an electron momentum of 62 and 25 MeV/c respectively. They calculated that the particle loses about 37 MeV/c of its momentum when it passes through the lead plate (the fitting circumferences are plotted in Fig. 54.2c). So they argued that the charged particle is coming from the bottom of the picture, where the curvature radius is larger, thus the velocity higher. Once known the direction of the magnetic field they were able to conclude that the tracks were left by a positive charged particle, the *positron*.

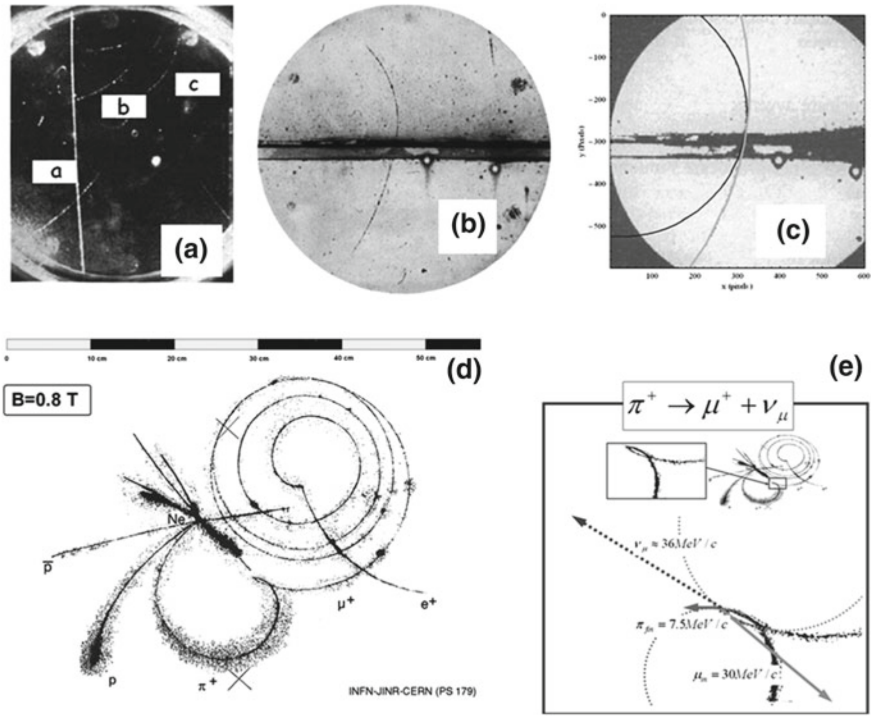


Fig. 54.2 **a** The density of ionization is consistent with the assumption that the track (*a*) corresponds to a proton [1]. **b** Cloud chamber photograph by Anderson of the first positron ever identified. A lead plate, 6 mm thick, separates the upper half of the chamber from the lower half. A positron passes through the lead plate and emerges as a electron. **c** Circumferences drawn by students to fit the particle tracks. **d** The streamer chamber photo of a $\pi \rightarrow \mu \rightarrow$ e decay chain resulting from an antiproton annihilation. At each decay the tracks change direction sharply, indicating simultaneous emission of an unseen neutrino. **e** Analysis carried out by a group of students, who measured the linear momentum of each particle. They added the momenta (*dashed arrows*) of unseen neutral particles

54.4 The Pion-Muon-Electron Decay and the Invisible neutrinos

In the last experiment students examine a streamer chamber photo of a pion-muon-electron (pi-mu-e) decay chain resulting from antiproton annihilation [5]. The image (Fig. 54.2d), recorded in 1983, shows the annihilation of antimatter: an antiproton ($\bar{p}$) produced in the accelerator at CERN, annihilates in the collision with a nucleus of neon (Ne). This produces a pion (π^+) slow moving along a spiral before decaying into a muon (μ^+), and a neutrino(ν_μ) which cannot be observed as a track as it has no electric charge. The students are guided to measure the linear momentum of each particle focusing on the decay events where the tracks change direction sharply. They measure the radius of curvature of the last part of the pion track and obtain the value of the momentum of this particle. In the same way they obtain the initial muon momentum. Using these results students can plot the vectors corresponding to the pion and muon momenta respectively and verify that the total momenta after the pi-mu decay and before the reaction are quite different both in the magnitude and in the direction (Fig. 54.2d). They are guided to interpret this result as the production of new neutral particles.

54.5 Results and final remarks

At the end of the work we asked the students answer some questions on the direction of the force experienced by a charged particle moving through a magnetic field, drawn from the literature [1]. Two different representations of the field are used: by magnetic poles and by magnetic field lines (Fig. 54.3a). Results reported in Fig. 54.3 refer to one of the questions used. They show that the work on the pictures improved high school students' understanding about the direction of the magnetic force and the experience acquired step by step in doing the activities also reduced the frequency of sign errors. We think that image analysis of particle tracks carried out by the students in small groups helped them in focusing on main characteristics of magnetic force (direction and dependence both on the particle momentum and on the magnetic field). Students experimented how the knowledge acquired allowed them to make the same type of analysis and predictions made by scientists in the original experiments and this awareness seemed not only to enhance their interest in the topic, but also to favor their understanding.

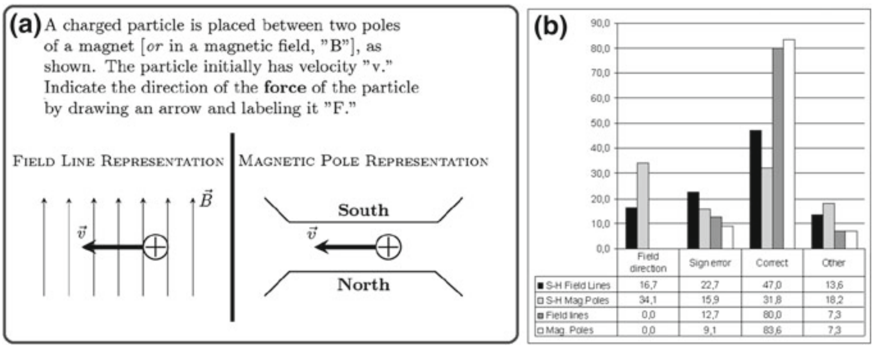


Fig. 54.3 Text and illustrations for the pole and field representations questions used in the study described in Ref. [1] and answered by our students. **b** Proportions of correct, sign error, and field direction answers in the study described by Scaife and Heckler's for the pole and field representations after instruction on magnetic force (S–H) compared with the answers given by our students after the activity sequence

References

1. Anderson CD, Neddermeyer SH (1936) Cloud chamber observations of cosmic rays at 4300 meters elevation and near sea-level. Phys Rev 50:263
2. Anderson CD (1933) The positive electron. Phys Rev 43(6):491–494
3. Galili I (1995) Mechanics background influences students conceptions in electromagnetism. Int J Sci Educ 17:371–387
4. Guisasola J, Almudi JM, Zubimendi JL (2004) Difficulties in learning the introductory magnetic field theory in the first years of university. Sci Educ 88:443–464
5. Piragino G et al (1985) Experiment PS179 at CERN. Nucl Instr Methods Phys Res A234:30–41
6. Sağlam M, Millar R (2006) Upper high school students understanding of electromagnetism. Int J Sci Educ 28:543–566
7. Scaife TM, Heckler AF (2010) Student understanding of the direction of the magnetic force on a charged particle. Am J Phys 78:869
8. Scaife TM, Heckler AF (2007) The effect of field representation on student responses to magnetic force questions. AIP Conf Proc 951:180

Chapter 55
Active Learning by Innovation in Teaching (Alit)

Dina Izadi and Marina Milner-Bolotin

Abstract Today more than ever before, the future depends on students' ability to apply the knowledge they learn in the classroom to solve real life problems such as global warming, climate change, air pollution, waste disposal, energy generation, world poverty and food production. In the incessantly changing world, students of the twenty-first century are very different from the students of the past. This requires educators to think continuously about how to change their teaching to empower and engage modern students, which makes educational innovations imminent. Contemporary students must be proactive in seeking relevant information and applying it to solve real life problems. However, the way we teach hasn't changed sufficiently to reflect these changes. Like in the earlier centuries, the dominant pedagogy in many contemporary science classrooms is still teacher-centered instruction, relying on route memorization and passive learning. To help science educators make a transition from passive to active learning in order to engage students in meaningful learning process, "Active Learning by Innovation in Teaching" (ALIT) model is introduced. This model offers a way of finding different approaches to engage students in meaningful science learning and apply their knowledge to solve real life problems.

55.1 Introduction

In order to be able to solve the problems faced by modern societies, the students should learn how to apply science they learn in the classroom to the world around them. Unfortunately passive learning strategies and the testing systems utilized by many science teachers, create a gap between the "classroom science" and the "real

D. Izadi (✉)
Ariaian Young Innovative Minds Institute (AYIMI), Tehran, Iran

M. Milner-Bolotin
Faculty of Education, University of British Columbia, Vancouver, Canada

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_55,

world science" that has a potential to affect modern society. In countries such as Iran students who passed the entrance exam (Konkoor) with higher scores are more likely to be accepted to universities. However, most of the times the majors these students choose do not entirely reflect their interests. Often, students pass science courses and study science topics in the syllabus without acquiring any deep and meaningful knowledge of the subject. This lack of conceptual understanding is also reflected in how students' success is measured: via end of semester grades obtained by solving plug-and-chug problems on science exams. An important aim of modern education should be helping students acquire and develop science knowledge, skills and abilities that they can apply to solve real life challenges. Research shows that students also need to develop communication and collaboration skills. This can be achieved via encouraging them to engage in active and creative learning in as many ways as possible [19]. A key emphasis in active learning pedagogy is placed on the combination of the theoretical knowledge with experimental—practical skills, which are crucial for motivating students. Active learning environment also requires a cooperative and collaborative atmosphere, where students are encouraged to ask questions and work together to seek answers. One of the challenges active learning pedagogy aims to address is helping students see and critically examine the environment around them, think about the phenomena they encounter in the real life as scientists, and solve the problems faced by the society using novel approaches.

A growing number of modern high school and college teachers have already realized the importance of moving from a passive to an active learning environment in order to motivate students via engaging them in a meaningful learning process. However, many teachers feel a need for support in implementing active learning pedagogies in the classrooms: designing and implementing activities to be used inside and outside of the classroom, as well as choosing pedagogically effective activities relevant to their science curriculum from the plethora of activities produced by other educators [9–11, 18, 19, 21].

55.2 Active Learning by Innovation in Teaching Model to Develop Student-Centered Education

In a teacher-centered learning environment, such as the traditional instruction, the teacher's primary functions are lecturing, designing assignments and tests, and grading. Some instructors use short quizzes at the beginning of every period for this purpose; others who don't want to spend a lot of class time administering and grading quizzes prefer to include questions on the readings assignments in their regularly scheduled examinations. In the student-centered learning environment, the students must take responsibility for their own learning [14, 22]. The students must identify what they need to learn to have a better understanding of the problem, and determine where to get the necessary information (books, magazines, teachers, other students, the internet, etc.). This approach forces students to become active learners, to take ownership of their learning and to work cooperatively. This process also increases students' motivation to learn, improves their retention of material, promotes deep conceptual understanding, and encourages more positive attitudes toward the subject [4, 5, 16]. In a teacher-centered learning environment, instructors often have implicit expectations about what students should learn and how they should learn it [14]. At the first glance, this approach seems to be a safer one. Yet it is less efficient, if we want to encourage our students to acquire critical thinking skills and develop positive attitudes about science. Sometimes, however, the students who have been accustomed to traditional pedagogies may resist student-centered pedagogies and collaborative learning approaches. They might prefer to learn individually without cooperating with other students. These students like to be more distinguished among other students and might not see how they will benefit from collaboration with peers. How can we encourage more students to become active learners instead of sitting in class, listening to their instructors, transcribing and memorizing meaningless formulae in order to get a passing score on the final exam? As educators we should consider how to help students become independent learners while applying their knowledge to solve real life problems, such as global warming, generation of renewable energy, reduction of carbon dioxide emissions.

Ariaian Young Innovative Minds Institute (AYIMI) is a scientific institute located in Iran. It was founded in 2009 (http://www.ayimi.org). It aims at improving science learning via providing opportunities for the students to apply science to solve real life problems. Its other goal is to investigate students' views about the nature of science and about what it takes to learn science [6, 12, 13]. Assessment of student learning is completely different in AYIMI activities. To make a transition from passive to active learning, and engage students in the learning process, (ALIT) model is introduced (Fig. 55.1). This model offers students different approaches to solving problems and to investigating suggested science topics. AYIMI follows the ALIT model for science teaching. In this model, the students learn to use science to improve their society and

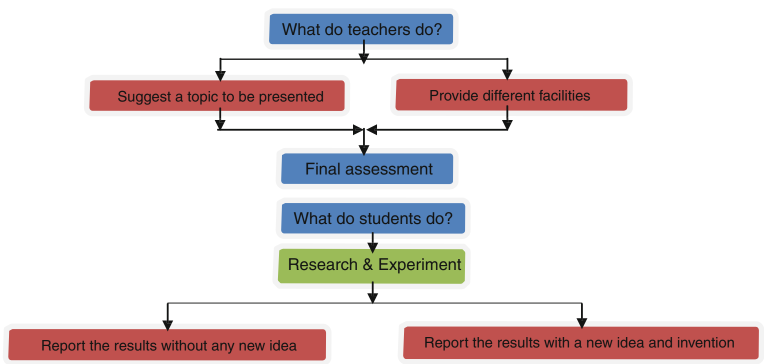


Fig. 55.1 Active learning by innovation in teaching (ALIT) model

solve some of its most important problems. ALIT pedagogy encourages students to work cooperatively to achieve common goals.

55.2.1 What Do Teachers Do?

55.2.1.1 Suggest a Topic to be Presented

A science teacher discusses with the students different topics relevant to the curriculum and relevant to the students that might require further investigation. The class can brainstorm possible problems that the students will investigate and later present to the class. Alternatively, a teacher might suggest a topic for investigation in class and then ask the students to think about possible experiments that they can perform to investigate this topic. These experiments might be conducted by the students working in groups outside of class, or a teacher might support the students by providing them with necessary facilities at school. To help the students build on their prior knowledge and motivate them to explore the topic of interest to them, a teacher might ask the students to think about the topic and write down what they already know on this subject. Then the students discuss their notes with peers and research related subjects to this topic in scientific books or on the internet. Teachers can devise a method of approaching the subject from the students' point of view. It is important to consider what was taught in previous classes, how the topic might be of interest to the students, how the topic of investigation might link to other school subjects. Then the teacher invites the students to discuss what they need to study and perform their experiment in front of the entire class.

55.2.1.2 Provide Different Facilities

According to the topic and what the students need to do to conduct an experiment, a teacher might provide them with different facilities and resources. These resources are used in order to conduct the preliminary experiments in class. However, students can design a more complicated experimental apparatus by themselves in order to obtain more advanced results. In a laboratory, the students work in groups to design their own experimental setup and collect the best possible results to address the problem under investigation.

55.2.1.3 Final Assessment

In traditional learning environments, the focus is most often placed on summative assessment: quizzes, exams, final papers [1]. For many students, this feedback is not useful in terms of their learning outcomes. A formative, ongoing assessment has been proven to be more effective in helping students learn science [2, 7, 8, 17, 20].

In ALIT model teachers can provide continuous ongoing feedback on students' work: evaluate the groups' interpretations, their results, their experimental setup, etc. At the capstone event of the ALIT model, different groups present their projects in a competition. Students and teachers also comprise the jury that judges the projects. Students' performance, the quality of their team work, the performance of project-related practical tasks, introduction of novel approaches in solving problems which can help them in real life, as well as students' confidence in project presentation are the most important parameters that guide the evaluation process.

The final assessment has the following components:

- Individual responsibility of team members
- Quality of student collaboration
- Quality of face-to-face interactions
- The level of students' self- confidence
- Quality and quantity of student learning
- The approach in finding the best solution
- The quality of the design of a method or apparatus related to the topic.

55.2.2 What Do Students Do?

55.2.2.1 Research and Experiment

Unlike the traditional teacher-centered learning environments, in the proposed model, teacher's primary role is to direct student thinking and help them integrated their prior knowledge, find useful resources via conducting independent research and do independent investigation. This model is akin to Project-Based Instruction discussed elsewhere [3, 15]. The students take notes and build models which are consistent with their preliminary findings. They brainstorm ideas in their groups to define the research focus—the problem they will investigate. The next step is conducting experiments which help students integrate science concepts and apply them to solve a specific problem. For example, a teacher asked the students to design an experiment that involved a ball. One of the teams decided to find parameters that help describe the collisions of this ball with different objects made of different materials. In physics, elasticity is defined as an ability of a material to return to its original shape after the stress (e.g. external forces) that made it deform has been removed. The weft and warp of different fabrics have been measured and with a high speed camera the falling objects (different balls) which heating targets (different fabrics) have been clearly analyzed. Then in laminar flow the relation between resilience coefficients and weft and warp of different fabrics have been analyzed with MATLAB (www.mathworks.com) software. Other students built a gun which is made from a solenoid and a ball and found the optimum velocity of this electromagnetic gun. Another group suggested a model to use tidal energy, as a renewable energy, to generate electricity by making the tidal water enter a specially designed cylinder with a piston. In these experiments the students have an opportunity to show their creativity. Engaging in a long-term project with a team of 3–6 students allows them to apply scientific concepts more thoroughly while solving a meaningful science problem under the guidance of a teacher.

55.2.2.2 Report the Results

The results from students' independent or group investigation are reported in class. Some of the results will show known relations between different parameters. However, sometimes the experiment will allow the students to unveil new previously unknown relations. Students are invited to take part in a competition with other groups and defend their projects. The groups who have worked in depth and are able to defend their findings and illustrate the research process in depth will win the competition.

55.3 Students' Role in ALIT

The students' role in ALIT can be described as follows:

- Forming a group in which each member takes specific responsibility
- Carrying out the research projects in each group
- Talking in class and producing a first stage report
- Designing a model to explain the observations
- Carrying out different experiments and explaining observations
- Collecting data
- Analyzing data and sometimes comparing these data to a simulation
- Presenting the results of the investigations in front of other teams and jury to challenge their topics with different team members
- Writing a final report.

55.4 Conclusions

To help students grow as scientists, teachers can suggest different project topics and encourage students to design experimental setup, to conduct the experiments on the topic and to find the best approaches and solutions to the problems. To support this process, the students are encouraged to collaborate with their peers inside and outside of class. (ALIT) model, will persuade students to act as a teacher in a way to help others understand science better. Cognitive research shows that memorizing information does not promote meaningful learning. However, when this knowledge is used to solve a practical problem relevant to students' lives, it becomes more meaningful. In conclusion, in ALIT model the students are encouraged:

1. To solve more practical problems and find solutions using a combination of theoretical and experimental approaches.
2. To find new methods in solving problems. This makes the students think more critically while trying to invent novel solutions to real problems. It also helps extend their scientific abilities.
3. To combine qualitative and quantitative approaches and deductions, to learn how to evaluate their data statistically and compare theoretical predictions with the collected data.
4. To report the results in front of audiences in competitions. This positively affects students' communication skills and their self-confidence about science learning.
5. To acquire collaborative and leadership skills while working in a group to solve a science problem.

In conclusion, from our experience, the ALIT model helps promote conceptual science learning, excite students about science and allow them to apply what they learn in the classroom to solve real life problems.

References

1. Angelo TK, Cross KP (1993) Classroom assessment techniques: a handbook for college teachers, 2nd edn. Jossey-Bass Publishers, San Francisco
2. Beatty ID, Feldman A, Leonard WJ, William J, Gerace, Cyr KS, Lee H et al (2008) Teacher learning of technology-enhanced formative assessment. Paper presented at the National Association for Research in Science Teaching. http://arxiv.org/ftp/arxiv/papers/0806/0806.4924.pdf
3. Blumenfeld PC, Soloway E, Max RW, Krajcik J, Guzdial M, Palincsar A (1991) Motivating project-based learning: sustaining the doing, supporting the learning. Educ Psychol 26(3,4):369–398
4. Bonwell CC, Eison JA (1991) Active learning: Creating excitement in the classroom (ASHE-ERIC Higher Education Report No1 No. Report One). The George Washington University: Washington DC , School of Education and Human Development
5. Bonwell CC, Sutherland TE (1996) The active learning continuum: choosing activities to engage students in the classrooms. In: Menges RJ, Svinicki MD (eds) Using active learning in college classes: a range of options for faculty, vol 67. Jossey-Bass Publishers, San Francisco, pp 3–15
6. Buffler A, Lubben F, Ibrahim B (2009) The relationship between students' views of the nature of science and their views of the nature of scientific measurement. Int J Sci Educ 31(9):1137–1156
7. Etkina E (2000) Weekly reports: a two-way feedback tool. Sci Educ 84(5):594–605
8. Feldman A, Capobianco BM (2008) Teacher learning of technology enhanced formative assessment. J Sci Educ Technol 17:82–89
9. Freeman S, O'Connor E, Parks JW, Cunningham M, Hurley D, Haak D et al (2007) Prescribed active learning increases performance in introductory biology. Life Sci Educ, 6(Summer), 132–139.
10. Hoffman C, Goodwin S (2006) A clicker for your thoughts: technology for active learning. New Libr World 107(1228/1229):422–433

11. Kalman C, Antimirova T, Milner-Bolotin M (2009). Conceptual conflict collaborative group and peer instruction. Paper presented at the canadian association of physicists congress 2009 comparison of two active learning teaching methods
12. Kalman CS (2009) Reading the book of nature: the hermeneutical circle in science. In: Gueldry M (ed) Languages mean business: integrating languages and cultures in/for the professions. The Mellen Press, Wales
13. Kalman CS, Milner-Bolotin M, Antimirova T, Aulls M, Charles ES, Huang X et al (2012) Understanding the nature of science and nonscientific modes of thinking in gateway science courses paper presented at the national association for research in science teaching annual conference
14. Kember D, Gow L (1994) Orientations to teaching and their effects on the quality of student learning. J Highr Educ 65(1):58–74
15. Krajcik J, Blumenfeld PC, Max RW, Bass KM, Fredricks J, Soloway E (1998). Inquiry in project- based science classrooms: initial attempts by middle school students. The J Learn Sci, 7(3,4); 313–350.
16. McKeachie WJ(1994) Teaching tips: strategies, research, and theory for college and university teachers Vol 1, 9th edn. D.C. Heath and Company: Lexington
17. Mishra P, Koehler MJ, Henriksen D (2011) The Seven Trans-Disciplinary Habits of Mind: Extending the TPACK Framework towards 21st Century Learning. [Reports - Evaluative]. Educ Technol 51(2):22–28
18. Perkins K, Adams W, Dubson M, Finkelstein N, Reid S, Wieman C et al (2006) PhET: interactive simulations for teaching and learning physics. The phys teach, 44(January), 18–23
19. Ramsier RD (2001) A hybrid approach to active learning. Phys Educ 36:124–128
20. Rodriguez AJ (1998) Strategies for counterresistance: toward sociotransformative constructivism and learning to teach science for diversity and for understanding. J Res Sci Teach 35(6):589–622
21. Sokoloff DR, Thornton RK (2004) Interactive lecture demonstrations: active learning in introductory physics. Wiley. New York
22. Voogt J, Tilya F, van den Akker J (2009) Science teacher learning of MBL-supported student-centered science education in the context of secondary education in tanzania. [Reports - Research]. J Sci Educ Technol 18(5):429–438

Chapter 56
Capacitors, Tanks, Springs and the Like: A Multimedia Tutorial

Assunta Bonanno, Michele Camarca and Peppino Sapia

Abstract In this work we present an interactive multimedia tutorial allowing to comparatively explore the dynamical behavior of the "Two-capacitor system" and of others systems, quite different, but showing a very similar energetic behavior: (i) two communicating tanks; (ii) two coupled lossy springs "sharing elongation"; (iii) plastic collision between two material points; (iv) two coaxial, rotationally coupled, disks spinning around an axis. The aim of the work is to give learners useful insights on the fundamental subject of energy transformation. The tutorial, appropriate for high school and first year university students, is implemented in a form that makes it also suitable for classroom use with an Interactive White Board.

56.1 Introduction

There is a well-known textbook problem [5, 6] that every once in a while, since forty years, appears in the literature [1, 2, 4, 7–9]: the so called "Two Capacitor Problem", also present in on-line forums as "The Capacitor Conundrum" [9]. The problem concerns two capacitors, one of which (C_1) charged, the other (C_2) uncharged, which are connected at time t = 0. Simple calculations show that in the process a fraction $C_2/(C_1 + C_2)$ of the initial electrostatic energy disappears, apparently giving rise to a paradox. The behavior of this system asks some questions usually worrying learners (first of all: where does lost electrostatic energy go?). The search for answering these questions allows, from a didactical point of view, to highlight many subtleties about energy transformation processes. In this regard, it is very useful comparing the two capacitors system with others systems, quite different, but showing a very similar energetic behavior: (i) two communicating tanks containing a fluid, initially located in one of the two; (ii) two coupled lossy springs "sharing elongation"; (iii) plastic collision between two material points; (iv) two coaxial, rotationally coupled, disks spinning around an axis.

A. Bonanno (✉) · M. Camarca · P. Sapia
P.E.R. Group, Physics Department, University of Calabria, 87036 Rende, CS, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_56,

In this connection we present here an interactive multimedia tutorial allowing to comparatively explore the dynamical behavior of described systems. The tutorial (implemented in a form that makes it also suitable for classroom use with an Interactive White Board) is appropriate for high school and first year university students and has manifold goals; among others: (i) to give learners useful insights on the fundamental subject of energy transformation; (ii) to shed light on the role of dissipative phenomena in the achievement of equilibrium; (iii) to illustrate the role of analogies in the study of physical phenomena, in the aim to allow students to get used to a proper identification of similarities between apparently different scientific subjects; (iv) to pave the way for further quantitative treatment of the problem.

In the present paper, after summarizing in Sect. 56.2 the two-capacitor-problem family, we will describe, in Sect. 56.3, the tutorial. In Sect. 4, eventually, we draw some interesting conclusion.

56.2 The Two-Capacitor(-like) Problem(s)

A capacitor is charged (Fig. 56.1, left) by a given potential difference (p.d.) V_1; then it is disconnected from the e.m.f. source and connected to an uncharged capacitor, with which it shares its electric charge (Fig. 56.1, right). To determine the final state one must impose the conservation of charge and the requirement of a common final p.d. through the two capacitors. Simple calculations show that the "initial" stored electrostatic energy ($U_i = 1/2CV_1^2$) halves in the "final" equilibrium state ($U_f = 1/4CV_1^2$); so that a definite fraction of the initial energy "disappears". For sake of simplicity we consider the case in which capacitors have equal capacities, so that this fraction is 1/2. More generally, one can easily show that the "disappearing" fraction of the initial energy is given by:

$$\frac{C_2}{C_1 + C_2} \tag{56.1}$$

where the subscripts 1 and 2 refer, respectively, to the initially charged or uncharged capacitor.

Fig. 56.1 The two capacitor problem. A capacitor is initially charged, the other uncharged (*left*). Then capacitors are connected so to share the electric charge (*right*). In this process half the electrostatic energy "disappears"

A very interesting consideration, on a pedagogical point of view, is the following: there are different mechanical systems (apparently dissimilar from each other, and even more so than the system of two capacitors) that show exactly the same kind of energetic behavior (Fig. 56.2). In fact, Simple calculations allow to verify that for each of the systems represented in Fig. 56.2 the energy loss for the attainment of the final state of equilibrium is equal to half the initial energy (Table 56.1).

A closer comparison among described systems permits to highlight the following similarities: (i) they are two-component systems; (ii) their energy depends quadratically on two symmetric intensive variables; (iii) their approach to equilibrium is driven by the tendency to equalize these two variables; (iv) they all need some form of dissipation to reach equilibrium, and lose half the energy in reaching equilibrium. These analogies, and their conceptual implications are explored in a paper we have recently written [3].

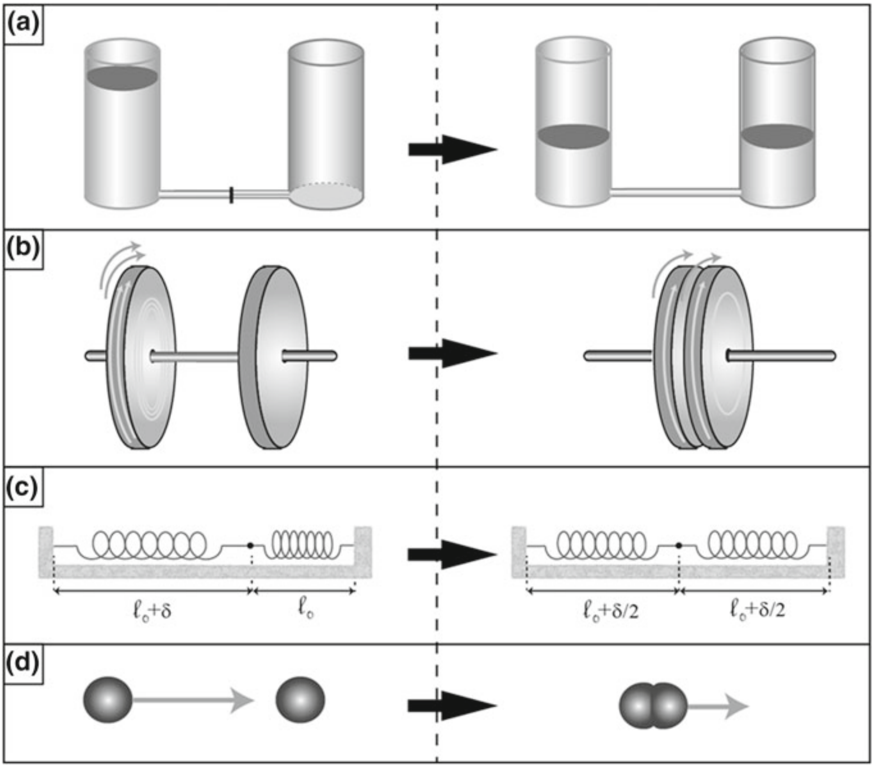
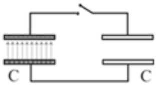
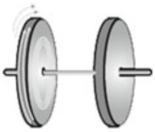


Fig. 56.2 Mechanical systems exhibiting the same behavior as the two capacitor system: **a** Two tanks sharing an amount of liquid; **b** Two coaxial rotating disks sharing angular momentum; **c** Two connected springs sharing elongation; **d** Two plastically colliding masses sharing momentum. In each case the equilibrium is driven by equating an intensive variable: **a** Liquid height; **b** Angular velocity; **c** Elastic force; **d** Velocity

Table 56.1 Two-capacitors and related systems

System	System parameter	System variables (equilibrium condition)	Energy "before"	Lost fraction of initial energy
	Capacitors' capacity (C)	V_1, V_2 : potential difference across capacitors ($V_1 = V_2$)	$\frac{1}{2}CV_1^2$	$\frac{1}{2}$
	Tanks' sectional area (S)	h_1, h_2 : heights of the liquid in the tanks ($h_1 = h_2$)	$\frac{1}{2}\rho g S h_1^2$	$\frac{1}{2}$
	Moment of inertia of the disks (I)	ω_1, ω_2 : Angular velocities of the disks ($\omega_1 = \omega_2$)	$\frac{1}{2}I\omega_1^2$	$\frac{1}{2}$
	Elastic constant of the springs (k)	Springs' elongations ($\delta_1 = \delta_2$)	$\frac{1}{2}k\delta_1^2$	$\frac{1}{2}$
	Masses of colliding material points (m)	Speeds of the material points ($v_1 = v_2$)	$\frac{1}{2}mv_1^2$	$\frac{1}{2}$

This table establishes a formal correspondence among the systems by comparing: characteristic parameters (second column); system variables and related equilibrium condition (third column); expression of the initial energy in terms of system variables (fourth column); fraction of the initial energy lost after equilibrium has been reached

A useful visual tool in order to understand the energetic behavior of two-capacitor-like systems is a graphical representation [9] of the system state in a Cartesian plane whose coordinates are, for each one of the systems in Table 56.1, the corresponding couple of variables in the third column. As an example, Fig. 56.3 shows the representation for the two-tanks case. A little algebra allows to see that the fixed energy[1] states of each system lie on an ellipse in that Cartesian plane; while equilibrium

[1] With reference to Table 56.1, the various kind of energy are, respectively: electrostatic, gravitational potential, rotational kinetic, elastic potential, translational kinetic.

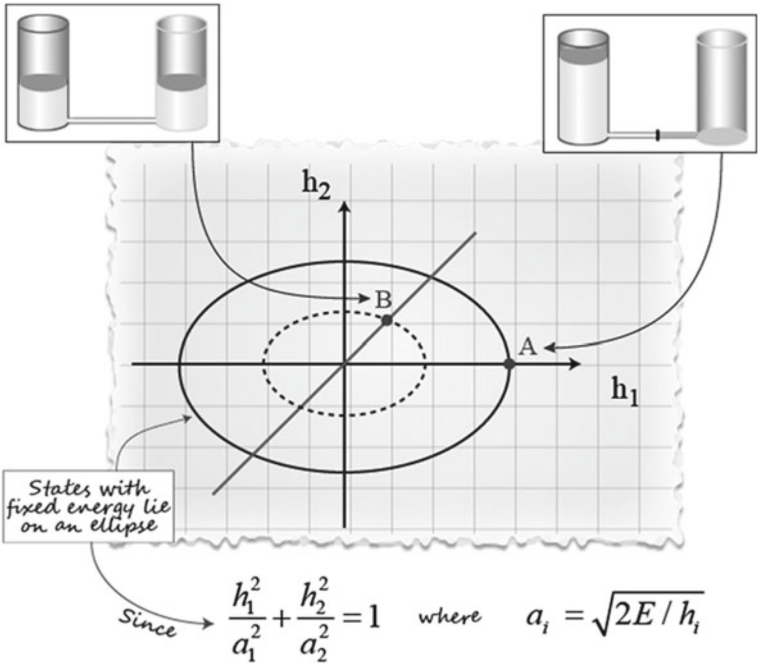


Fig. 56.3 A useful graphical representation of the energetic behavior of the two-tanks system. A given state is represented in a Cartesian plane whose coordinates are the liquid's levels (h_1 and h_2) in the two subsystems. Fixed-energy states lie on an ellipse, while equilibrium states belong to the bisector of the first and third quadrants. The example depicted in the figure refers to an initial state (point A) in which the liquid is entirely contained in the right tank, and a final (equilibrium) state in which liquid's heights in the two subsystems are equal

states between the two subsystems are represented by points belonging to the bisector of the first and third quadrants. Reaching the equilibrium state, starting from an imbalance between the two subsystems, requires a "jump" from an outer ellipse to the intersection point between the bisector and an inner ellipse: the "jump" requires the dissipation of a given amount of the system's energy: this amount is always one half, in terms of the energy corresponding to the starting ellipse.

56.3 The Multimedia Tutorial

The interactive tutorial, implemented as a web-browsable hypertext, is intended to introduce in a visually appealing way the two-capacitor-like systems, in the aim to discuss the subtleties regarding the transfer of energy between systems. In particular it helps shedding light on the role of dissipative phenomena in the achievement of equilibrium, suitably for high school (and first year university) students. On a

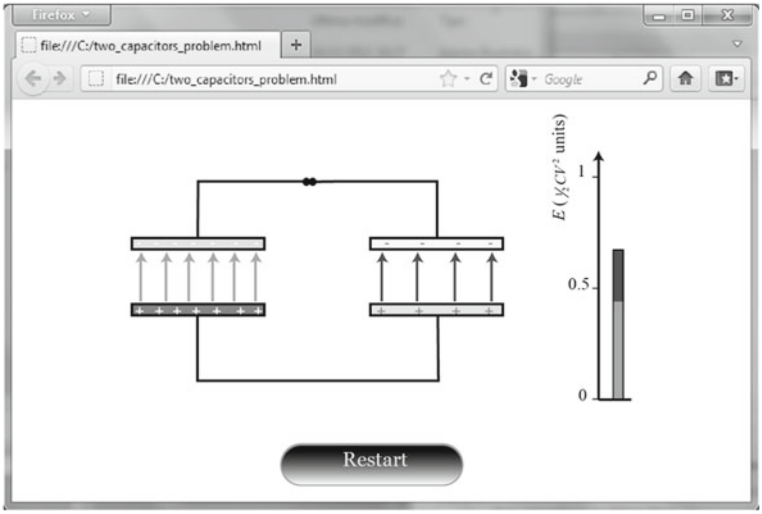


Fig. 56.4 A screenshot from the tutorial, illustrating by means of an animation the distribution of the electrostatic energy between the two subsystems (single capacitors) as the whole system approaches equilibrium

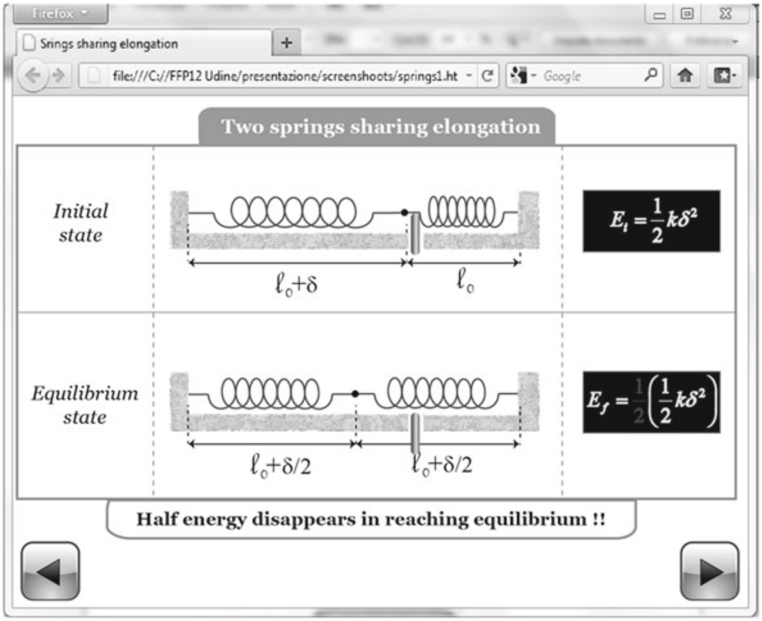


Fig. 56.5 A screenshot from the tutorial, illustrating the mandated energy loss in a system constituted by two springs sharing elongation

more strictly pedagogical point of view, the proposed multimedia permits learners to acknowledge the same role of different variables in unlike systems, so to enlighten the power of analogies in the study of physical phenomena. Finally, it paves the way for further quantitative treatment of the problem, examples of which can be found in [3]. Figures 56.4 and 56.5 show some screenshot of the tutorial, which can be freely accessed on the web at the following URL: www.fis.unical.it/didattica/FFP12/capacitors.

References

1. Al-Jaber SM, Salih SK (2000) Energy consideration in the two-capacitor problem. Eur. J. Phys. 21:341–345
2. Boykin TB, Hite D, Singh N (2002) The two-capacitor problem with radiation. Am. J. Phys. 70(4):415–420
3. Bonanno A, Camarca M, Sapia P (2012) Reaching the equilibrium: the role of dissipation in analogous systems. Eur. J. Phys, Submitted
4. Cuvaj C (1968) On conservation of energy in electric circuits. Am. J. Phys. 36:909–910
5. Halliday D, Resnick R, Walker J (1997) Fundamentals of physics, 5th edn. Wiley, New York
6. Lorrain P, Corson D, Lorrain F (1988) Electromagnetic fields and waves, 3rd edn. Freeman, New York, p 117
7. Mita K, Boufaida M (1999) Ideal capacitor circuits and energy conservation. Am. J. Phys. 67(3):737–739
8. Powell RA (1979) Two-capacitor problem: a more realistic view. Am. J. Phys. 47(5):460–462
9. Sciamanda RJ (1996) Mandated energy dissipation- e pluribus unum. Am. J. Phys. 64(10):1291–1295

Chapter 57
Energy Exchange by Thermal Radiation: Hints and Suggestions for an Inquiry Based Lab Approach

Onofrio Rosario Battaglia, Claudio Fazio, Nicola Pizzolato and Rosa Maria Sperandeo Mineo

Abstract In this paper we present some laboratory activities developed in the framework of an inquiry-based approach to the study of energy exchange by thermal radiation. These activities were developed in the context of "Establish", a FP7 European Project aimed at promoting and developing Inquiry Based Science Education in European Secondary Schools. By starting from real life students are engaged in designing and carrying out laboratory activities by collecting, processing and analysing data. Particular attention is paid in building data interpretation by taking into account the effects of parameters like the environmental temperature.

57.1 Introduction

Several researches in science education support the view that teaching science as inquiry enables students to obtain an experience that is similar to that of scientists, making their learning more meaningful and improving their scientific understanding [1, 2]. By inquiry we refer to learning experiences that engage students in various integrated activities of identifying questions, collecting and interpreting evidence, formulating explanations, and communicating their findings, that are consistent with science standards and recent reports [3, 4]. In the majority of the developed Inquiry Based (IB) approaches, a relevant role is played by laboratory activities where students are directly engaged in finding answers to their questions, developed in appropriate contexts. On the contrary, in traditional content-to-context based physics education, the focus of laboratory activities developed by students is mostly dedicated on verifying information previously transferred by the teacher.

It is well acknowledged that science education has to make students aware that science is not just a body of knowledge that reflects current understanding of

O. R. Battaglia (✉) · C. Fazio · N. Pizzolato · R. M. S. Mineo
Dipartimento di Fisica, University of Palermo, Palermo, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_57,

the world; it is also a set of practices used to establish, extend, and refine that knowledge. Both elements, knowledge and practice, are essential [5].

In all IB approaches to science teaching, students are themselves engaged in the practices and do not merely learn about them second-hand.

In this view, we are developing a teaching-learning sequence regarding the investigation of energy exchange by thermal radiation, conduction and convection, within the context of ESTABLISH (http://www.establish-fp7.eu), a FP7 European Project aimed at promoting and developing IBSE in European secondary schools. Our teaching-learning sequence is organized by firstly engaging students by means of scientifically oriented questions about real life situations, such as thermal insulation of houses and the use of energy saving materials. Students are invited to design, build, model by themselves a scale model of an energy-efficient house.

As a part of this inquiry-oriented teaching/learning sequence, here we present a laboratory activity aimed at investigating the physics of energy transfer by thermal radiation. A scientific investigation on the energy exchange between a powered resistor and its surrounding environment during the heating and cooling processes is proposed. Students are stimulated to plan and carry out their own laboratory activity by collecting, processing and analysing data, in order to learn new concepts or laws and obtain more meaningful conceptual understanding of the physics underlying the process of energy exchange by thermal radiation. In particular, we report as problems and experiments were developed in a workshop for engineering students attending the second year of their university curriculum.

In the following section, we present the theoretical framework that grounded the organization of the workshop. In Sects. 57.3 and 57.4 we describe an example of inquiry-lab learning path where engineering students are directly involved in experimental activities and in the related investigations aimed at giving theoretical meaning to their experimental results. Finally, we discuss the pedagogical experiment by drawing conclusions about the way to improve its efficacy.

57.2 Theoretical Background

The term "inquiry," has been interpreted over time in many different ways throughout the science education community. For this reason, many new documents better specify what is meant by inquiry in science and the range of cognitive, social, and physical practices that it requires.

The new Framework for K-12 Science Education [5] describes a set of practices for K-12 students derived from those that scientists and engineers actually engage in as part of their work. It is recognized that students cannot reach the level of competence of professional scientists and engineers, as well as that to give students opportunities to immerse themselves in these practices and to explore why they are central to science and engineering is critical to appreciate the skill of the expert and the nature of his or her enterprise.

Eight practices are considered to be essential elements of the K-12 science and engineering curriculum:

1. Asking questions (for science) and defining problems (for engineering);
2. Developing and using models;
3. Planning and carrying out investigations;
4. Analysing and interpreting data;
5. Using mathematics, information and computer technology, and computational thinking;
6. Constructing explanations (for science) and designing solutions (for engineering);
7. Engaging in argument from evidence;
8. Obtaining, evaluating, and communicating information.

These may be, a fortiori, considered relevant for undergraduate students. However, typical pedagogical methods used in introductory physics courses mainly engage students in lectures and the related laboratory experiments can be grouped into two categories:

- A theoretical relationship between two or more variables is already known (or at least suspected) and an experiment is needed to verify or quantify this relationship.
- A theoretical relationship between two or more variables is not available but rather sought through an experiment.

Such kind of activities obscures the inquiry characteristics of scientific investigation and very often students perform experiments by slavishly following the recipes supplied by instructor.

Many papers have pointed out the fairly general process that one can follow to design an experiment under any circumstances [6, 7]. This process can serve as a tool for teaching students experimental design as well as to introduce in it many of the practices previously described [5].

Relevant elements of such a process have been considered the following [6]:

1. Define the goals and objectives of the experiment;
2. Research any relevant theory (or idea) about what to expect from the experiment;
3. Select the dependent and independent variable(s) to be measured;
4. Select appropriate methods for measuring these variables;
5. Choose appropriate equipment and instrumentation;
6. Select the proper range of the independent variable(s);
7. Determine an appropriate number of data points needed for each type of measurement.

An easy solution, of course, is to intend all the information in steps 1 through 7 as a prescription by providing a "cookbook" lab. However, these types of approach should not supply an environment suitable to develop competences in experimental design as well as to engage students in inquiring. Our approach is aimed at defining and characterising practices in parallel at applying them at the solution of well-posed problems. We think that this approach is relevant for making students aware of the involved practical/thinking process in a way that make them able to transfer these in different contexts.

Here, we describe a 4 h Worksop (W) where students were put in a real problematic situation that made them face with a real problem, whose solution required the activation of the typical operative processes (theoretical and experimental) of scientific investigation.

57.3 Method

The W involved 60 engineering students that have completed their two courses of Introductory Physics mainly consisting in lectures and activities devoted to the solution of numerical problems. They participated to the W on a voluntary basis.

The lab activities were introduced through a phase of topic exploration, mainly based on students' prior knowledge, that was followed by a discussion about the relevant physical quantities, which should be taken into account in a contest of a laboratory activity of measurements. From their previous studies, students should know that the energy exchange between an emitting body and its surrounding is caused by the processes of conduction, convection and radiation. By discussing between each other, students should gather that, in the absence of a medium, as in a vacuum system, the conduction and convection cannot take place. The inquiry–lab activity on thermal radiation proceeded as following:

1. A scientific investigation on the energy exchange between a powered resistor and its surrounding environment during the heating and cooling processes was proposed;
2. Students were stimulated to design experimental activities aimed at testing previously formulated hypotheses;
3. Then, they were introduced to the laboratory equipment and invited to adapt their plans about conducting the experiments with the available measurement facility for the study of energy exchange by thermal radiation;
4. Finally, they carried out their own experiments by collecting, processing and analysing data, by focusing on new concepts and laws and constructing models supplying a meaningful conceptual understanding of the physics underlying the process of energy exchange by thermal radiation.

57.4 The Lab Activities

At the end of the first two steps previously described, all students agreed that in order to eliminate effects due to conduction and convection, the cooling experiments have to be conducted in a vacuum environment. Moreover it is necessary to measure the

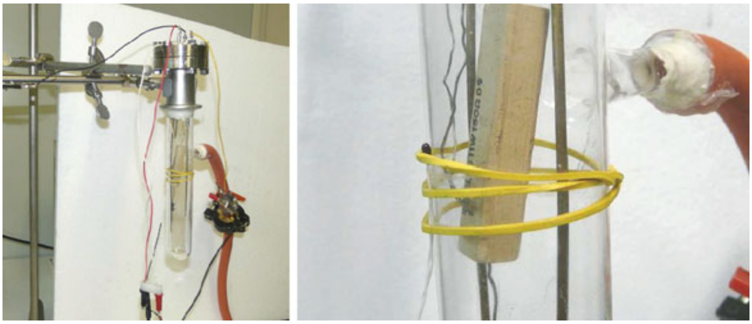


Fig. 57.1 The vacuum tube containing the resistor and details of the ceramic-type resistor and the thermocouple

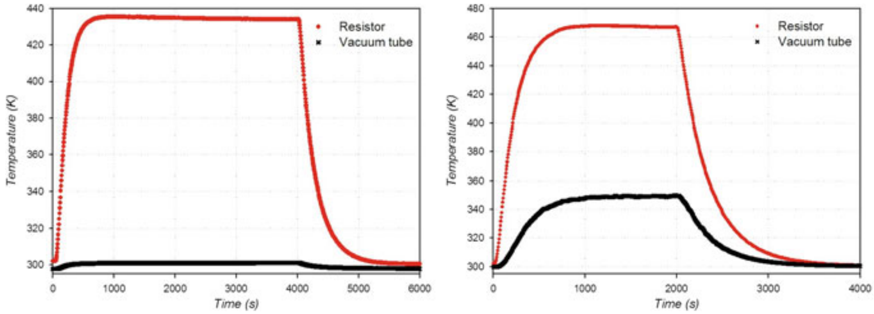


Fig. 57.2 Heating and cooling curves of resistor and glass vacuum tube: on the *left* the temperature of the tube is almost constant; on the *right* the glass tube temperature increases and then decreases as the resistor temperature increases/decreases

variation of the resistor temperature as well as that of the environment (for example a glass tube surrounding the resistor). Figure 57.1 shows the proposed equipment.

The system is composed by a vacuum pump which is connected to the vacuum tube containing the resistor; the pump can lower the pressure inside the tube up to 0.1 mbar. The temperature of the resistor is measured by means of a thermocouple which is placed inside the vacuum tube and whose tip is in contact with the resistor surface. The thermocouple is externally connected to an amplifier module and interfaced to the lab computer. The resistor is a ceramic-type with the following physical characteristics: 11 W, 150 Ω. The temperature of the external surface of the vacuum tube is measured by using a surface temperature sensor connected to a Vernier LabQuest computer interface. A power management system drives the electric current into the resistor.

The first experiment was carried out by the measuring the temperature of the resistor during the heating and cooling process. Figure 57.2 shows two typical results: the first one involving a modest variation of the environment temperature, the second one showing a relevant heating of the environment temperature.

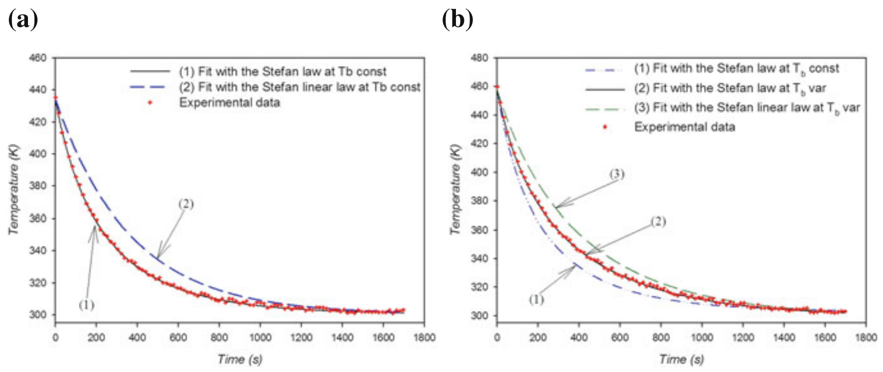


Fig. 57.3 Fitting of cooling experimental data: on the *left* (the case of constant temperature environment), data are well fitted with the Stefan's law (curve (*1*)) and not with its linear approximation; on the *right* (the case with a decreasing temperature environment) curves (*1*) and (*3*) do not fit well experimental data that are well fitted with the Stefan's law when the environment temperature is supposed exponentially decreasing, curve (*2*)

Students decided to investigate the cooling process following the saturation regime, just after the power supply has been switched off. They discussed about the collected data and tried to compare their experimental findings with the theory of thermal emission of radiation expressed by the Stefan law:

$$C\frac{dT}{dt} = -\varepsilon\sigma S(T^4 - T_b^4) \tag{57.1}$$

where T is the temperature of the emitting object, C the heat capacity, ε the emissivity, σ the Stefan constant, S the object surface and T_b the temperature of the surrounding ambient. Both the heat capacity and the emissivity are known. Students realised that, in order to perform such comparison, it was necessary to integrate the Stefan law or to approximate it (as reported in their textbooks) to an exponential law:

$$T_b(t) = (T_{b0} - T_{inf})e^{-t/b} + T_{inf} \tag{57.2}$$

where T_{b0} is the maximum value of ambient temperature, T_{inf} is the equilibrium temperature and b is the time constant.

They tried to change these parameters in order to match the theory with the experimental data, and found that in the case of Fig. 57.3a a satisfactory agreement is not obtained for the exponential approximation, but the agreement is acceptable in the case of numerical integration of Stefan law. For the case of Fig. 57.3b students tried to identify the mean temperature of the glass tube as environment temperature during the cooling, but in both case the agreement is not satisfactory. They deduced that the problem needed further investigations. By analysing their thinking-aloud reasoning we verified that the experiment made students aware that the interpretation of experimental finding needs to put in action a lot of approximations, in particular the need to understand the importance of the boundary conditions. A group of students analysing the cooling of the glass tube and hypothesizing that in this case the effect of conduction/convection can be the most relevant factor tried to fit data by imposing also an exponential decreasing of the glass tube temperature by obtaining the curve (3) (see Fig. 57.3b) that fits well experimental data.

57.5 Discussion and Conclusions

In this paper we present a teaching/learning path based on a set of inquiry-laboratory activities aimed at driving students to the understanding of the physics underlying the process of thermal radiation. Students are, initially, engaged by questions regarding their everyday experience that lend themselves to scientific investigations and they are stimulated to design their own experiments. Such experiments are discussed in the classroom and adapted to the equipment at disposal. The experimental findings are compared with the available theories and students discuss about the results of the comparison. The question is: “Do our experimental data meet the theory?”. If the answer is “yes”, everything is fine. If the answer is “no” (as usual) students should check their experiment and think about physical phenomena that can affect the data. Then, they can turn back to the theory, possibly modifying the previous one, and conduct a new comparison. In this way, students are involved into a real scientific research that can bring up to results unexpected even for the teachers, such as the importance of the temperature of the surrounding environment into the study of the energy exchange by thermal radiation. If the agreement between theory and collected data is still not completely satisfying, as the most usual case, students could argue that maybe they used a not perfectly correct value of heat capacity and/or emissivity in the numerical solution of the Stefan law and perform further investigations on how the accordance between data and theory can be improved by changing these parameters.

The results from the W described in this paper seem quite promising. Students showed to gain extensive experience in design of experiments by taking responsibility to define objectives, select appropriate methods to measure their variables, and analysing data.

The experimentation of a simplified version of this inquiry-lab learning path on high school students is under implementation and the results will be the subject of a forthcoming paper.

References

1. Minstrell J, van Zee E (eds) (2000) Inquiring intoInquiry learning and teaching in science. American Association for the Advancement of Science, Washington, DC
2. Windschitl M, Thompson J (2006) Trancending simple forms of school science investigation: the impact of preservice instruction on teachers’ understandings of model-based inquiry. Am

Educ Res J 43:783–835
3. Duschl RA, Heidi A, Schweingruber HA, Shouse AW (eds) (2007) Taking science to school. National Academies Press, Washington, DC
4. Singer SR, Hilton ML, Schweingruber HA (eds) (2005) America's lab report: Investigations in high school science. The National Academies Press, Washington, DC. http://www.nap.edu/books/0309096715/html
5. NAS (2011) A Framework for K-12 science education: practices, crosscutting concepts, and core ideas, The National Academies Press at http://www.nap.edu/catalog.php?record_id=13165
6. Du WY, Furman BJ, Mourtos NJ (2005) On the ability to design engineering experiments. In: 8th UICEE annual conference on engineering education, Kingston, Jamaica, 7–11 Feb 2005
7. Hatzikraniotis E, Kallery M, Molohidis A, Psillos D (2010) Students' design of experiments: an inquiry module on the conduction of heat. Phys Educ 45:335–344

Chapter 58
Investigating Teacher Pedagogical Content Knowledge of Scientific Inquiry

Claudio Fazio, Giovanni Tarantino and Rosa M. Sperandeo Mineo

Abstract This paper analyses the scientific inquiry in the context of modelling physical systems and point out some related teaching/learning strategies. Two relevant points are also discussed: (i) the different representations of scientific inquiry held by a sample of secondary school physics teachers and their competencies in conducting investigations aimed at solving scientific problems; (ii) what kind of supports can be supplied to teachers in order to modify their misconceptions about scientific inquiry and develop adequate competences related to Pedagogical Content Knowledge of Scientific Inquiry. The topic will be analysed in the light of some preliminary results of a course for in-service physics teacher education held at University of Palermo in the framework of Establish, an FP7 EU Project.

58.1 Introduction

In designing innovative pedagogical materials usually supports for teachers' subject matter knowledge and pedagogical content knowledge for students' ideas (e.g., misconceptions), but rarely for pedagogical content knowledge of scientific inquiry, are provided. By inquiry we refer to learning experiences that engage students in various integrated activities of identifying questions, collecting and interpreting evidence, formulating explanations, and communicating their findings, that are consistent with science standards and recent reports [1–3]. Many researchers have shown that Inquiry-Based (IB) approaches to learning are able to increase student motivation, interest, understanding and development [3, 4]. However, despite the consensus found in educational research, teachers may have different ideas about the meaning of Inquiry-Based instruction and it has been shown that misconceptions abound [2]. These mistaken notions about inquiry can, perhaps, deter efforts to reform science education.

C. Fazio (✉) · G. Tarantino · R. M. Sperandeo Mineo
Dipartimento di Fisica, Università di Palermo, Palermo, Italia

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_58,

In this paper, we analyse scientific inquiry in the context of modelling physical systems and point out some related teaching/learning strategies and how these are perceived by physics teachers. The topic will be analysed on the light of some preliminary results of a course for in-service physics teacher education held at University of Palermo in the framework of Establish [5], an FP7 Project aimed at extending the use of IBSE (Inquiry Based Science Education) in second level schools across Europe.

58.2 Theoretical Background

Despite the consensus found in educational research about the efficacy of IB approaches, it has been pointed out that misconceptions about IB instruction abound and serve to deter efforts to reform science education [2]. The more relevant misconception involves the idea that " *IB instruction is the application of the scientific method"* Many teachers learned as students that the process of science can be reduced to a series of five or six simple steps. It has been shown that the notion that scientific inquiry can be reduced to a simple step-by-step procedure is misleading and fails to acknowledge the creativity inherent in the scientific process. Research has connected this view with what science's nature is perceived from samples of pupils and teachers [6, 7]. Some studies showing that science teachers persist in holding views about nature of science qualified as empirical or naïvely empirical, identify two peculiar aspects in their thinking, involving: the role accorded to observation, which is seen as giving experimental data an absolute value; the role played by theory in conducting experiments and in making observations, along with the value of scientific knowledge as a means of explaining and predicting [8–10]. Other studies report that teachers carry positivistic views of their discipline: i.e. they teach only the knowledge aspects of science and emphasize vocabulary rather than balance knowledge claims with knowledge generation and evaluation, and present science as "the method" of understanding the world [11]. Additional classroom consequences may include a decreased emphasis on inquiry-oriented and problem solving teaching methods that positively impact pupils' conceptions of science [7, 11].

Teaching IB science entails ambitious learning goals for students and thus is complex and difficult for teachers to enact [12, 13]. Moreover, most of teachers also have not experienced IB instruction as learners and thus need guidance in enacting this type of instruction [14]. Researches specifically aimed at the implementation of the IB approaches to physics education have shown that teachers aren't able to make the transition from a purely transmitting didactics to an IB one only through the illustration of the new methods and strategies [15]. Training experiences based on new theoretical models have to be provided. Among these, models that underline the necessity of collaborative construction of understanding and reflection on the enactment of new practices in classrooms and on the consequent adaptation of materials and practices show a relevant efficacy. Such procedures require an accurate designing of the training activities where the roles of the different materials, the disciplinary conceptual knots, the problems related to the introduction of the innovative methods are evident. It has been pointed out that a conceptual change approach is not only relevant to teaching in the content areas, but it is also applicable to the professional development of teachers. For example, as constructivist approaches to teaching gain popularity, the role of the teacher changes. Teachers must learn different instructional strategies as well as also re-conceptualize or change their conception about the meaning of teaching.

Founding on research related to the Inquiry Based (IB) methodology and to models of the teachers training, our research takes as theoretical framework the following key points:

(a) to analyze conceptions, reasoning schemes and teachers' knowledge in the light of the specific operative processes of an IB approach and to study how these can enhance or thwart the introduction of innovative strategies and contents;
(b) to develop and to experiment a Training Action (TA) that proposes subjects and strategies focused on the specific operative processes of an IB approach.

58.3 Method

The TAs developed by our local project propose subjects and strategies focused on the specific operative processes of an IB approach, as well as on analysed ways to integrate them in a pedagogy aimed at pointing out relevant elements of a adequate Pedagogical Content Knowledge [16] for SI. Some pilot teacher workshops (W) have been designed aimed at finding operative answers to defined research questions. Here, we report about some preliminary results of a 4 hours W where teachers were put in a real problematic situation that made them face with a real problem, whose solution required the activation of the typical operative processes (theoretical and experimental) of scientific investigation. Our research questions are the following;

1) Which kinds of approaches to a complex problem are preferred by teachers? Which cognitive resources are involved?
2) Which relevant procedures of scientific investigation are put in action and how are these intended by the teachers?

15 in-service secondary school teachers participated to the pilot workshop. They had different backgrounds of graduation, pre-service training and kind of teaching experience in different grade levels (junior or high school level). The W has been developed into two phases: (a) analysis of a real problematic situation and development of a solution; (b) reflection about the proposed solution and definition of the characteristics of the procedures put in action for the searching of the solution. During the first phase a real problematic situation involving the phenomenon of heat conduction through different materials was proposed to teachers (see Fig. 58.1).

In particular, teachers were requested, through a questionnaire, to look at seven different flat plates, different in material or mass, area and thickness and to predict

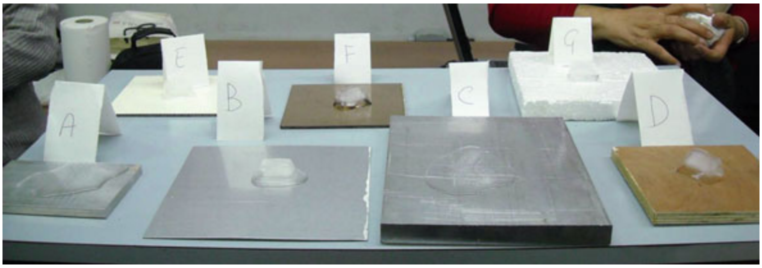


Fig. 58.1 The experimental situation, first only theoretically proposed, then practically analyzed by the teacher group. Some ice cubes are placed on plates of different material, mass, area and thickness, identified by letters A, B, C, D, E, F, G, H

what happens if seven identical ice cubes are placed upon them, i. e. to predict the time sequence of ice cubes melting. As a second question, they were requested to put into evidence the parameters they considered relevant in influencing the melting process, and to design a set of experiments devoted at checking the relevance of such parameters. Then, the experience was really performed and they were required to compare their predictions with the experimental results, writing down their comments.

The questionnaire sheets were photocopied and then returned to teachers for further processing. As a second phase teachers were divided into 3 groups (5 teachers in each group) to discuss their solutions with the guide of a researcher. The 3 researchers had previously developed common discussion guide-lines and questions in order to make explicit teachers conceptions about the meanings of investigation procedures as to make a hypothesis, to construct a descriptive model, to find an explanation or to compare different kinds of explanations. All the discussions have been audio taped and the detailed analysis is in progress.

58.4 Results

A detailed analysis of teacher incorrect predictions with the related explanations is in progress. We note here that all teachers correctly predicted that the ice cubes melt sooner when placed on aluminum plates, but only 3 teachers performed the correct prediction by appropriately ranking plates with different masses and thickness. By considering at a finer grain detail the predicted melting time sequence for the three aluminum plates (of different geometrical characteristics), as well as the teachers' explanations about the parameters that have to be taken into account for the correct description and explanation of the phenomenon, some considerations can be drawn about the procedures they followed in formulating their hypotheses.

The correct analysis of the proposed situation must take into account several parameters, i.e. the plates' geometrical characteristics (surface area and thickness), thermal capacity, thermal conductivity and the temperature difference between the two plate faces. All these parameters have to be considered together to correctly explain the phenomenon. An approach based on the typical theoretical knowledge about thermology, resulting from classical university education can easily guide the teacher to search for an explanation to the phenomenon based on thermal conduction, i.e. on the idea of heath flux between two surfaces at different temperatures. From the Fourier law, we know that this flux is proportional to the surface area and inversely proportional to the plate thickness, so thinner plates should make ice cubes melt quicker, as they allow a bigger heath flow. Indeed, for a clear understanding of the phenomenon other factors involved in the analysis of thermal interaction between two bodies (the ice cube and the plate that exchange thermal energy until equilibrium is reached) must be taken into account. This second kind of approach actually produces opposite predictions with respect to the previous ones, i.e. a greater melting speed of ice cubes placed on the thicker (and so heavier) plate. Our data show that 6 teachers appear to describe the phenomenon only on the basis of the Fourier law by considering as more relevant factor the plate thickness and 5 teachers that seem to consider thermal capacity as relevant for the explanation of the phenomenon, but not plate thickness or thermal conductivity coefficient. It is worthy to note that the 4 teachers performing the correct ranking (2 graduated in Biology, 1 in Physics and 1 in Natural Science) analysed the experimental situation by trying to point out the different characteristics of the various bodies and evaluating the relevance of each of them and not by searching, in their knowledge, for laws to apply to the experimental situation.

In comparing their predictions with experimental results, the majority of teachers that predicted the melting of the ice cubes on the basis of a partial acknowledgement of relevant variables expressed, in different ways, a sort of "surprise" with respect to the results. They did not try to search for an explanation of the differences by showing that their predictions were mainly driven by memory of subjects studied during past courses. Only a few made reference to the necessary comparison between what they remembered from textbooks and real-life experience. The great majority of teachers at last commented that the approach to the proposed situation posed them some difficulties, as it is one that is not part of their theoretical knowledge about thermology.

As previously mentioned, the group discussion during the second phase of the W was devoted to deepen the teacher conception of the different procedures involving in a scientific investigation stimulated by the real situation analysed. A preliminary analysis of the audio taped discussion shows the relevant factors described in the following.

During the discussion, all teachers mentioned (at least one time) the need to apply "the" Scientific Method (SM) in order to solve the different problems involved in the analysis of an experimental situation. When requested to clarify what this means, most teachers recited from memory the steps of this process (with only minor variations): observe, develop a hypothesis, conduct an experiment, analyze data, state conclusions framed in some theory and generate new questions. In their idea, these unproblematic rules give them the guarantee to always find a solution to the problems posed. In order to understand this full confidence in the SM it must be taken into account that the first chapter of all the high school physics textbook has the title “ The Scientific Method”.

Another relevant factor is connected with teachers idea of scientific explanation; very few teachers are aware that explanations are representations of scientific phenomena that link observable features of that phenomenon with hypothesized events, properties, or structures that are not directly observable because of their inaccessibility (as for example the molecular movement) or conceptual nature (as for example forces, energy,….). They did not show a clear distinction between explanation and description, by often using observations or empirical laws as explicative or conjectural models. Finally the fully confidence in the mathematical laws as explicative framework was a common characteristic; our whole teacher sample was completely confident in the explicative value of the Fourier’s law and no one was fully convinced that the Fourier’s law is an empirical law based on observation.

These results obtained in the second phase of our W, integrated with the answers to the problem posed in the first phase account for the need of an epistemological clarification about what the term ‘scientific investigation’ means and how this can be related to the construction of scientific knowledge.

58.5 Discussion and Conclusions

The analysis of data previously reported allow us to draw some conclusions with respect to our research questions.

The majority of teachers showed an approach mainly involving the activation of cognitive resources as memory of past learning experience in order to make sense of reality. It seems that the way the proposed situation is first considered, or “read-out”, plays a crucial role in achieving an accurate description and activating the correct strategies to select the relevant variables. In some cases, these strategies seems to activate “textbook-like” cognitive resources like memory and formulas, acting as conceptual obstacles to the IB approach. This, in some ways works like a sort of “short-circuit of knowledge”, avoiding a phenomenological approach to the problem and a complete formulation of hypotheses. This last conclusion seems to be enforced by the consideration that 3 out of the 4 teachers that correctly explained the ice melting process are not graduated in Physics and probably have superficially analyzed thermal conduction in their University courses. So, it seems that previous knowledge (particularly that coming from university) built in environments not linked with real life experience, is not really significant for the learner and acts as an obstacle for the inquiry competences that we want to develop in teachers. Practices as formulating hypotheses or design appropriate experiments are strongly influenced by the searching for appropriate physical laws.

The second phase of our W showed that our teacher sample had an uninformed view that the scientific method is a fixed step-by-step process; many teachers explicitly declared that “science is a systematic process and that only by following these steps in a orderly way, valid scientific knowledge can be constructed.” Teachers may have these uninformed views about scientific inquiry as a result of the traditional portrayal of recipe-like experiments in science textbooks, as textbooks often play a vital role in understanding the process of science [17]. It may therefore be reasonable to argue that science textbooks should be revised in line with the contemporary conception that there is no single scientific method to be used in developing scientific knowledge [17–20].

Guide lines for future TAs suggested by the approaches followed by our teachers can be synthesizes as follows. During the training time, characteristics of IB practices must be explicitly addressed from a epistemological point of view, as well as problem based activities that are not too much focused on specific disciplinary knowledge. Preliminary results about teacher TAs involving environments based on problems and situations not belonging to the field where teachers are expert show a greater involvement of teachers that pay a greater attention to inquiry procedures rather than to the correct application of disciplinary knowledge. In this way the previously cited "short-circuit of knowledge" seems to be avoided and teachers can activate the reasoning resources necessary for a profitable development of their Pedagogical Content Knoeledge about SI. Moreover teachers need to gain a more epistemically congruent representation of how contemporary science is done by developing activities across different domains of inquiry and many types of investigations.

References

1. Duschl RA, Heidi A, Schweingruber HA, Shouse AW (eds) (2007) Taking science to school. National Academies Press, Washington
2. National Research Council (2000) Inquiry and the national science education standards. National Academy Press, Washington
3. Singer SR, Hilton ML, Schweingruber HA (eds) (2005) America's lab report: investigations in high school science. Washington, The National Academy Press.http://www.nap.edu/books/0309096715/html
4. Collins A (1997) National science education standards: looking backward and forward. Elementary Sch J 97(4):299–314
5. ESTABLISH (2010), www.establish-fp7.eu
6. Akerson VL, Abd-El-Khalick F, Lederman NG (2000) Influence of a reflective explicit activity-based approach on elementary teachers' conceptions of nature of science. J Res Sci Teach 37(4):295–317
7. Lederman NG (1992) Students' and teachers' conceptions of the nature of science: a review of the research. J Res Sci Teach 29:331–359
8. Abd-El-Khalick F (2005) Developing deeper understandings of nature of science: the impact of a philosophy of science course on preservice science teachers' views and instructional planning. Int J Sci Educ 27(1):15–42
9. Glasson GE, Bentley ML (2000) Epistemological undercurrents in scientists' reporting of research to teachers. Sci Educ 84:469–485
10. Van Driel JH, Verloop N, De Vos W (1998) Developing science teachers' pedagogical content knowledge. J Res Sci Teach 35(6):673–695
11. Gess-Newsome J, Lederman NG (eds) (1999) Examining pedagogical content knowledge. Kluwer Academic Publishers, Dordrecht
12. Marx RW, Blumenfeld PC, Krajcik JS, Soloway E (1997) Enacting project-based science. Elementary Sch J 97(4):341–358

13. Roehrig GH, Luft JA (2004) Constraints experienced by beginning secondary science teachers in implementing scientific inquiry lessons. Int J Sci Educ 26(1):3–24
14. Windschitl M (2003) Inquiry projects in science teacher education: what can investigative experiences reveal about teacher thinking and eventual classroom practice? Sci Educ 87(1):112–143
15. Pinto' R (2004) Introducing curriculum innovations in science: identifying teachers' transformations and the design of related teacher education. Sci Educ 89:1–12
16. Shulman LS (1986) Those who understand: knowledge growth in teaching. Educ Researcher 15(5):4–14
17. Abd-El-Khalick F, Waters M, Le A-P (2008) Representations of nature of science in high school chemistry textbooks over the past four decades. J Res Sci Teach 45(7):835–855
18. Abd-El-Khalick F, Lederman NG (2000) The influence of history of science courses on students' views of nature of science. J Res Sci Teach 37(10):1057–1095
19. Lederman NG (2004) Syntax of nature of science within inquiry and science instruction. In: Flick LB, Lederman NG (eds) Scientific inquiry and nature of science. Kluwer Academic Publishers, Bordrecht, pp 301–317
20. McComas WF, Almazroa H, Clough MP The nature of science in science education: an introduction. Sci Educ 7(6):511–532
21. Beyer CJ, Delgado C, Davis EA, Krajcik JS (2009) Investigating teacher learning supports in high school biology curricular programs to inform the design of educative curriculum materials. J Res Sci Teach 46(9):977–998
22. Borko H, Putnam RT (1996) Learning to teach. In: Berliner DC, Calfee RC (eds) Handbook of educational psychology. Macmillan, New York, pp 673–708
23. Clark DB, Linn MC (2003) Scaffolding knowledge integration through curricular depth. J Learn Sci 12(4):451–494
24. Linn MC, Hsi S (2000) Computers, teachers, peers. Erlbaum, Mahwah
25. Magnusson S, Krajcik J, Borko H (1999) Nature, sources, and development of pedagogical content knowledge for science teaching. In: Gess-Newsome J, Lederman NG (eds) Examining pedagogical content knowledge. Kluwer Academic Publishers, Dordrecht
26. Singer S, Hilton M, Schweingruber H (2005) Needing a new approach to science labs. Sci Teach 72(7):10

Chapter 59
Learning Knots on Electrical Conduction in Metals

Giuseppe Fera

Abstract A number of studies related to the conceptual understanding of DC circuits highlighted several difficulties in learning about: the closing of the circuit [1], differentiation of the concepts of electrical energy, voltage and current [2], the determination of the equivalent circuit as a descriptor of multiple loop circuit and use of systemic reasoning instead of local reasoning [3]. Others researchers identified the following issues, which are relevant to learning: the phenomenological relationship between electrostatic and electrodynamics [4–7] and the link between the macroscopic and microscopic level of processes description [8–11]. The dilemma of displaying the microscopic world is still under discussion [12, 13]: students are confronted every day with colorful pictures of atoms and molecules which are spread both by the media and the textbooks, not in accordance with the actual proportions between the physical properties of the particles (size and speed) and can generate misunderstandings about the microscopic processes. The comprehensive discussion of these issues leads to design learning paths on those specific concepts which favor a parallel approach on macroscopic and microscopic levels of the phenomena.

59.1 Introduction

This study is founded on literature survey and presents some research results on teaching/learning of electric conduction in solids. Electricity is one of the most important topics in physic, both looking at the wide contexts of everyday life where it is applied, both for the relevant role of the involved concepts (charge, current, field, potential, …). These aspects make electricity a subject for teaching in every level of instruction. It is also one of the most investigated fields as regards the learning difficulties of students of any age. From the perspective of physics education it's

G. Fera (✉)
Physics Education Research Unit of University of Udine, Udine, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_59,

important to recognize learning knots of the specific context. As concern the electric conduction in metals, there are several knots related to different aspects:

(1) Electrical circuits, (2) Phenomenological relationship between electrostatic and electrodynamics, (3) Link between the macroscopic and microscopic level of processes description, (4) Use of simulations and analogies.

59.2 Electrical Circuits

The researches developed starting from 1980 [1, 14, 15] have analyzed how children working with batteries, wires and light bulbs to build simple circuits. A number of studies shows that the basic concepts related to the d.c. circuits are not intuitive and that electrical phenomena are interpreted by young people through persistent and widespread non scientific schemes. These alternative conceptions coexist in parallel with the scientific view and are not integrated into it, even to the end of education path. The transition from the common sense to the scientific view, according to [2], is favored by the measurement procedures that allow to identify and recognize the physical quantities voltage and current and their role in explaining the functioning of the circuit.

As in the context of electrostatics, learning difficulties are compounded by the fact that the processes involved in the electrical phenomena are far from direct sensory perception. It is not easy to connect the observable effects to abstract concepts like "electric charge", "current" and "energy". Since the students know that the batteries "are consumed" and you pay for household electricity consumption, they believe that in the circuits "something is consumed" and, for many of them, the most reasonable thing to "consume" is the electricity itself, conceived as energy at times, sometimes as electric current. Some authors find that the concept of conservation of electric charge is intuitive for children and yet coexists with the view that the current is consumed. In addition, the concept of electric potential is already quite difficult to acquire in the context of electrostatics, for most students its relevance and applicability in the case of electrical circuits remains unclear. The main learning knots identified, as regards electrical circuits, are:

(a) need of closing the circuit for the current circulation
(b) topological knot: recognizing the equivalent circuit as describer of many loop circuit [3]
(c) sequential knot: the current in a series circuit element is independent of its position
(d) construction of formal thinking: representing the circuit with a symbolic scheme
(e) role of the battery: it is not a source of constant current regardless of the connected circuit
(f) use of systemic reasoning in place of local reasoning: the change of a circuit element can generate changes in all other circuit elements

(g) interpretation and differentiation of the concepts of energy, voltage and electric current [2]
(h) incorrect use of Ohm's law: the absence of current implies absence of voltage.

59.3 Phenomenological Relationship Between Electrostatic and Electrodynamics

Many authors [4, 10, 16] tend to identify in the missing link between electrostatic and electrodynamics the main reason for incomplete students' conceptual knowledge about electrical circuits. These authors argue that students whose education has focused on the microscopic processes involved in the functioning of electrical circuits develop a better understanding not only of simple macroscopic phenomena, but also more complex phenomena such as transient.

Specifically [7] argue that provide students with an explanatory mechanism facilitates the understanding of physical phenomena. They, to teach the electrical conduction, choose the electric field as conceptual referent and the surface charges as a mechanism. In this way, they also realize the connection between the macroscopic and microscopic description of the processes. Main learning knots of context are:

(a) transient processes occurring in electric circuits with capacitors
(b) electric charge concept and physical properties
(c) differentiation of the voltage propagation in the circuit by the current circulation
(d) voltage concept as engine of the charge transfer
(e) link between macroscopic concept of voltage and microscopic concept of electric potential
(f) needs of surface charges on the wire so that the electric field inside the wire is parallel to the wire longitudinal axis [7].

59.4 Link Between the Macroscopic and Microscopic Level of Processes Description

The link between electrostatics and electrodynamics for many authors is integrated with the connection between the macroscopic and microscopic description of the processes of electrical conduction. In contrast, other studies [17, 18] show a cautious attitude on the educational value of the microscopic model of learning by identifying nodes as:

(a) overcoming the local model focused on the behaviour of single atom (as in chemistry)

(b) integration of different models: current as a flow of charge carriers, motion of electrons in metals (random motion in the absence of field, drift motion in the presence of field), description of states and processes in the quantum band model
(c) evidences for the existence of the electron and characterization of its physical properties both intrinsic and interaction with the ionic lattice of the material.

59.5 Use of Simulations

A model of conduction in metals based on the free electron gas [19] can be used to explain the observed relation between resistance and section of the wire (Ohm's second law): the resistance decreases by increasing the section not because the conduction electrons have more space available to move (the average number of collisions of an electron per unit time does not change with respect to the thin wire) but because a greater number of electrons through a section of the wire into time. Similar representations are proposed by the simulations of the interior of a wire with flowing current. According [20], computer simulation tools can be an opportunity to promote the development of scientific thought in the students on issues of particular complexity, such as the relationship between macroscopic and microscopic properties of material.

Simulations can show a phenomenon and/or enable one to make predictions about its trend. Phet Interactive Simulations or Easy Java Simulations provide a microscopic representation of electrical transport in solids. Showing current as moving electron spheres to directly counter the spontaneous model that current is consumed, but the physical proportions among electrons and lattice ions are not respected. However such simulations can help to show how macroscopic properties emerge from a microscopic model at the level of material individual constituent and their interactions.

The use of simulations also involves learning knots among which we highlight the following:

(a) physical proportions among electrons and lattice ions are not respected
(b) free electrons model of metals in Supercomet simulation: the average electric field into a metal is zero, then opposite sign charges don't interact by Coulomb law.

59.6 Analogies

In teaching practice the use of hydraulic analogies is common in teaching of electrical conduction in wires. Under this analogy the wires are likened to pipes and current to a fluid flowing in them, the battery to a pump, etc. These analogies are founded, from the mathematical perspective, on the fact that the transport equations defined by the laws of Fourier, Ohm, Poseuille have the same form to describe the flow of heat, electricity and a viscous fluid, respectively. [21] argue that "using analogies not only promoted profound understanding of complex scientific concepts (such as electricity), but it also helped students overcome their misconceptions of these concepts". However, we can identify knots in the context:

(a) water is a liquid substance and electric current is not
(b) the current repartition between two parallel bulbs connected to the same battery
(c) the current repartition between two different bulbs in series.

59.7 Implications for Physics Education

Common reasoning must be taken into account in teaching: they should be considered not only as creating an obstacle to learning physics, but also as resources at the learner's disposal.

The comprehensive discussion of these issues leads to design learning paths on those specific concepts which favor a parallel approach on macroscopic and microscopic levels of the phenomena.

The relative learning process should be planned as an aware and systematic transition between different levels of description and interpretation of phenomena.

References

1. Osborne RJ (1983) Towards modifying children's ideas about electric current. Res Sci Technol Educ 1(1):73–82
2. Psillos D (1998) Teaching introductory electricity. In: Tiberghien A, Leonard Jossem E, Barojas J (eds) Connecting research in physics education with teacher education
3. McDermott LC, Shaffer PS (1992) Research as a guide for curriculum development: an example from introductory electricity. Part I: investigation of student understanding. Part II: design of instructional strategies. Am J Phys 60(11):994–1013
4. Eylon BS, Ganiel U (1990) Macro-micro relationship: the missing link between electrostatics and electrodynamics in students' reasoning. Int J Sci Educ 12(1):79–94
5. Stocklmayer SM, Treagust DF (1996) Images of electricity: how do novices and experts model electric current? Int J Sci Educ 18(2):163–178
6. Viennot L (2001) Reasoning in Physics, The Part of Common Sense. Kluwer, France
7. Chabay RW, Sherwood BA (2007) Matter and interactions II: electric and magnetic interactions. Wiley, New York
8. Licht P (1991) Teaching electrical energy, voltage and current: an alternative approach. Phys Educ 25:271–277
9. De Posada JM (1997) Conceptions of high school students concerning the internal structure of metals and their electric conduction: structure and evolution. Sci Educ 81(4):445–467
10. Thacker BA, Ganiel U, Boys D (1999) Macroscopic phenomena and microscopic processes: student understanding of transients in direct current electric circuits. Am J Phys 67(7):S25–S31
11. Wittmann MC, Steinberg RN, Redish EF (2002) Investigating student understanding of quantum physics: spontaneous models of conductivity. Am J Phys 70(3):218–226

12. Mashhadi A, Woolnough B (1999) Insights into students' understanding of quantum physics: visualizing quantum entities. Eur J Phys 20:511–516
13. Seifert S, Fischler H (2003) A multidimensional approach for analyzing and constructing teaching and learning processes about particle models. http://www.ipn.uni-kiel.de/projekte/esera/book/s021+fis.pdf
14. Fredette NH, Clement JJ (1981) Students' misconceptions of an electric circuit: what do they mean? J Coll Sci Teach 10(5):280–285
15. Shipstone DM (1984) A study of children's understanding of electricity in simple DC circuits. Eur J Sci Educ 6(2):185–188
16. Tveita J (1997) Constructivistic teaching methods helping students to develop particle models in science. http://www.mlrg.org/proc3pdfs/Tveita_KineticParticleModel.pdf Accessed May 2011
17. Benseghir A, Closset JL (1996) The electrostatics-electrokinetics transition: historical and educational difficulties. Int J Sci Educ 18(2):179–191
18. Borges AT, Gilbert JK (1999) Mental models of electricity. Int J Sci Educ 21(1):95–117
19. Stephens S-A, McRobbie C, Lucas K (1999) Model-based reasoning in a year 10 classroom. Res Sci Educ 29(2):189–208
20. Sins PHM, Savelsbergh ER, van Joolingen WR, van Hout-Wolters BHAM (2009) The relation between students' epistemological understanding of computer models and their cognitive processing on a modelling task. Int J Sci Educ 31(9):1205–1229
21. Chiu MH, Lin JW (2005) Promoting fourth graders' conceptual change of their understanding of electric current via multiple analogies. J Res Sci Teach 42(4):429–464

Chapter 60
Measures of Radioactivity: A Tool for Understanding Statistical Data Analysis

Vera Montalbano and Sonia Quattrini

Abstract A learning path on radioactivity in the last class of high school is presented. An introduction to radioactivity and nuclear phenomenology is followed by measurements of natural radioactivity. Background and weak sources are monitored for days or weeks. The data are analyzed in order to understand the importance of statistical analysis in modern physics.

60.1 An Early Experience

Since 2006, the Pigelleto's Summer School of Physics is an important appointment for orienting students toward physics in our territorial area [2].

Forty students from high school are selected to attend a full immersion summer school of physics in the Pigelleto Natural Reserve, on the south east side of Mount Amiata in the province of Siena.

During the 2009 edition, titled *The Achievements of Modern Physics* [1], a learning path on radioactivity was proposed by one of us (VM) to small groups of students. After a brief introduction to the nuclear phenomenology, they were involved in measures of radioactivity from a weak source of Uranium in order to characterize the emitted ionizing radiation by using an educational device provided by a school.

A teacher (SQ) was impressed by student's involvement and suggested of elaborating a learning path on Nuclear Physics in which the students are active and perform directly measures of radioactivity, in order to propose this activity in her school.

V. Montalbano (✉)
Phys. Department, University of Siena, Ciena, Italy

S. Quattrini
Istituto Tecnico Tecnologico Sarrocchi, Ciena, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_60,

60.2 A Learning Path on Nuclear Phenomena

We planned a path which was proposed as a laboratory within the National Plan for Science Degree and the school obtained founds for it from a regional project supported by Tuscany (Progetto Ponte, i.e. Bridge Project) whose purpose was to promote relations between Universities and High Schools for orienting students.

In order to integrate this learning path in the ordinary school program, the activities were planned for the last year of high school (5a Liceo Tecnologico) after the lessons on electromagnetism.

Teachers make few lessons in class in which presented and discussed the phenomenology of nuclear physics. Afterward, the experimental setup was presented in order to perform measure of background radioactivity in the laboratory.

The next step was to divide students in small groups that are involved in measures of alpha, beta and gamma emission from weak sources of uranium and a very weak source with trace of uranium ore.

During measurements, it rose soon the necessity of performing many measures and the problem of valuating uncertainties. The students were ready to recall all their knowledge on statistics for facing this problem. Table 60.1 shows the topics treated by the teacher in class, in physics laboratory and in computer lab.

60.2.1 A Brief Introduction to Nuclear Phenomena

The lessons give an historical sight on the discovery of radioactivity, and connect the previous knowledge of students to these phenomena so far from their daily experience.

Many examples and easy computations are performed in order to clarify concepts such as binding force, the relative intensities of fundamental interactions at atomic and nuclear distances, mass defect and binding energy, equivalence mass-energy and so on.

Topics related to fission and fusion can be very stimulating for students, specially for our society in these times of discussion about use of nuclear energy and security.

60.2.2 The Experimental Setup

The measures are realized by using a commercial Geiger counter usually used for dosimetry. It can detect alpha, beta and gamma rays and it is possible to collect data without any unit conversion (counts) and download them into a computer.

The detector and the sources are aligned by means of supports on an optical bench as shown in Fig. 60.1a. The sources are shown with their supports in Fig. 60.1b (from

Table 60.1 Covered topics

A brief introduction to nuclear phenomena	Physics laboratory	An introduction to statistical data analysis	Computer lab
Discovery of radioactivity	Background measures	Distributions: Bernoulli, Poisson, Gauss	Verify that radioactive decay follows Poisson's distributions
Pierre e Marie Curie	Measurement of Uranium sources at different distances	Representation of casual phenomena	Testing samples by giving 10, 20, 30, 50 data from the background
Neutron discovery and atom structure	Measurement of Uranium sources at different distances	Mean, median, variance and standard deviation	Compare means and standard deviations of previous samples with the mean and standard deviation of very long measure (more then 1000 data)
Dimensions of atom and its constituents	α, β and γ rays measurement	Bernoulli, Poisson and Gauss distributions	
Natural radioactivity phenomenology (in particular α, β and γ emission)	α, β and γ rays measurement with different shields	Examples	
	Measures with a small Uranium source	The existence of a parent distribution	Verify that for mean values greater than … the Poisson's distribution becomes indistinguishable from the normal one
Isotopes	Measures with a smaller Uranium source	What is a good statistical sample	
Forces into atoms and nuclei (how strong or weak with numerical examples)	Measures with a very weak Uranium ore source	Mean and variance of a statistical sample	Compare measures of background and very weak source
Nuclei can die…			
Law of radioactive decay			
Natural radioactive chains			
Binding energy and mass of a bound state			
Equivalence mass-energy			
Binding energy per nucleon			
Fission and related topics			
Fusion and related topics			

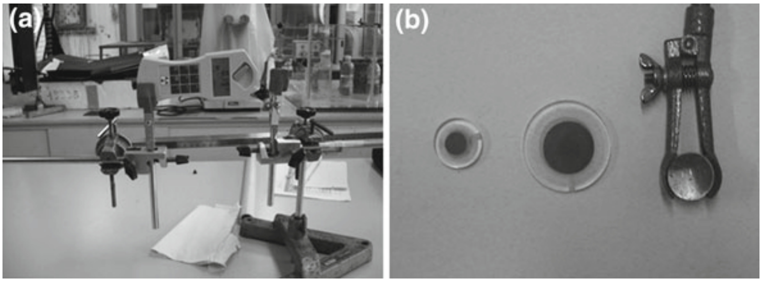


Fig. 60.1 **a** The experimental setup, **b** Radioactive source in their supports

left) small Uranium sheets in plastic holder ($m_1 = 0.152$ g), other Uranium sheets ($m_2 = 0.888$ g), fluorescent marble.

60.2.3 An Introduction to Statistical Data Analysis

We started from a question. There are limits to the precision of a measurement? By answering to this question, we can fix the cases in which the use of statistical data analysis make sense.

The first step is to recall all previous knowledge of students on statistics (there are many but almost never used in physics by teachers), such as mean, variance, standard deviation, distributions [3]. At this point, it is possible to introduce the pupils to some elements of sampling theory.[1] They are ready to deal with the statistical data analysis of our measures.

60.3 Measures and Statistical Data Analysis

The data are collected in the meantime by the Geiger detector and can be downloaded in a computer. The teacher can divide the data in sets in order of giving different tasks to any group of students in computer lab.

In Fig. 60.2, a confront between two slightly different measures can be made. Figure 60.1a shows the data which are divided into to groups ($N_1 = 188$, with mean 182, standard deviation 16, sample standard deviation 1 and $N_2 = 437$, with mean 194.5, standard deviation 14, sample standard deviation 0.7). It is easy to compute the difference between means but it is overwhelmed by the uncertainty if one uses standard deviation, on the contrary the use of sample standard deviation allows to estimate quantitatively the difference.

[1] [4] for a synthetic review.

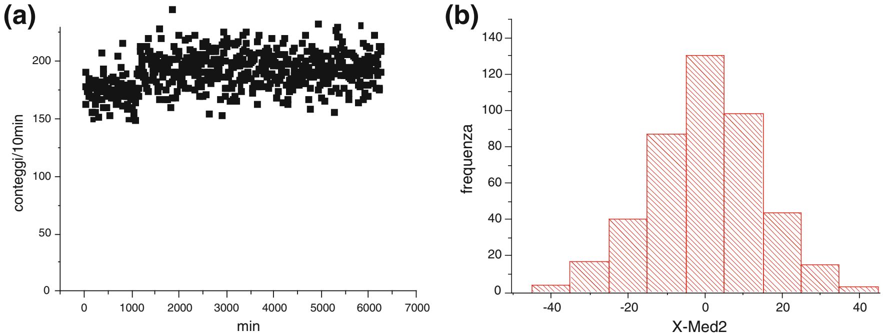


Fig. 60.2 **a** The data from Geiger detector, **b** Frequency histogram for N_2 data

60.4 Conclusion

This learning path on nuclear phenomena was performed on three last classes in 2010 and 2011. We presented the revised path, after analysing the learning difficulties encountered by students in these years. All classes made the introduction to nuclear phenomena in class and in laboratory, only one class made some activity in computer lab. In our opinion, this path is very interesting for students but hardly feasible in last class due to graduation exam. It can be an excellent optional proposal for interested students but if a teacher want to implement it in class it needs to be revised the presentation of statistical topics in previous years.

References

1. Benedetti R., Di Renzone S., Mariotti E., Montalbano V., Porri A (2011) Un percorso di orientamento sulla fisica moderna (An orienting path on modern physics), ComunicareFisica 2010, in press
2. Benedetti R., Mariotti E., Montalbano V., Porri A (2012) Active and cooperative learning paths in the Pigelleto's Summer School of Physics, Proceedings 12^{th} Symposium Frontiers of Fundamental Physics, Udine 21–23 November 2011, submitted
3. Bevington PR (1969) Data Reduction and Errors Analysis for Physical Science. McGraw-Hill Book Company, New York
4. Stat Trek, http://stattrek.com/Lesson3/SamplingTheory.aspx?Tutorial=Stat, accessed 2011 November

Chapter 61
Active and Cooperative Learning Paths in the Pigelleto's Summer School of Physics

Roberto Benedetti, Emilio Mariotti, Vera Montalbano and Antonella Porri

Abstract Since 2006, the Pigelleto's Summer School of Physics is an important appointment for orienting students toward physics. It is organized as a full immersion school on actual topics in physics or in fields rarely pursued in high school, i.e. quantum mechanics, new materials, energy resources. The students, usually forty, are engaged in many activities in laboratory and forced to become active participants. Furthermore, they are encouraged in cooperating in small groups in order to present and share the achieved results. In the last years, the school became a training opportunity for younger teachers which are involved in programming and realization of selected activities. The laboratory activities with students are usually supervised by a young and an expert teacher in order to fix the correct methodology.

61.1 The Pigelleto's Summer School of Physics

In the last years, many Italian Universities are involved in a large national project[1] in order to enhance the interest of high school students towards scientific degrees, in particular in Physics, Mathematics and Chemistry.

In this context, we started to organize a full immersion summer school of physics in the Pigelleto's Natural Reserve, on the south east side of Mount Amiata in the province of Siena. Our purpose is offering to really motivated students an opportunity of testing the scientific method, the laboratory experience in a stimulating context, by deepening an interesting and relevant topic in order to orienting them towards physics.

[1] Progetto Lauree Scientifiche, i. e.Scientific Degree Project from 2006 to 2009.

R. Benedetti (✉) · E. Mariotti · V. Montalbano
Physics Department, University of Siena, Siena, Italy
e-mail: benedetti.2@OSU.edu

A. Porri
Regional Scholastic Office of Tuscany, Arezzo territorial area, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_61,

Forty students from high school are selected by their teachers in a wide network of schools within the National Plan for Science Degree[2] and come from the south Tuscany (Arezzo, Siena. Grosseto).

The school is usually held in the beginning of September and lasts for four days. The 2011 edition was titled *Thousand and one energy: from sun to Fukushima* [5]. Some previous editions were *Light, colour, sky: how and why we see the world* [10, 11], *Store, convert, save, transfer, measure energy, and more*…[10, 11], *The achievements of modern physics* [1, 12], *Exploring the physics of materials* [4, 8].

Topics are chosen so that students are involved in activities rarely pursued in high school, aspects and relationship with society are underlined and discussed.

In the morning, we usually propose lectures in which, by stimulating the active involvement, we give the necessary background for the following activities in laboratory. In the afternoon, small groups of students from different school and classes are engaged in laboratories activities, playing an active role. Every group is supported by one or two teachers that are available for discussions and hints. We propose different laboratories, usually eight in different places. Sometimes they focus on the same physics phenomena explored by different point of view. Every group has the task of preparing a brief presentation in which describing to other students what it has been learned in lab. In the evening, a sky observation is proposed. If it is cloudy, we can propose an indoor problem solving.

In the last years, the school became a training opportunity for younger teachers which are involved in programming and realization of selected activities.

Many educational choices are made in programming the summer school:

- Main target and methodologies are discussed and selected with the teachers involved in PLS
- Almost all laboratories are made with poor materials or educational devices provided by some schools
- The groups are inhomogeneous and formed by following the teachers suggestions in order to promote the best collaboration
- The main topic is related to all activities and must be not trivial
- Almost all laboratories lead to at least one measure and its error evaluation.

Active learning [2, 9] plays a central role throughout all school activities. During the morning lectures, students are engaged in more activities than just listening. They are involved in discussion of or thinking about issue before any theory is presented in lecture or after several conflicting theories have been presented. We often present them a paradox or a puzzle involving the concept at issue and have them struggle towards a solution, by forcing the students *to work it out* without some authority's solution.

In the laboratory activities, they are encouraged in cooperating in small groups in order to present and share the achieved results. The laboratories are organized in such a way that cooperative learning is implemented. A closer analysis of our laboratories

[2] i.e. Piano nazionale Lauree Scientifiche, which continues the legacy of the previous project since 2010.

shows that all requests for achieving a cooperative learning are satisfied [3, 7], such as positive interdependence, individual accountability, face-to-face promotive interaction, social skills and group processing.

61.2 Some Learning Path

Two examples of learning paths made in Pigelleto's summer school are shown in Table 61.1.

We emphasize that it was always planned in the lab at least an activity that we can call *How it works*. This type of activity is very stimulating for students that literally *discover* the underlying physics, as expected in a context-based approach [6].

61.3 Communicate Physics

The oral talk in which students share what they have learned in laboratory is stressed as the central part of the school. Students are not used to speak in public and fear this moment but many of them respond well to the challenge. They can choose the

Table 61.1 Two examples of learning paths

The achievements of modern physics 7th–10th September 2009	Thousand and one energy: from sun to Fukushima 5th–8th September 2011
Morning lectures	*Morning lectures*
On the validity of a physical theory On concepts of classical physics disproved by modern physics Atomic models: from classical to modern ones How laser works Radioactivity, nuclei and surrounding On nuclear energy and energy resources	On Energy Energy from nuclei Energy from sky and sun Energy from the ground On photovoltaic panels On fuel cells Conductors and superconductors
Laboratories	*Laboratories*[5]
How it works: a spectroscope	How it works: a photovoltaic panel
Measurement of Plank's constant with a LED[1] Photoelectric effect[1] We characterize a diode laser We characterize the natural radioactivity Measure of the speed of rotation of a star[12]	How it works: Stirling machine, solar oven, coffeepot Measurement of the mechanical equivalent of heat Electromagnetic induction and energy dissipation Newton's cradles and Yo-Yos Radioactivity background versus. small Uranium sources

form and the means utilized in the communication. Everyone can talk or a group can delegate a speaker, they can present the data in the blackboard or with slides. Furthermore, they can show devices or material for better explain the topic. At the end a brief discussion is performed between the communicators and the public.

In the beginning the communications were proposed in order to share the results of each lab to everyone. It was a necessity because we haven't enough time and space to perform all experiments with everyone. In the years, these communications are became the central activity that stimulates students to be active and cooperative in the learning path.

61.4 Conclusion

The Pigelleto's Summer School of Physics plays a central role for orienting students toward physics in our territorial area. Our main goal is the excellent feedback that returns from students, their families and teachers. In the last years, the applications are increased up to double the available positions. In our opinion, the main reason of this success is the active and cooperative learning of interesting topics in physics that students have experienced during the summer school.

References

1. Benedetti R, Di Renzone S, Mariotti E, Montalbano V, Porri A (2011) Un percorso di orientamento sulla fisica moderna (An orienting path on modern physics). ComunicareFisica 2010 (in press)
2. Bonwell CC, and Eison JA (1991) Active learning: Creating excitement in the classroom. (ASHE-ERIC Higher Education report No. 1, 1991) Washington, D.C.
3. Cuseo J (1992) Cooperative learning vs small group discussions and group projects: the critical differences. Coop Learn Coll Teach 2(3):5–10
4. Di Renzone S, Mariotti E, Montalbano V, Porri A (2011) La Scuola estiva del Pigelleto e la fisica dei materiali (The Pigelleto's summer school and the physics of materials), in atti del XLIX Congresso Nazionale AIF. La Fisica nella Scuola, XLIV suppl. 2:54–57
5. Di Renzone S, Mariotti E, Montalbano V, Porri A (2012) Trasformando l'energia (Transforming Energy), in atti del L Congresso Nazionale AIF, La Fisica nella Scuola, submitted.
6. Kortland, J. (2005), Physics in personal, social and scientific contexts: A retrospective view on the Dutch Physics Curriculum Development Project PLON. In P. Nentwig, D. Waddington (Eds.), Making it relevant. Context based learning of science, Munster, Waxmann, pp 67–89
7. Johnson DW, Johnson RT (1999) Making cooperative learning. Theor Pract 38(2):67–73
8. Montalbano V, Porri A (2011) Ferrofluidi e campo magnetico (Ferrofluids and magnetic field), in atti del XLIX Congresso Nazionale AIF. La Fisica nella Scuola, XLIV suppl. 2:38–42
9. Paulson DR, Faust JL 2011 Active learning for the college classroom http://www.calstatela.edu/dept/chem/chem2/Active/index.htm,. Accessed Nov 2011
10. Porri A, Benedetti R, Mariotti E, Millucci V, Montalbano V (2008a) La Scuola estiva del Pigelleto e I Giocattoli al Liceo (The Pigelleto's summer school and toys at high school) in atti del XLVI Congresso Nazionale AIF. La Fisica nella Scuola, XLI suppl. 3:130–136

11. Porri A, Benedetti R, Mariotti E, Millucci V, Montalbano V (2008b) Esperienze significative del Progetto Lauree Scientifiche di Siena (Meaningful experiences of Siena's Scientific Degree Project), ComunicareFisica07. Frascati Phys Ser, Ital Collect - Scienza Aperta II:121–122
12. Porri A, Benedetti R, Di Renzone S, Mariotti E, Montalbano V (2010) La Scuola estiva del Pigelleto e la misura della velocità di rotazione di una stella (The Pigelleto's summer school and rotation speed measurement of a star) in atti del XLVIII Congresso Nazionale AIF. La Fisica nella Scuola, XLIII suppl. 4:182–188

Chapter 62
The Challenge of Contemporary Society on Science Education: The Case of Global Warming

Barbara Pecori, Giulia Tasquier, Olivia Levrini, Francesca Pongiglione and Margherita Venturi

Abstract The big problems that contemporary society needs to address (e.g. climate change) challenge our traditional idea of education and require to revise the goals of science education research. Such problems are indeed so complex as to require a wide range of competencies to be engaged in producing and implementing solution strategies. Science education is forced to take into account the many dimensions that characterize contemporary science and to face the task of bringing together the potential of all the different perspectives [8]. An example of this kind of research, concerning environmental problems, will be briefly described and its first encouraging results illustrated.

62.1 A Research on Teaching About Complexity

The Global Warming (GW) issue is a problem which existed from more than 20 years. However, it is difficult to capture a deep interest and attention by the general public due to its complexity and the interweaving between different levels/dimensions: scientific, cultural, psychological-behavioral, political-economic and ethical.

From the perspective of Science Education the environmental issues require a rethinking of scientific content and a review of the traditional curricula because of (i) the inter-disciplinary nature of the topic; (ii) the disciplinary problems linked to

B. Pecori (✉) · G. Tasquier
Department of Physics, CIRE, University of Bologna, Bologna, Italy

O. Levrini
Department of Physics, University of Bologna, Bologna, Italy

F. Pongiglione
Department of Philosophy, University of Bologna, Bologna, Italy

M. Venturi
Department of Chemistry, University of Bologna, Bologna, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_62,

the construction of some important Physic concepts (i.e.: transparency, absorption, emissivity); (iii) the effect and the role of the scientific controversy/debate.

From the perspective of Sociological/Behavioral research, the environmental issues require to investigate the emotional attitude of citizens with respect complex issues, like global changes, that have important impact future [4].

A research has been started at the Physics Department, University of Bologna, supported by the CIRE (Interdepartmental Center for Educational Research), aiming at designing a teaching proposal on GW based on the contributions of different fields of research [9].

62.2 First Results of the Research Project

The research work carried out so far has been divided into three phases, that will be described in details in the paragraphs below:

(a) To identify the main emotional, cognitive and cultural barriers that prevent the individuals from engaging in actions that can limit GW (contribution from psychological/behavioral research);
(b) To identify operational criteria for rethinking the disciplinary contents coming from the analysis of barriers (joint contribution of different fields of research);
(c) To make explicit the criteria by applying them to simple models of explanation of GW (contribution of Science Education research).

(a) Identifying Emotional, Cognitive and Cultural Barriers

The first phase consisted in the analysis of some important European reports [1, 3] where some behavioral barriers, which come from the way people perceive the risk of GW, can be identified. We summarized the research results in three kind of barriers that are described below.

i. *Not individual but collective.* The lack of understanding of some basic climate dynamics can cause the climate problem to be perceived by people as concerning only a global and not an individual level: the individual does not take its role as "causal agent" in the interaction between man and nature.
ii. *Too big or too small.* Two opposing perceptions of the risks associated with climate change can both cause reactions of "denial" of the problem. If risks are seen as too big, the individual feels he can do nothing about them and therefore does not care; if risks are seen as too small, the individual feels to be entitled to neglect them. The uncertainty that characterizes some aspects of the scientific debate, produces disorientation, loneliness and sense of guilty.
iii. *Too far.* To individuals it is not clear what are short-term and what are long-term risks: the present generation feels to be called to "pay" for a probable good of the future generations.

The research in social/behavioral science shows that there is an emotional attitude which manifests itself in divided feelings as guilt, confusion, frustration. Faced with the complexity of the debates on environmental issues, like global warming, people have a different kind of “fear” than respect to cataclysms or disasters. This kind of fear acts at a deeper level and causes people to keep a safe distance from the issue of climate priorities. The unknown, the uncertain, the complex, the debated, all generate a sense of disorientation and a feeling of inability to act [5].

(b) Defining Operational Criteria for Disciplinary Reconstruction

The second phase consisted in defining some operational criteria to guide the process of discipline reconstruction that can allow science education to contribute to overcoming the emotional/cognitive/cultural barriers:

i. As far as the barrier *Not individual but collective* is concerned the corresponding criteria is to make the role of the individual explicit - what interaction between man and nature - in the modeling of global warming.
ii. For the barrier *Too big or too small* the criteria is: to make it explicit, in the modeling of global warming: what is shared and what is under discussion; examples of causal connections, typical of complex systems, between man-nature-technology and examples of feedback to show that small causes can have large effects and back again.
iii. For the barrier *Too far* the criteria is: to place, the examples already treated, on a time scale, typical of an evolutionary approach to complex systems, to reflect on possible future scenarios (i.e.: the melting of glaciers).

(c) Applying the Criteria to the GW Model

The third phase consisted into applying the criteria to a simple model of GW explanation [6], where the main role is played by the*coefficient of absorption in atmosphere (a).*

In the model we have chosen it is possible to apply criterion (i) because it is possible to recognize the central role of human action in the variation of the parameter a.

It is also possible to apply criterion (ii) because it is possible to show examples of causal links and feedback mechanisms (i.e.: If a increases $\rightarrow$ then T increases; if T increases $\rightarrow$ the glaciers are melting; if the glaciers are melting $\rightarrow$ a increases).

As far as criterion (iii) is concerned the model offers the possibility to raise the issue of the predictive character of the models that study the evolution of complex phenomena, like the melting of glaciers.

62.3 Future Developments

This synthetic description is intended to highlight the importance and the potential of a reflection on the changing role of Science Education to face big problems that society needs to address.

The results already obtained support the hypothesis that addressing the issue of GW from different perspectives may help to develop teaching materials on this topic able to convey the message that:

– The scientific controversy does not concern the existence or not of GW, but it concerns the predictions of the models in terms of consequences of the GW;
– The models of GW cannot show a power of prediction in classical terms (like in Classical Mechanics) because the phenomenon is intrinsically complex and multi-dimensional.

To introduce the epistemological dimension, in particular epistemology of complexity seems potentially able to provide students with cultural tools to rationally navigate in the jungle of ideological wars about environmental issues.

This is an important goal for Science Education in the challenge of finding an effective perspective, not only for teaching Physics, but also for educating young people to Scientific Citizenship.

The future research development includes design of a teaching proposal for introducing GW as a topic in physics teaching at higher secondary level, integrated in a wider teaching path about Thermodynamics, already produced and tried out by the Physics Education research group of the Bologna University [2, 7].

References

1. IPCC (2007) Climate change: Synthesis Report
2. Levrini O, Fantini P, Gagliardi M, Tasquier G, Pecori B (2010) A Longitudinal approach to appropriation of science ideas: A study of students' trajectories in thermodynamics. Proc. 9th ICLS conference, Chicago (IL)
3. Norgaard K M. (2009) Cognitive and behavioral challenges in responding to climate change. Background paper to the 2010 world development report. The world bank development economics
4. Pongiglione F (2011) Climate change and individual decision making. FEEM Working Paper
5. Pulcini E (2009) La cura del mondo. Paura e responsabilità nell'età globale, Bollati Boringhieri, Torino
6. Rizzi R (2011) Atmospheric Physics. Material from lectures, A.A, pp 2010–2011
7. Tasquier G (2009) Un esperimento di termodinamica in una classe IV liceo scientifico: analisi della fattibilità di un percorso innovativo. Tesi di laurea specialistica. Olivia Levrini, Relatore
8. Tasquier G, Nonni C (2011) Per una educazione scientifica che investe sul futuro: il patrimonio della tradizione e la sfida del mondo contemporaneo. Università di Bologna, Giornata di studi del CIRE
9. Tasquier G, Pongiglione F (2011) Problemi ambientali e loro complessità scientifica e sociale: Una sfida per la ricerca in Didattica delle Scienze. Atti XCVII Congresso Nazionale SIF, L'Aquila, 26–30 settembre, 2011

Chapter 63
Magnetic Field as Pseudovector Entity in Physics Education

Carlo Cecchini, Marisa Michelini, Alessandra Mossenta, Lorenzo Santi, Alberto Stefanel and Stefano Vercellati

Abstract The analysis of the nature of the magnetic field offers the ideal framework in which students could address the mutual integration between mathematics and physical aspects facing experimentally the analysis of the phenomenology. Took look how students face experiments in which the phenomenology founds the theory and the mathematics offers its formalism as the best language to describe the explored phenomena, an activity concerning the pseudovectorial nature of the magnetic field was performed looking at the way in which the students' reasoning evolve.

63.1 Introduction

As in several physics situations, the knowledge of the symmetries of the physical systems analyzes the situations in a more effective and simple way, but the individuation of the symmetries could not be done without passing through the knowledge of the laws of transformation of the single entities that take part in the systems [1, 3, 6, 8]. For instance, symmetries are related with laws of conservation through the use of the Neother's first theorem [7].

Unfortunately the role of symmetries in high school physics education is often underestimated. The ways in which entities are transformed by symmetry operations are usually not explicitly adressed, the usual goals of the high school physics courses are only aimed to the definition of the structures of the entities, and this approach creates an intellectual gap between students' studies in high school and university courses [2].

C. Cecchini (✉)
Physics Department, University of Udine, Udine, Italy

M. Michelini · A. Mossenta · L. Santi · A. Stefanel · S. Vercellati
Mathematics Department, University of Udine, Udine, Italy

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_63,

The main example of this intellectual gap is related to the distinction of axial and polar vectors. In high school this distinction is usually neglected and this learning knot lies submersed until students had to face it in the university courses.

This research is aimed to study how high school students' reasoning evolves in the construction of formal thinking when they face an explicit situation in which the pseudovectorial nature of the magnetic field is highlighted.

63.2 Context and Sample

Research work was done in the context of the Summer School of Modern Physics held in Udine during the summer of the 2011. The students participating to this school were the best 42 students coming from the last 2 years of the Italian high school (students are 17–18 years old; school grades 12th–13th). The opening course of this intensive school was the course of electromagnetism. It was a 16 hours long course distributed on 3 days.

63.3 Instruments and Methods

During this course of electromagnetism, students, divided in 7 groups of 6 components each one (4 groups of 12th grade students and 3 of 13th grade), follow an inquired based learning path that was constructed on the experimental exploration of the formal properties of the magnetic field with the aim to construct a formal representation of it. In particular the works on the exploration of the pseudovectorial properties of magnetic field was proposed in this learning path as an interactive lecture demonstration using an inquired based tutorial constructed with the Prevision-Experiment-Comparison (PEC) strategy [4, 5].

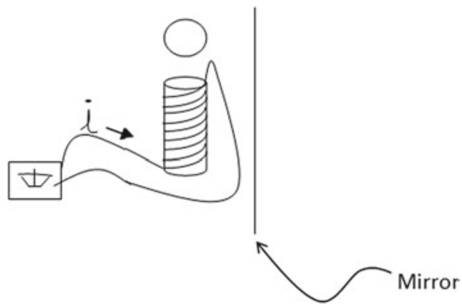


Fig. 63.1 Situation proposed to the students in the first step

The situation proposed was the one represented in Fig. 63.1. Taking into account the given circuit, students had to draw its mirrored image representing the needles of the compass, the needle of the reflected compass and analyze the picture (the compass is represented as the circle placed above the coils).

The second step was the realization of the experiment and the analysis of the original and the mirrored situations. Then, in the third part was asked to the students to rise up some considerations starting from the experimental observations.

63.4 Data

Data were collected using personal inquired based worksheets. In the first phase, the provisional phase, 36 % of the students represent the mirrored image of the needle as simple reflection, 28 % represents the needle of the compass with opposite verse, 5 % represent only the mirrored apparatus and 36 % did not replay to this task (Fig. 63.2).

At the second phase, 19 % of the students highlighted that the verse of the mirrored image of the needle is the same without justify it, while 14 % highlights that the verse is different justifying their answers saying that the verse is different due to the right hand rule (2 %), because the verse of rotation of the coils change (2 %), because there is a change in the direction of the magnetic field (Fig. 63.3).

In phase three, 64 % of the student highlight that the magnetic field is not a vector reporting only the representative matrix wrote by the teacher on the blackboard 31 %, highlighting the rule in the recognition of the nature by the different rule of reflection 29 %, pseudo magnetic vector is like the angular momentum or quantities that are defined through a vector product, 14 % because the scalar product is anti-commutative 5 %, or give other argumentation 4 %, 5 % did not motivate and the remaining 36 % did not reply (Fig. 63.4).

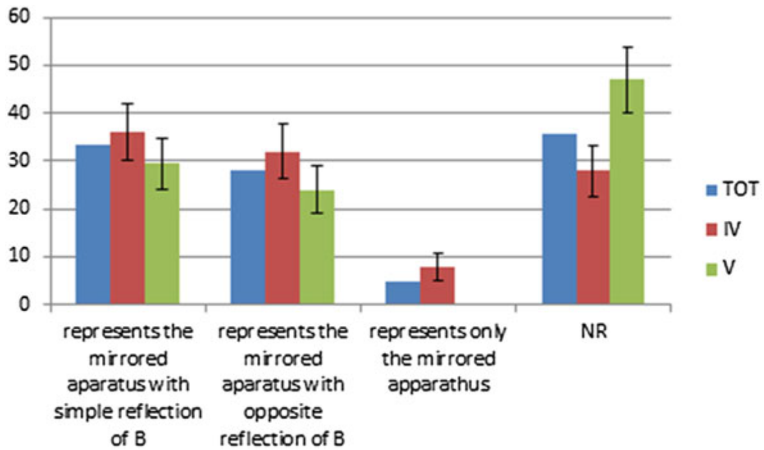


Fig. 63.2 Representation of the mirrored image of the needle done by the students

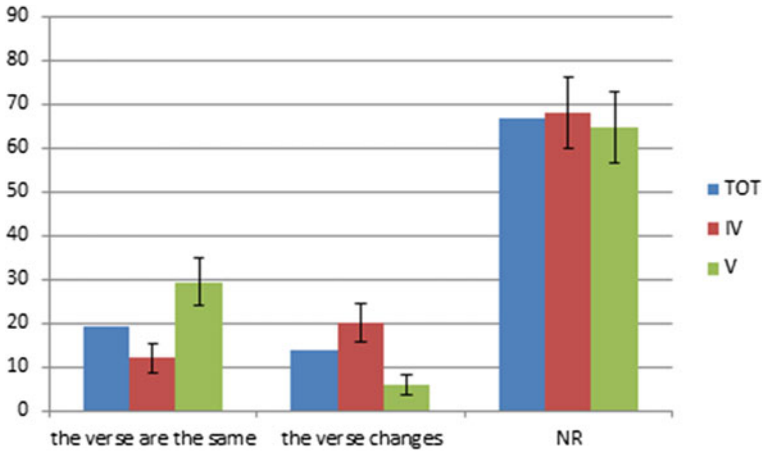


Fig. 63.3 Observation done by the students of the experimental results

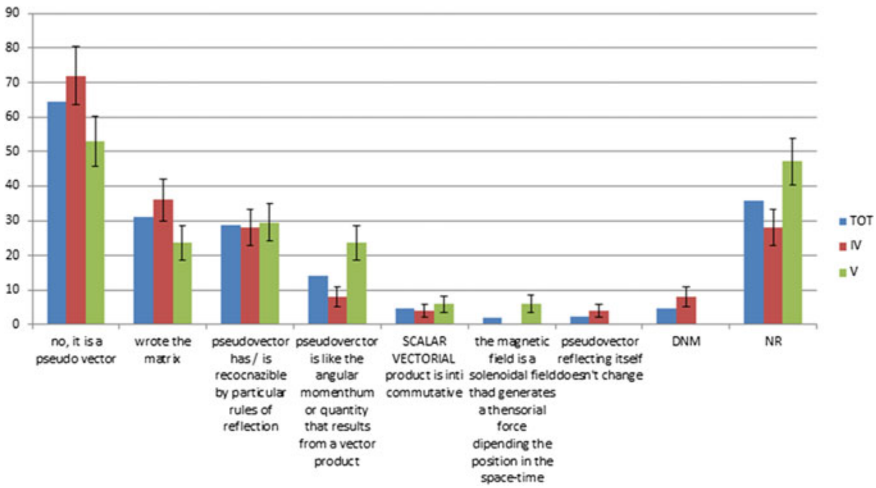


Fig. 63.4 Consideration done by the students on the nature of vector of the magnetic field

63.5 Data Analysis

Neglecting the students who did not replay and the ones that represented only the mirrored apparatus, students' spontaneous approaches to the analysis of the mirrored situation are almost distributed equally on two different way of analysis: the first one is the representation of the needle of the compass as a simple reflection of it and the other is the drawing of the compass needle starting from physical consideration of the mirrored apparatus. No significant differences were highlighted between the 13th and 12th grade students.

In phase two, the analysis of the experimental situation, could be noticed how decrease the number of students that thought that the compass needle (i.e. the magnetic field) is reflected in the standard way, but not to a negligible percentage highlighting two main approaches. The first approach followed, by that the students of 13th grade (that had already faced the magnetic field description during the previous school year), is to look at the experiment as a confirmation of their ideas and not as an investigation opened to new findings, do not allowing them to notice keys elements that are in opposite with their thesis. The second approach, followed mainly by the 12th grade students is more open to the acceptance of eventually discording results respect to their prevision and so they can noticed and highlighted these discrepancies.

After the discussion, in the third phase, no one of the students recognize the magnetic field as a 'standard' vector, but the motivations that they gave are distributed on a wide spectrum that is reported in Fig. 63.4. Interesting to be noticed is that, even almost one third of the students justified this aspect reporting only the formal structure wrote on the blackboard by the teacher, the remaining part highlights experimental evidences to support this inference. In addition the data also highlight how mainly the 13th grade students propose comparisons and analogies with other quantities that they had already faced during their previous study (as for instance the angular momentum).

63.6 Conclusions and Further Remarks

This experimentation, done on the introduction to high school students of the nature of pseudovector of the magnetic field, highlight how through a simple experiment student could face this formal characteristic, recognize it and do comparison between this characteristic and analogous previous physic entity that they already know. In addition, looking data concerning question two, were highlighted how the standard way of teaching used in the high school represent an obstacle to the acceptation of this 'new' property by the student in they further studies.

References

1. Foot R, Volkas RR (1995) Neutrino physics and the mirror world: how exact parity symmetry explains the solar neutrino deficit, the atmospheric neutrino anomaly, and the LSND experiment. Phys Rev D 52:6595–6606
2. Kolecki JC (2002) An introduction to tensors for students of physics and engineering. Technology report, NASA Glenn Research Center, Cleveland, Sept
3. Kozlov VV (1995) Symmetries, topology and resonances in Hamiltonian mechanics. Springer, Germany
4. Martongelli R, Michelini M, Santi L, Stefanel A (2000) Educational proposals using new technologies and telematic net for physics, In: Physics teacher education beyond

5. Michelini M, Ragazzon R, Santi L, Stefanel A (2004) Implementing a formative module on quantum physics for pre-service teacher training. In: Michelini M (ed) Quality development in the teacher education and training. Girep Book of Selected Papers, Udine, Forum, pp 429–435
6. Mohapatra RN, Senjanovi'c G (1981) Neutrino masses and mixings in gauge models with spontaneous parity violation, Phys Rev D 23:165
7. Noether E (1918) Invariante Variationsprobleme. Nachr. D. König. Gesellsch. D. Wiss. Zu Göttingen, Math-phys Klasse 1918:235–257
8. Redlich AN (1984) Parity violation and gauge non-invariance of the effective gauge field action in three dimensions. Phys Rev D 29:2366

Chapter 64
A Model of Concept Learning in Physics

Wagner Clemens and Vaterlaus Andreas

Abstract Learning concepts in physics is difficult due to the often simultaneous presence of misconceptions. We have developed a dynamic model of concept learning, which includes the dynamics of misconceptions. In its simplest form the model simulates the learning of the concept and the unlearning of the misconception with a two-dimensional differential equation system. The major conclusion from our model simulations is that while teaching a concept, misconceptions should be intensively addressed. Rapid decay of concept knowledge observed in experiments after concept teaching is explained in our model with the persistent presence of a high level of misconception.

64.1 Introduction

Recently, exploring the time dependence of learning physics concepts has gathered increased attention [1–3]. The data about concept learning shows convincingly that the acquisition of concepts is a dynamic process. As a consequence the interpretation of quantities like the normalized gain, for example, requires great care since the results strongly depend on the time points when the data are recorded. Therefore a more thorough understanding of the dynamic process of learning is required. Heckler et al. [2] investigated concept learning of an introductory physics course at the university level. Data were collected from more than 1,600 students. Every few days or every week 12 students (in the average) were asked to solve a set of multiple-choice questions concerning physics concepts. Students were asked before, while and after the topic was taught. The proportion of correct answers was then plotted against time. It turned out that the most effective learning event was the homework activity since the homework due time showed the highest level of concept knowledge.

W. Clemens (✉) · V. Andreas
Solid State Dynamics and Education, ETH Zurich, Schafmattstr. 16, 8093 Zurich, Switzerland

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_64,

The goal of this work is to construct a model, which is able to simulate the time course of concept learning. The model we present is based on previous work done by Lei Bao [1]. The major difference is that our model consists of two variables, one for the concept c and one for the misconception m. It distinguishes between not knowing the concept $(1 - c)$ and the level of misconception m. The differential equation system we present to simulate concept learning has two stable fixed points, a concept fixed point and a misconception fixed point. Learning of concepts means in terms of our model, to move the system away from the misconception fixed point so that it relaxes on the concept fixed point after the learning phase.

The major conclusion from our model is that if misconceptions are not properly addressed in class, resulting in a reduced misconception level, teaching of concepts in physics will be almost futile. Student's concept knowledge rapidly decays, after the learning phase. Thus an initial unlearning phase of misconceptions is necessary (if misconceptions are initially present) to settle concept knowledge on the long run.

64.2 The Model

Our model of concept learning consists of two coupled differential equations, one for the concept c and one for the misconception m

$$dc/dt = \alpha c(1 - c)(1 - m) - \gamma c\,m,$$
$$dm/dt = -\beta \mathrm{m}(1 - \mathrm{m})\mathrm{c} + \mu(1 - \mathrm{m}).$$

The parameters α and γ are the coefficients for learning and unlearning of concepts. The misconception is learned with rate coefficient μ and unlearned with rate coefficient β. The levels of the concept and the misconception are in the interval [0, 1]. The variable c is defined as probability to solve a concept question correctly at the given time point. Although we call the variable m misconception level it summarizes all detrimental effects on the concept knowledge. Therefore m hampers the learning and promotes the unlearning of concepts. We assume that the learning of concepts is hard at the beginning and at the end. This justifies that the learning rate is proportional to $c(1 - c)$. The most progress is made for concept learning of $c = 0.5$. The same argument also holds for the unlearning of misconceptions (unlearning is proportional to $m(1 - m)$. However, learning of concepts does not only depend on the concept c but also on the level of misconception, m. Thus learning of concepts is easier the larger the factor $(1 - m)$ is. The stronger the misconceptions are the slower is the learning of concepts. The unlearning of concepts is proportional to the product of the level of the concept and the misconception. If there is no concept there is no unlearning and the stronger the misconception the stronger the unlearning. The unlearning of misconception is amplified by the presence of a high concept level. Therefore we multiplied the misconception unlearning term $-\beta m(1 - m)$ by c. Finally, misconceptions have a permanent learning term, $(1 - m)$, due to everyday experience (for example: there is no real world experiment without friction) or due to intuitive perception (for example: adding a resistor to a circuit should increase the equivalent resistance).

64.3 Results

In order to find reasonable values for the learning and unlearning parameters we used the data provided by Heckler and Sayre [2]. They applied multiple-choice questions to students in order to test their understanding of physics concepts in Mechanics and Electricity and Magnetism. A data point corresponds to the normalized answer of 12 students (average). At every time point a different representative sample of students was asked the same questions. The data points of the electric circuit problem are shown in Fig. 64.1.

If a topic is taught directly by lectures, labwork or homework we increase α and assume an enhanced learning ability α_{enh}. To obtain the fit curve we used the procedure (*fminsearch*) provided by *Matlab*. The procedure was allowed to adjust the following parameters providing the results $\alpha = 0.4$, $\alpha_{\text{enh}} = \tilde{3}44$, $\beta = 0.24$, $\gamma = 0.043$, $\mu = 0.0015$ and the initial value of the misconception level $m_{\text{init}} = 0.8$. The time points for the onset and the end of the enhanced learning phase were kept constant at $t_1 = 28$ days and $t_2 = 39$ days. These time points correspond to the onset and offset of the teaching period. We selected for the initial condition of the concept the first point of the measurement $c_{\text{init}} = c_5 = 0.47$ (circuit question). The fitted misconception level at the beginning is approximately $m_{\text{init}} = 0.8$ and seems to be rather high. The rate constant for enhanced learning is about ten times higher $\alpha_{\text{enh}} = \tilde{3}44$ than the basic coefficient $\alpha = 0.4$. It is also remarkable that the rate for unlearning misconceptions is of one magnitude smaller than the rate

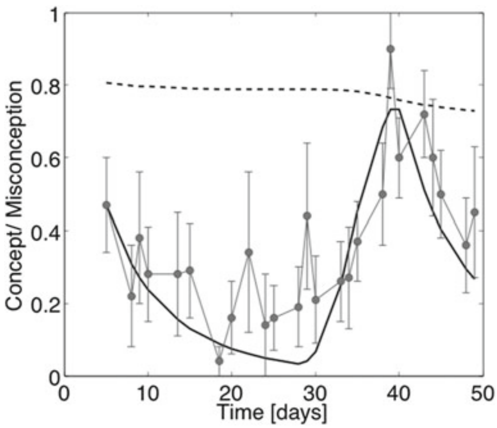


Fig. 64.1 Fit of the model to the data provided by Heckler and Sayre [2] (electric circuit problem). The *full line* represents the concept knowledge and the *dashed line* the misconception level

of concept learning. As a consequence, the unlearning of misconception has to be done in an efficient way in order to successfully teach concepts. If the misconception level is not substantially reduced (when starting at $m_{\text{init}} \cong 0.8$) during the learning phase, the concept level decays quite rapidly after the enhanced learning phase (see Fig. 64.1). Finally students end up in the misconception fixed point and all efforts to understand the concept were in vain. The long-term behavior of the system is shown in Fig. 64.2 (thin full line: concept; thin dashed line: misconception). It emphasis our point of view that we not only have to teach concepts, but we also have to unteach misconceptions.

What happens to the student knowledge if misconceptions are unlearned at the same time as the concept is taught is shown in Fig. 64.2. In order to do the corresponding simulations we introduced an enhanced unlearning coefficient for the misconception $\beta_{\text{enh}} = 0.37$ during the concept teaching period. Between days 28 and 39 we changed the misconception unlearning rate from β to β_{enh}. During this time period the misconception level drops markedly so that after the learning phase the system drives towards the concept fixed point.

64.4 Discussion

In order to provide a better understanding of the evolution of concept knowledge we have developed a new model, which allows for simulating the interaction between the concept and the misconception. Using the model we fit the data about concept knowledge recorded by Heckler and Sayre [2]. In our model students start at a medium level of concept knowledge and at a strong misconception level. For the first 27 days the model moves towards the misconception fixed point. Between day 28 and 39 the learning activities take place leading to a change of the learning coefficient from α to α_{enh}. After day 39 the topic is assumed to change and without

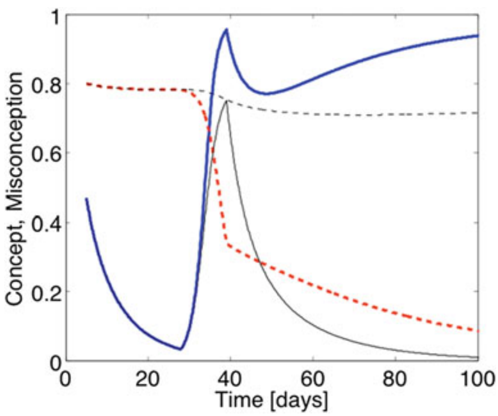


Fig. 64.2 Simulation of unlearning misconceptions. The *thick full line* represents the concept knowledge and the *thick dashed line* is the level of misconception. The *thin lines* represent the simulation of the concept (*full*) and the misconception (*dashed*) due to the data shown in Fig. 64.1 for comparison

unteaching misconceptions the system moves back to the misconception fixed point. In contrast, if the misconceptions are properly addressed during the learning phase the misconception level drops substantially so that after the learning phase the system ends up in the concept fixed point on the long run.

References

1. Bao L (2008) arXiv:0710.1375v2 physics-ed-ph
2. Heckler AF, Sayre EC (2010) The effects of physical and virtual manipulatives on students conceptual learning about pulleys. Am J Phys 78:768–777
3. Sayre EC, Heckler AF (2009) Peaks and decays of student knowledge in and introductory E & M course. Phys Rev 5:013101 (STPER)

Part XI
Popularization of Physics

Chapter 65
Invitation to Physics not Only for Gifted Pupils

Stanislav Zelenda

Abstract In Talnet–we create new opportunities for gifted pupils in science. We suppose any level of giftedness or interest must have practical opportunities to be found and developed. We focus on study, research, inquiry, collaborative and modeling activities where the learning itself is not a primary goal. The activities combine online with face to face and are designed for 13–19 years old pupils. More than 1,000 pupils has participated in 28 different yearlong courses, more then 500 pupils in tens of different "T-excursions" and collaborative research activities. Gifted pupils involved such systemic and authentic activities can play an important role on the pupils' side of communication with "physics".

65.1 Introduction

It is a matter of fact that gifted pupils and students are of a great potential for physics as future researcher, explorer, experts etc. This contribution is about a secondary, but effective, potential of gifted pupils for popularization in physics (communication of physics). It is based on our eight year long experience of the national project "Talnet — online to science" that is focused on 13–19 years old gifted pupils interested in science, mathematics and technology. Despite the discussions about giftedness the practical approach for identification and development of gifted pupils we use is a very practical approach. The approach to giftedness is close to Renzulli's definition of giftedness "…an interaction among three basic clusters of human traits: above-average general and/or specific abilities, high levels of task commitment (motivation), and high levels of creativity…" [1]. We consider giftedness as a complex notion including not only passive—segregation bur preferably active—development stimulating moments.

S. Zelenda (✉)
Faculty of Math and Physics, Charles University, Prague, Czech Republic

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_65,

We looked at real situations in physics teaching in schools, opportunities for pupils who are interested in science (physics) in or out of schools in the Czech Republic from the point of present knowledge about gifted pupils. In the classroom there are very limited opportunities for individual and differential care about gifted. The national system of extracurricular activities consists of:

– *On national level*

Olympiads and other national contests [2]
Secondary School Expert Activity [3]

– *On regional or local level*

“Correspondence seminars”, “Child’s universities”
“School club” type activities
“Camps”
“Stints” (e.g. at research institutions)
Individual or series of lectures
Visits to (interactive) exhibitions in museums, “Science fairs” etc.

Most of these activities (namely that of higher population capacity) are more competitive than collaborative; more individualized than team oriented, more isolated than inter-related, systemic, and continuously developing pupil’s ability with respect to individual potentially gifted pupil’s abilities and needs. It is hard to find activities supporting systemic development of creativity, critical thinking, enquiry skills and emotional, (meta) cognitional, communicational and other social needs of gifted pupils.

There is only a very limited experience in identifying and teaching gifted pupils among schoolteachers and among university teachers.

Talnet [4] aims to cover some of the gaps and invite subject experts to meet, communicate and make an authentic more continuous contact (most of time moderated via computer and web) with potentially gifted pupils.

With the help of online (web) technology and face to face activities Talnet offers a variety of different systemic, continuous activities to our pupils and teachers accessible in some sense sufficiently in time and geographically. The activities are specially designed with respect to present knowledge about gifted pupils. Discipline experts and facilitators are guiding or supporting pupils in their activity. Two important principles make the system usable as a practical tool for identification of gifted pupils interested in a discipline. 1. Suggested relatively autonomous blocks of activities (topics, enquiry etc.) should be designed and offered in several different levels (different assignments) to be meaningful and at least partially completed by pupils of different (pre) knowledge, interests and ability levels. 2. The second principle: The available support and usage of the Talnet environment (including communications with experts) is limited only by pupil’s purposive activity.

Pupils, nominated mostly by their teachers or experienced friends, reveal important characteristics of gifted and talented pupil’s profiles (cf.: [5]) in such an online environment and in a variety of structured activities. Keeping them in their school among classmates and simultaneously in their "gifted pupil's" (virtual) community we support them in accomplishment (filling up with people) their two different "Zones of Proximal Development" [6] in each of them they can play different roles. That is the important factor of their potential for communication of physics to their class or school mates.

65.2 Types of Activities in Talnet

Activities differ by subject (by subject are organized in series) and by individual design that should fit with the guide's (instructor–the subject expert) intentions and styles. The types of activities (Table 65.1) differ namely by organizational forms, durations, outputs an expected assets to kids.

T-courses mostly offer online individual studies of several topics, based on communication between a student and an expert when working on assignments. The year-long study of such a course is divided into two online study period and a period to complete individual or team research like 'seminary work'. When completed it is online discussed and defended among all pupils of all courses of Talnet. The authors are invited to present it at one (of three) face-to-face meetings of the courses faced to a jury. Talnet T-courses in physics (and mathematics support) realized annually follows:

For ages 13–15:
Mathematics 0-Functions and Transformation
For ages 14–15:
Astro and Modeling I-Astronomy in Black and White for Now, and Learning Modeling on Computers
Mathematics I-Mathematic Algorithms and Their Geometric Representations
Materials and Crystals I-Experimenting and Measuring
For ages 15–16:
Astro and Modeling II–Astronomy in Color and Our Models are Usable
Mathematics II–Geometry in Standstill and Motion
Materials and Crystals II–Exploring Structures
For ages 15–18:
Doing Physics in a Team
For ages 16–17:
Mathematics III–Combinatorics Games
Meteo and Modern Physics–World in Terms of Meteorology, Solar Energy Uses
For ages 17–18:
Astro and Modeling IV–The Space in Ultra and Getting Better with Our Models
Programmable Machines–Simulation and Control
SW Mathematica for the Inquisitive-Application I
Selected Chapters of the Theory of Relativity–Special and General

The T-pro-seminars train more general types of skills, e.g. how to make presentations, to process measurement data, to use certain SW, etc. The online support of long term face-to-face activities, such as a series of lectures concerning the Mathematics

Table 65.1 Types of activities in Talnet

Activity	Nature	Main objective
Structured, individual and team		
T-excursions (combined)	Education	Opportunity to demonstrate interest
T-courses (online)	Education	Knowledge deepening and … to learn to study
	Research	Application of ability/interest … to learn to research
T-pro-seminars (online)	Education and training	Acquire general skills
T-projects (combined)	Study	Deeper study of selected problem
	Research	Authentic research
Online support (e.g. courses for Math. Olympiad participants…)	Education	To offer more active forms of preparation for existing activities
	Research	To offer more active forms of preparation for existing activities
International activities (combined)	Communication and Research	To experience international interaction … the need of language skills
Loosely structured		
TalnetSpace, Café Talnet, T-journal (pending)	Communication	Own suggestions, information exchange within the community…
	Assistance	To apply the acquired knowledge in helping others
	Production	Presentation of the results and the community
	Games	To apply and develop the intellect
Expedition	Project	To apply the range of knowledge, skills, and volatile qualities
Workshop	Experience	To complement the online activities

Olympiad, allows the participants to practice, deepen and consult their knowledge gained during the traditional lectures. The most demanding activities are the multinational international activities (T-international). Another extreme, in terms of demands on knowledge, creativity, and observation, experimenting and research skills, are the team's enquiry activities (e.g. Doing Physics in a Team).

65.2.1 Team's Enquiry Activities-Doing Physics in a Team

This activity uses annually published problems of International Young Physicists' Tournament to be explored, investigated and explained in a virtual team consisting

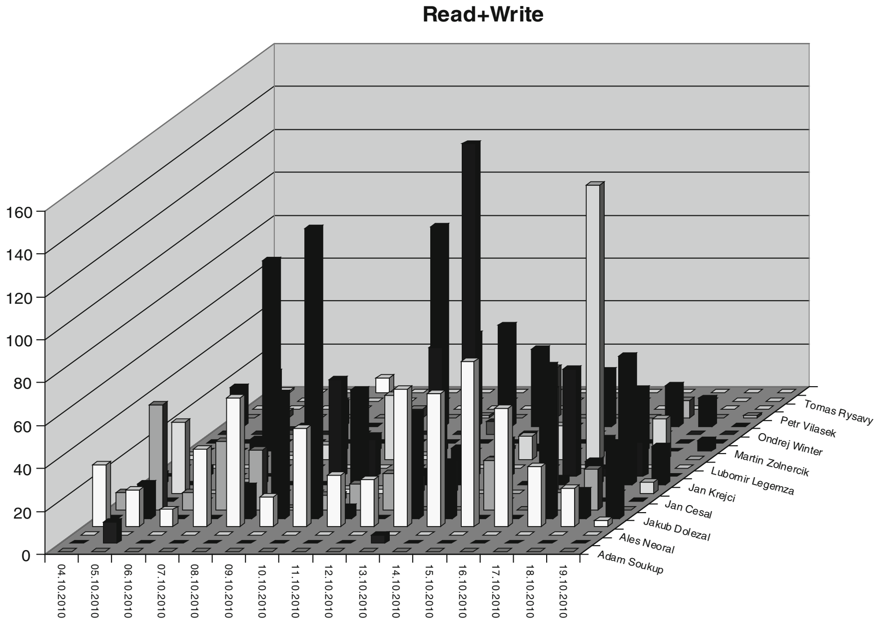


Fig. 65.1 Frequency of pupil's activities in the Team's Enquiry Course (by day and pupil)

of pupils from different schools and regions. It is a very difficult task for facilitators to manage up to 30 pupils in such ambitious activities. The Tournament competition is developed for a five member team in a school. But lack of capacity of physics teachers and lack of pupils' motivation (typically only one or two pupils of the school are interested and ready to spend so much time by in physics enquiry, limits the possibility of Young Physicists' Tournament in the school practice. The Talnet course Doing Physics in a Team get together individual pupils who have no partners in the school and on the other side small teams (typically a couple or triple) of interested pupils from the same school get needed support and training useful for participation in the Young Physicists' Tournament regardless if they are gifted or not. The course environment includes T-pro-seminars, sources, (modeling and other) tools carefully prepared to be used by students when needed in studying the problems.

The above Figs. 65.1 and 65.2 illustrate "character of collaborative enquiry and mutual learning" in the virtual environment. A closer analysis of communications shows that communications are horizontal (among pupils), more regular, more frequent and topic related. The difficulty and complexity of enquiry problems make pupils to meet more frequently than in other courses.

This type of courses aims to create some opportunities to solve well known epistemological problems of school science (teaching) in almost all countries: the absence of development of science creativity by reducing science to the only moment of logical and experimental defense. This way the alumni (also future researchers, scientists) are not ready for discovering.

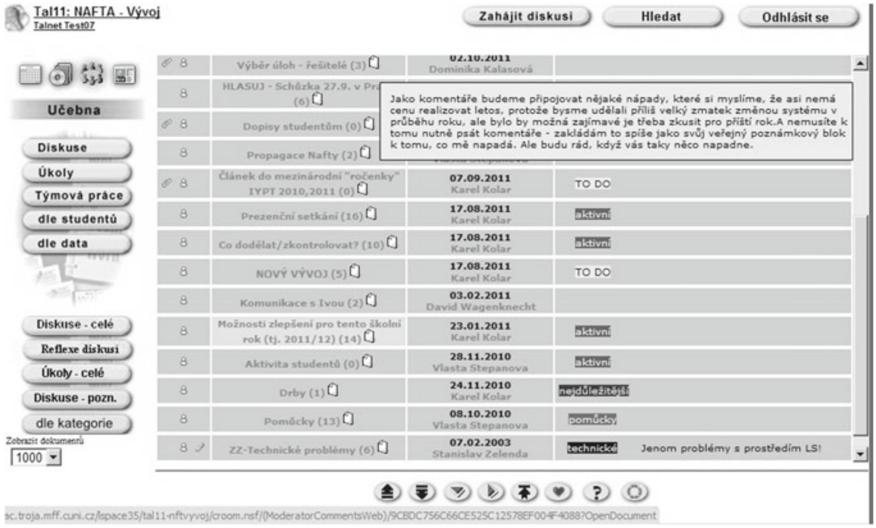


Fig. 65.2 Special tools and meta layers were developed for massive communications

In our seminars for school teachers they are (in their words) "fascinated" when see pupils–participants of the course-working and can interact with them. The teachers mention strong positive feelings and a will to look at their teaching and pupils in a new perspective.

65.3 Conclusions

The recent experience and results, despite the limits given by computer mediated communication, shows that there are students who are willing and able to go deep in physics, does not matter whether we call them gifted or not (interestingly some seems to be independent learners). They have a great potential to communicate their live and authentic experience, gained in "doing physics with a real physicist" or in "physics enquiry with mates sensitively guided by a physicist", with their classmates in an acceptable way and the language of the group knowledge and age level.

In Talnet we offer study or developmental trajectories that include team's enquiry activities and simultaneously we open gates to "real physics" (as complement to school physics) not only to interested potentially gifted pupils (nominated by teachers) pupils but also their mates.

For higher effect of gifted pupils' influence in the classroom we offer teachers special opportunities (seminars with online support) to get experience in nomination of gifted pupils and get inspirations for involving pupils participating in Talnet to "popularize physics" in their classrooms or schools.

There is still another critical point: limited possibilities and motivations of physicists to participate in such a complicated and time consuming “intergeneration communications”.

References

1. Renzulli JS (1978) What makes giftedness? Re-examining a definition. Phi delta kappa 60:180–181
2. FO - Physics Olympiad (2011) http://fyzikalniolympiada.cz Accessed 20 Nov 2011
3. SOC - Secondary School Expert Activity (2011) http://www.soc.cz Accessed 20 Nov 2011
4. Talnet (2011) http://talnet.cz Accessed 20 Nov 2011
5. Betts G, Neihart M (1988) Profiles of the gifted and talented. Gifted child quarterly, NAGC
6. Mcleod SA (2010) Simply psychology; zone of proximal, development

Chapter 66
On INFN 2010 Physics Popularization School: Video Report

Santo Reito

Abstract In this short video we report on the four days INFN school "Comunicazione e Divulgazione della Fisica 2010" held in Perugia (Italy) on November 9/12. This meeting was conceived in the frame of a bigger training programme for INFN people interested in the popularization of physics. It was a full immersion school with traditional lectures, laboratories, working group and some sessions called "A tavola con lo scienziato" (i.e. dinners in the course of students could discuss with active scientist expressly invited). In the video are included summaries of the lessons and interviews of the teachers. In particular we paid some special attention on the "A tavola con lo scienziato" event. During the lunches the INFN school students met some teachers and student of the secondary school and ordinary people. This video is not just a chronicle of an event, it collects ideas, experiences and expectations of the people who shared this happening throughout four days. We were particularly interested in the opinions of the students and the teachers interviewed during "A tavola con lo scienziato" event because one of our goals was the understanding of the "learning science problem" directed to decrease the distance gap between science and youth in the era of learning in the wild and citizen-science projects.

The science plays a crucial role in the modern society. How do people learn science ? Nature in an editorial, titled "Learning in the wild" told us that much of what people know adout science is learned informally.

The seemingly endless debate about how to improve US science education seems to make the tacit assumption that learning happens only in the classroom. As a result, the arguments tend to focus on issues such as curricula specifying, say, what information pre-college students should be expected to learn at each grade level and, as in US President Barack Obama's recent proposals to reform the No Child Left Behind policy, on the best way to hold schools to rigorous standards of student achievement.

S. Reito (✉)
Istituto Nazionale Fisica Nucleare, Frascati, Itlay

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_66,

However, researchers who study learning are increasingly questioning this assumption. Their evidence strongly suggests that most of what the general public knows about science is picked up outside school, through things such as television programmes, websites, magazine articles, visits to zoos and museums and even through hobbies such as gardening and bird watching.

This process of ’informal science education’ is patchy, ad hoc and at the mercy of individual whim, all of which makes it much more difficult to measure than formal instruction. But it is also pervasive, cumulative and often much more effective at getting people excited about science and an individual’s realization that he or she can work things out unaided promotes a profoundly motivating sense of empowerment.

Nowadays *science popularization* plays a crucial role. The *science popularization* is an attempt to reduce the distance standing between science specialists and the public.

Science popularization is interpretation of scientific information (science) intended for a general audience, rather than for other experts or students. The primary objective of the *popularization of science* is to increase public understanding of science. Since a few years ago, INFN has undertaken a strong program for training its Scientists in Science Outreach and supports many Outreach initiatives. In 2009 started a four days INFN school “Comunicazione e Divulgazione della Fisica 2009”. Goal of the Course was to provide participants with tools and insights instrumental to fulfil basic needs of Communication and Outreach about Science in general and Physics in particular.

- 4 days—full immersion
- 20 h of plenary session lectures:
  Communication Psychology
  Targets of Outreach
  Writing Lab
  Communication Protocols
  Communication and the WEB2.0
- 4 h of debate workshops
- 12 h of hands-on work organized in Working Groups
- 12 h of “socialising” communication: dining with a scientist.

In this school had taken part 40 participants (INFN researchers), 7 coordinators, 20 teachers and 50 guests. During the lunches the INFN school students met guests for “socialising” communication. The hands-on work was organized in 7 Working Groups. The school students simulated real events.

Working group 1:
Advertising and publishing products;
Working group 2:
Radio and TV communication;
Working group 3:
Physics on the street;
Working group 4:

Open Lab;
Working group 5, 6, 7:
LHC Physics, neutrins, dark matter.

In 2010 the second edition of school INFN "Comunicazione e Divulgazione della Fisica 2010" was held in Perugia (Italy) on November 9/12. Aim of the School "Comunicazione e Divulgazione della Fisica" was to provide participants with basic tools and knowledge about verbal and non-verbal communication to be used in Outreach of Science in general and Physics in particular:

- 4 days—full immersion
- 16 h of plenary session lectures
- Written communication; advertisement communication; non-verbal communication; Physics on the street; Physics outreach in the WEB 2.0 era; Science communication: features and processes; Physics outreach in the schools and learning processes; Physics outreach and adult involvement: goals and results; the Perugia Science Festival; visual Communication: from gestalt to coordinated image design; the role of social media in the daily work of scientists; virtual environments for teaching and outreach.
- 4 h of debate workshops
- 12 h of hands-on work organized in Working Groups
- 12 h of "socialising" communication: dining with a scientist.

In this school had taken part 29 participants (INFN researchers), 4 coordinators, 17 teachers and 40 guests. During the hands-on work the school students realized a web site for child students. These schools have been a very good experience for partecipants, for teachers and in particular for guests.
Particularly appreciated is the "dining with a scientists" session, which sees a quite lively participation of students and guests. The next edition will be in Torino in 2012. Follow this link you can watch video about "Comunicazione e Divulgazione della Fisica 2010": http://www.youtube.com/watch?v=TviDcNgLLKg

Reference

1. (2010) Learning in the wild. Nature 464:813–814

Chapter 67
Popularisation of Physics in the Wild

Beatrice Boccardi, Michela Fragona and Giovanna Parolini

Abstract Science popularisation is important to inform tax payers on how their money is used and to inform the general public of the progresses of research. The web 2.0, comprising the 3D virtual worlds, offers science and physics popularisation new communication channels and tools constituting, as a whole, a large part of "learning in the wild", with the consequent *citizen science* projects. Tags: physics popularisation, informal learning, web 2.0, virtual worlds, Second Life.

67.1 Introduction

Science and physics popularisation is important as a means to promote further investments through the information of the tax payers on how their money has been used and on the improvements that scientific research can bring in their lives, so the communication of scientific knowledge to the general public is both a consequence and a precondition for the ongoing scientific-technological progress. Science popularisation gives the non-scientists an up-to-date knowledge, resulting in the constitution of a cultural environment where science is valued and supported.

The web 2.0 has not just quickened or simplified the existing procedures, it has opened new domains to science popularisation and to the active involvement of the non-scientists. Seminars, lectures, science museums, etc., are the traditional formats for science popularisation, while the web 2.0 provides tools that make it accessible to everybody. The cultural revolution of the Internet and the web 2.0 is expressed

B. Boccardi (✉)
Liceo Scientifico A. Labriola, Napoli, Italy
URL: www.liceolabriolanapoli.com

B. Boccardi · M. Fragona · G. Parolini
Esplica no-profit, Laboratorio per la Divulgazione Culturale e Scientifica nell'Era Digitale, Villafranca Di Verona, VR, Italy
URL: www.esplica.it

B. G. Sidharth et al. (eds.), *Frontiers of Fundamental Physics and Physics Education Research*, Springer Proceedings in Physics 145, DOI: 10.1007/978-3-319-00297-2_67,

in the triad of information-communication-sociality. The multimedia on the web, with music, videos and games has made it predominant for sociality, knowledge and learning.

In this paper, we will describe in particular the experiences of physics popularisation in the virtual worlds. These are a specific instance of the cyberspace, true 3D social networks where the users interact directly through their graphical representation, the avatar, and "build" the world. Though they may appear like videogames, they constitute places where new forms of science popularisation are experimented.

The definitions *learning in the wild* and *citizen scientist* — not created, but made popular by *Nature* — summarize the change in progress in learning and the popularisation of science: "Much of what people know about science is learned informally" (*Learning in the wild*, 2010.) [1]. In this context, wild stands for *informal* learning, without fixed schedules and programs. The result of this diffused learning mode is the *citizen scientist*, contributing to the *citizen science* projects with observations and data gathering, mostly thanks to the digital technologies.

67.2 Learning in the Wild and Citizen Science

Learning in the wild, made up of freely chosen experiences without an express formative aim, is "often much more effective [than formal learning] at getting people excited about science" (*Learning in the wild* 2010) and involves people of every age, so it can be the foundations of school learning and of lifelong learning.

Forms of science popularisation and citizen science have existed since the XVIII century, but the web 2.0 has given the *science enthusiasts* the tools to communicate among themselves wherever they are based in the world, in a continuous junction between the physical world, where the facts are observed, and the cyberspace, where the data are stored and discussed, strengthening free learning through sharing and giving a sense of accomplishment and participation.

The wilderness is rich of engaging events, science festivals, plays like *Copenhagen* (Frayn 1998) [2] and *The Children of Uranium* (Greenaway 2006) [3], popular essays and fiction of science. The web hosts an amount of scientific sites, blogs and magazines. *Symphony of Science* (Boswell 2011) [4] presents famous scientists "singing" lyrics inspired to various aspects of scientific knowledge, Wikiversity collects resources for formal and informal learning.

Yet the direct participation of the general public to experiences of science popularisation is fostered by the citizen science projects. The oldest of these is *SETI@home* (SETI 2011, *a*) [5] - since 1999, 2,000,000 members in the world, almost 500 in Italy (SETI 2011, *b*) [6]. SETI analyzes the narrow-bandwidth radio signals from space, which do not known occur naturally, "so a detection would provide evidence of extraterrestrial technology". It covers "greater frequency ranges with more sensitivity" exploiting the virtual super computer composed of the members' computers connected to it. *Galaxy Zoo* uses the collaboration of the citizen scientists to classify the galaxies. The participants watch images of galaxies and choose among some features that mark out different types of galaxies (Galaxy Zoo 2011) [7]. *GridRepublic Volunteer Computing* is a container of citizen science projects (GridRepublic 2011) [8]. The members install a software on their computer, through which it works on research when the machine is not in use. The *NASA* recognises that they the citizen scientists "have helped to answer serious scientific questions" (NASA *a*) [9] and confirms the importante of constituting a supporting environment for the promotion of further investments.

The radon gas measurement is a precursor example of citizen science in Italy. As the decay product of radium, radon is present in the whole earth's crust and in the building materials taken from it. This natural source of radiation needs constant testing, in public and residential buildings, where the measuring kits are installed (ISPRA 2004) [10]. The *School Radon Survey* of Trieste university with INFN (Università di Trieste 2004) [11] in the years 2003–2004 engaged 4 schools in the region Friuli Venezia Giulia and about 200 students to test radon outdoor and in their homes. The *ENVIRAD-SPLASH* experiment of the INFN involved 110 schools all over Italy from 2002 to 2009. The students tested the schools for radon, calculated the concentrations on the basis of the readings of the detector, prepared the introductory sessions for new participants and ran the internet connection (Roca 2010) [12].

The *EEE-Extreme Energy Events Project*, began in 2005 by the Centro Fermi with the CERN, INFN, MIUR and EMFCSC, studies the cosmic rays with the involvement of students from all over Italy (Zichichi 2005) [13]. So far, 32 schools have participated. After the preliminary visit at the CERN, to assemble the telescope MRPC detector for the observation and the measurement of the cosmic muons, the engagement of the whole school is necessary, with turns to monitor the apparatus. The Liceo Cagnazzi at Altamura, Bari (Liceo Cagnazzi 2011) [14] is a positive example: information on the school website, an article in the local magazine, discussion of the results in a public Conference. In these projects the students learn through a hands-on approach, carry out true scientific project, build the measuring tools, study radioactivity in the field, apply the experimental method, evaluate the limits and validity of the measurements. The radon projects also improve the knowledge of the territory. These projects constitute physics laboratories and excellence courses and are the junction between the popularisation of physics and formal learning. While the schools need to change programs and methods, the general public is caught in a direct participation, through the installation of the radon measuring kit at home, or being invited to a conference.

67.3 Virtual Worlds

The virtual worlds, or metaverses, invented for playing, have turned out to be the most complete environments on the web to create and share culture. The features common to all of them are tridimensionality, the user present as an *avatar,* synchronous communication through text-chat and voice-based chat, but a few others are necessary to use them for the popularisation of physics: no pre-arranged storyline, no characters to interpret or levels to attain, user-created content. To the persistent base of the world designed by the developers, the users add regions (*lands*) designed by themselves - ancient Rome, Dracula's castle, islands devoted to culture, art, science - and create groups according to their specific interests, many of which active also on other social networks, websites and blogs.

The virtual worlds are fundamental channels for the popularisation of science and physics, for the users attending them. The majority belong to an age level higher than in the real games, are educated, have various cultural interests. The virtual worlds are consistent with the emerging trends, the gender asymmetry in disfavour of women reported by other projects of dissemination is absent and they are fit for the differently abled, which they foster socially and physically, solving some motor or expression difficulties, in a place where it is easier to contact people and make friends.

The social interaction with the avatars present in the same place, produces the sensation of "feeling present", just like in the physical world (Baumgartner 2008) [15]. The creativity prompted by building and scripting makes these worlds *generative* of new cultural contents. The user as the content creator and the social interaction in presence bear the *immersivity* of the environment, a positive emotion that prepares the avatar to a more active participation than in Real Life (Suler 1996) [16]. Universities, non-profit organization, scientific groups give courses, make conferences and research. All this does not replace the activities in the physical world. Rather, they are expanded and given a new presentation. The playful dimension, even when it is not the immediate aim of communication, helps the transmission of knowledge and learning (Flanagan 2010) [17].

67.4 Science Popularisation in the Virtual Worlds

In the metaverses science popularisation may happen in forms different or impossible in the physical world, with objects where the game is implicit. *Building* and *scripting* enable you to represent physics concepts or events through interactive objects that are themselves a form of play - the avatar can get into a black hole or "become" a subatomic particle, a playful dimension that is part of the appeal of science popularisationin the metaverses. The science event "within reach of teleport" encourages the participation of those who up to that moment had left it out, for want of interest or prevented by everyday circumstances. The metaverses facilitate the discovery of cultural fields so far ignored or considered intellectually distant. At the moment, the 3D world *Second Life* appears as the most complete and flexible for the popularisation of science. The group *Real Life Education in Second Life* has 4,800 members, educators in Real Life. Universities and institutes inworld are Stanford University, Media PlayLab (Singapore), Infolit iSchool (UK), NOOA - to mention but a few.

Among the scientific groups, physical and astrophysical topics are prevailing. The NASA land in Second Life exhibits space and terrestrial environments that illustrate its fields of research - glaciers, underwater laboratories, volcanoes. Through its website, the NASA provides a guide for the games inworld, for students of different age groups: they deal with the building of space rockets and their importance in the development of space research, or with the students working at the organization of a mission on the Moon and the study of its habitat (NASA *b*) [18]. Some of these games are suitable to promote a popular talk, as the playful dimension attracts also the adult receivers of science popularisation, who can play with the interactive objects or build some themselves. *MICA-Meta Institute for Computational Astrophysics* gathers international Real Life astronomers and astrophysicists to study the application of the virtual worlds and technologies to astronomical research. They give public lectures on Saturdays, which can be listened again from their website (MICA) [19]. *Science Friday* presents a course on the basic concepts of astronomy and optics, with interactive models of astronomical objects. It broadcasts inworld the scientific programs of NPR-National Public Radio. The avatar audience can ask questions to the guests in the Real Life studio. The anglophone *Kira Cafè* is a bar to socialize and discuss interdisciplinary scientific subjects. *Exploratorium* is a science museum inspired to the Real Life homonymous museum in San Francisco, CA. It hosts a presentation of nuclear physics, experiments of physics, astronomy, on perception, and an interactive Sun: the avatar travels to the Earth riding a sun ray. An installation of the Earth-Moon system shows how a solar eclipse appears from the Moon. *Second Life Physics Lab*-SLPL, founded by a physics professor in Real Life, exhibits a cannon shooting cannon balls following two different trajectories, either according to Buridanian or Newtonian physics: this shows how the tools of the virtual world can be used for science popularisation.

In the Italian community we can see the emergence of several groups that carry out extemporary activities or well established projects lasting for several years now. Italy has taken part in *Avatar Project* 2009–2011, to study the use of the virtual worlds as a classroom laboratory, co-financed by the European Commission. High school teachers and students experimented with the features of the virtual world: immersivity, creativity, learning made interactive and easier by the 3d ambient (Avatar) [20]. Italian groups dealing with teacher formation are *INDIRE* (INDIRE) [21], with courses of building, maths and Italian l2. *SecondAnitel*, association of e-tutors, with courses on teaching with the digital technologies (ANITEL) [22]. *Imparafacile* offers courses on e-learning and makes research on teaching methods in the virtual world (Imparafacile) [23]. *Immersiva.2life* is engaged in the popularisation of culture and music through live concerts and forums on topical issues. *EsplicaSL* is the virtual representative of Esplica no-profit (Esplica) [24], to link "the immersive creativity of the metaverse and the physical reality" and to promote Second Life as an aggregation place for the diversely abled. In 2010 it realized "Let's read in Second Life", a twinning with the eBookFest in Real Life (Cross-Universe) [25] and this year has presented its accomplishments at the exhibition for the 4th anniversary of Nonprofit Commons, a similar group in Second Life.

Second Physics, group for the popularisation of physics and science, from 2008, is an example of the many activities possible in the metaverse in this field. *Scienza on the Road*, in collaboration with Immersiva2Life, is an itinerant series of one-hour conversations started in 2009. It has won fame by now, with more than 4,000 avatars attending the lectures in more than a hundred different Italian lands. At the moment

Second Physics has more than 1,500 members – one of the highest number among all the scientific groups – comprising many non-Italian-speaking avatars, who receive the English translation on a special board. So far, the conversations have dealt with Physics, Astronomy, Particle Physics, Mathematics, Acoustics, Optics, Medicine, Cybernetics, Literature for Science Popularisation. The audience can ask questions and discuss with the lecturers. The conversations are broadcast in streaming web and announced in blogs, online magazines and on FaceBook. Second Physics has taken part to the conference Virtual Worlds Best Practices in Education 2010, with school and university groups active in Real and Second Life, with a poster session and an article on the Journal of Virtual Studies (Boccardi 2011) [26]. With SL Art,in 2010 Second Physics organized *Scien&Art*, a contest for the artists aiming at expressing "their creativity through scientific concepts" and the scientists trying to communicate scientific ideas in a creative way, which put into evidence the link between art and science. A 10-hour course *Beyond the Third Dimension*was held in 2009 and a course on the cosmic rays will be given in *Second Campus*, the project started this year with two 10-hour courses, on the literary fiction for science popularisation and on the graphics software*Blender*. The *Science Café* opened in 2010 at Second Physics, without a fixed schedule. On the occasion of scientific events in Real Life, an expert in the subject is invited to analyse the sense and the value of the piece of news. Besides the scheduled events, the land of Second Physics presents installations on the physicists' researches and anecdotes, as prompts for quizzes that, while attracting the visitors' attention, drive them to some googling to solve the enigmas. Other fixed installations are the optical illusions, and the presentations of past courses and exhibits. Finally, there are scientific sculptures, from Leonardo to the DNA helix to the cosmic rays.

We end this description of the activities of Second Physics mentioning the support group to the free online course on AI by S. Thrun and P. Norvig with Stanford University, set up in Second Campus Hall to discuss the lessons weekly. *Learning in the wild* arises new cultural interests and the virtual world gives a space to share them.

67.5 Conclusions

Learning in the wild is made up of freely chosen activities, without a direct formative aim. The web 2.0 has given the science enthusiasts the tools to contribute to scientific research through the citizen science projects making observations, collecting data, discussing with other citizen scientists online.

The metaverses are showing great possibilities for science popularisation, thanks to the direct relationship between the divulger and the receiver through the avatar, both a playful and "serious" mode, that makes learning easier.

This has changed the way we learn and formal learning, in its turn, will have to change programs, methods, evaluation systems, to keep up with tools that are not simply "technologies", but represent a true cultural revolution where sharing, playing, free choice, are the keywords.

References

1. Learning in the wild, Nature. Intl Weekly J Sci. www.nature.com/nature/journal/v464/n7290/full/464813b.html (2010).
2. Frayn M (1998) Copenhagen. Anchor Books, New York
3. Greenaway P, Boddeke S (2006) The Children of Uranium. Charta, Milan-New York (bilingual edition in English and Italian)
4. Boswell D, http://www.symphonyofscience.com
5. SETI a http://seti.berkeley.edu
6. SETI b, http://www.mconsult.it/seti-at-home.htm
7. Galaxy Zoo, http://www.galaxyzoo.org/
8. GridRepublic Volunteer Computing, http://www.gridrepublic.org/
9. NASA a, http://science.nasa.gov/citizen-scientists/
10. ISPRA, Istituto Superiore per la Protezione e la Ricerca Ambientale (2004) http://www.apat.gov.it/site/it-IT/Progetti/Legge_93_01_-_Disposizioni_in_campo_ambientale/Progetti_a_gestione_APAT/Progetto_13/
11. Universitá di Trieste (2004), Progetto Radon School Survey, http://www.fisica.units.it/didattica03/download/progettoradon.pdf
12. Roca V (2010) I risultati di un progetto scientifico didattico, Atti del 3 ° Convegno "Comunicare Fisica, e altre Scienze", Frascati, 12–16 Aprile 2010.
13. Zichichi A (2005) Progetto "La Scienza nelle Scuole". EEE-Extreme Energy Events, http://www.centrofermi.it/eee/documents/EEE.pdf
14. Liceo Cagnazzi, http://www.liceocagnazzi.it/. The material produced can be seen here: http://www.ba.infn.it/regano/EEE/
15. Baumgartner T, Speck D, Wettstein D, Masnari O, Beeli G, Jäncke L (2008) Feeling present in arousing virtual reality worlds: prefrontal brain regions differentially orchestrate presence experience in adults and children. Frontiers in Human Neuroscience, 2, 1–12, http://www.frontiersin.org/human_neuroscience/10.3389/neuro.09/008.2008/abstract, then click on fnhum 02–008_SUPPLEMENT.pdf for PDF
16. Suler J (1996) The Psychology of Cyberspace, revised 2006. Online treatise, http://users.rider.edu/ suler/psycyber/psycyber.html
17. Flanagan T, Delphin G, Fargis M, Lexington C (2010) Arts, mathematics and physics in second life. Paper for the Aplimat conference 2010, http://journal.aplimat.com/volume_3_2010/Journal_volume_3/Number_1.pdf
18. NASA b, http://search.nasa.gov/search/edFilterSearch.jsp?empty=true and http://core.nasa.gov/
19. MICA, http://mica-vw.org/wiki/index.php/Meta_Institute_for_Computational_Astrophysics and http://www.mica-vw.org/wiki/index.php/Popular_Talks
20. Avatar, http://www.avatarproject.eu/avatar/
21. INDIRE, http://www.secondlearning.it/
22. ANITEL, http://www.anitel.org/anitel/
23. Imparafacile (2011) http://imparafacile.ning.com/, accessed Dec 2011
24. Esplica, http://www.esplica.it
25. Cross-Universe EBookFest, http://www.esplica.it/images/stories/IMMAGINI_ARTICOLI/leggiamo_in_second_life.pdf
26. Boccardi B, Fabbri FL, Fragona M, Nardone A, Parolini G, Dos Santos RP (2011) Second physics: popularisation and outreach of science in the Italian second life community, http://ejournal.urockcliffe.com/index.php/JOVS/article/viewFile/20/3