Euro Jnl Phil Sci (2017) 7:127–150
DOI 10.1007/s13194-016-0149-x

ORIGINAL PAPER IN PHILOSOPHY OF SCIENCE

Using conceptual spaces to exhibit conceptual continuity through scientific theory change

George Masterton [1] · **Frank Zenker** [1,2,3] · **Peter Gärdenfors** [1,4,5]

Received: 25 November 2015 / Accepted: 29 April 2016 / Published online: 23 June 2016

Abstract There is a great deal of justified concern about continuity through scientific theory change. Our thesis is that, particularly in physics, such continuity can be appropriately captured at the level of conceptual frameworks (the level above the theories themselves) using conceptual space models. Indeed, we contend that the conceptual spaces of three of our most important physical theories—Classical Mechanics (CM), Special Relativity Theory (SRT), and Quantum Mechanics (QM)—have already been so modelled as phase-spaces. Working with their phase-space formulations, one can trace the conceptual changes and continuities in transitioning from CM to QM, and from CM to SRT. By offering a revised severity-ordering of changes that conceptual frameworks can undergo, we provide reasons to doubt the commonly held view that CM is conceptually closer to SRT than QM.

Keywords Radical theory change · Conceptual space · Classical mechanics · Special relativity theory · Quantum mechanics

1 Introduction

On the basis of an analysis of the phase-space formulations of three physical theories this paper argues for two theses: (1) Conceptual space models of conceptual

✉ George Masterton
george.masterton@fil.lu.se

[1] Department of Philosophy & Cognitive Science, Lund University, Lund, Sweden

[2] University of Konstanz, Konstanz, Germany

[3] Slovak Academy of Science, Bratislava, Slovak Republik

[4] Wissenschaftskolleg zu Berlin, Berlin, Germany

[5] University of Technology Sydney, Ultimo, Australia

frameworks bear a striking resemblance to the phase-spaces of physical theories and that this resemblance is strong enough to warrant the thesis that phase spaces *are* conceptual spaces. (2) By identifying the types of changes in the underlying conceptual space that occur when one theory supplants another, the degree of conceptual continuity between physical theories and their successors becomes apparent.

We introduce conceptual spaces as our modelling tool (Section 2), show how these models apply to scientific conceptual frameworks (3), introduce differentially severe changes to assess the comparative continuity through framework change (4), propose viewing the phase-spaces of physical systems as conceptual spaces (5), and then exhibit the phase/conceptual spaces of Classical Mechanics (CM) (5.1) and Quantum Mechanics (QM) (5.2). Detouring to compare the algebras of observables of these two (5.3), we finally present the phase/conceptual space of Special Relativity Theory (SRT) (5.4). On this basis, we gauge the continuity that obtains between their conceptual frameworks (6), then provide a discussion (7) and conclusions (8).

2 Conceptual spaces

Conceptual spaces (Gärdenfors 2000) model representations of cognitive systems, and contrast with two other common ways of modelling such representations. One of these alternatives is the *symbolic approach* (Pylyshyn 1986), which starts from the assumption that cognition is essentially computation involving symbol manipulation. The other alternative is *associationism*,[1] where associations between different kinds of information elements carry the main burden of representation. Conceptual spaces are not parts of a symbolic system with a syntactic or logical structure. Rather, they are topological (typically, but not necessarily, Finsler) manifolds that can be analyzed into their constitutive quality-dimensions, and how these are structured.

Part of the structuring of the dimensions is their allotment into domains; to represent a single concept, one often needs multiple dimensions. For instance, physical extension requires three spatial dimensions *integrated* into a single domain while temporal extension has historically, in physics, been considered a *separable* dimension that forms its own domain. Another part of this structuring is the geometry these domains are endowed with. For example, traditionally, physical space was taken to be Euclidean while physical time was taken to be isomorphic to the positive real number line. With the advent of relativity theory these same dimensions were all integrated into a single space-time domain with a Minkowskian geometry.

Raubal (2004) provides a rigorous formalization of conceptual spaces as vector spaces, defined as $C^n = \{(c_1, c_2, c_3, c_4 \ldots c_n) : c_i \in \mathbf{C}\}$, where each c_i is a quality domain $D^n = \{(d_1, d_2, d_3, d_4 \ldots d_n) : d_i \in \mathbf{D}\}$ and each d_i is a quality dimension. Since vector spaces have metrics, there will be a well-defined measure of qualitative similarity between points in the space provided all the dimensions have the same relative unit of measurement. Raubal (ibid.) proposes that all vectors in a conceptual space be z-transformed to ensure that they all have the same unit of measurement. The semantic distance between two instances of a concept—two particulars, that is, which instantiate

[1] Connectionism (Bechtel and Abrahamsen 2002) is a special case of associationism that models associations using artificial neuron networks.

the concept—is then the Euclidean distance between the z-transformed vectors of those instances. Defined as vector spaces, conceptual spaces are amenable to various transforms and mappings that formalise such notions as ‘change in the relative importance of dimensions’ (by introducing weights on the components of the vector space), or ‘how one can alter one conceptual space to generate another’, and so on. In the rest of this paper we will talk about these notions informally. However, it is to be understood that our claims and assertions can be formalised.

Psychologists use various empirical tests to determine whether two dimensions are treated as separable or integral (see, e.g., Maddox 1992). One such test concerns the metric used to measure dimensions. A commonly applied rule is that if the Euclidean metric ($ds^2 = dx^2 + dy^2$) fits the data best, the dimensions are classified as integral, while if the city-block metric ($ds = |dx| + |dy|$) fits the data better, then they are classified as separable. In the context of scientific theories, we can express this by saying that dimensions are integral if they are measured in the same way, while they are separable if measured by distinct methods.

Another separability/integrality test is whether the perception of values on one dimension *interferes* with perception of values on the other (Maddox 1992). In psychology, this test typically generates the same results as the metric test, provided that interfering dimensions are integral to a domain and non-interfering dimensions are separable. As we show in greater detail below (Section 6.1), however, although position and momentum remain two separable domains in the context of quantum mechanics (QM), measurements of the former nonetheless interfere with measurements of the latter.

For instance, in classical mechanics (CM) the commutator between position and momentum in the x-direction $[x, p_x]$ is set to zero, since the order in which one measures position and momentum is taken to be irrelevant on experimental grounds. Indeed, measurements of position and of momentum along any axis do not interfere with each other in CM, making position and momentum *non-interfering* domains in CM. But this commutator is non-zero, and such interference *is* present, in QM. Yet it is *prima facie* implausible that momentum in a direction and position along that direction should be treated as a single integral domain in QM. We will return to this conundrum below.

As asserted above, conceptual spaces come with a metric, that is, a measure of distance by which qualitative similarity can be gauged (see Fig. 1). Given a space with particular quality dimensions and a structure, or a geometry, an object’s qualitative state can be represented as a point or, where there is vagueness, as a region in that space (Douven et al. 2011).[2] Two objects, hence, are similar to the extent that they are close together in the space.[3]

A further aspect of conceptual spaces to observe—providing a reason why a scientific theory cannot be completely modelled as a conceptual space—is that conceptual spaces normally lack dynamics. That is, while one can represent the qualitative

[2] Given a temporal dimension, the qualitative evolution of an object can be represented by a trajectory through the conceptual space.

[3] Quality dimensions can be continuous or discrete, and discrete spaces can also have metrics allowing measures of similarity. For instance, if three objects, A, B, and C, differ along two discrete dimensions such that A differs from B only along one dimension, but A differs from C along both dimensions, then A is more similar to B than A is to C, which is to say that the distance from A to B in the (Euclidean) plane of these dimensions (taking the value 1) is less than the distance from A to C (taking the value $\sqrt{2}$) in that plane.

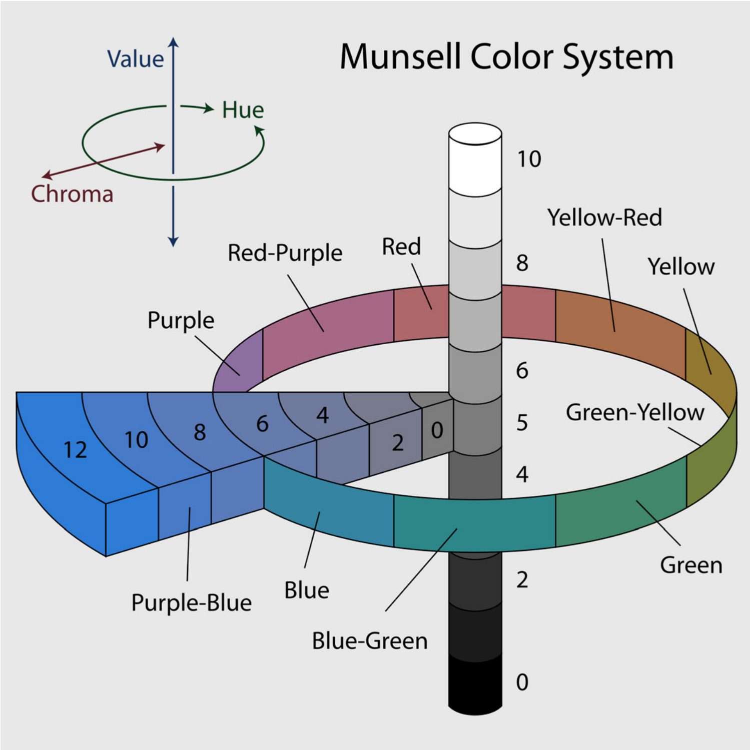

Fig. 1 The Munsell colour space. The metric by which we gauge the qualitative similarity of two colours in this space is the cylindrical metric: $ds^2 = dr^2 + r^2 d\psi^2 + dz^2$, where dr, $d\psi$ and dz are, respectively, the radial (chroma), angular (hue in radians) and axial (value) displacements

evolution of objects/systems using trajectories through the relevant conceptual space, principle constraints on those trajectories are, as a matter of contingent fact, not normally specified. Arguably, it is the addition of such principles to a conceptual space that generates a genuine *theory* of the systems located in that space that could be used to predict/postdict their future/past qualities from their initial qualitative states. For this, as well as other reasons such as the possibility of distinct theories sharing the same conceptual framework, conceptual spaces are not sufficiently fine grained to model theories themselves. As our thesis is that conceptual spaces can be used to model the conceptual frameworks within which theories are formulated, and not the theories themselves, this limitation does not trouble us.

3 Modelling scientific conceptual frameworks

An empirical theory always presupposes, but, as we just saw, remains distinct from, a specific conceptual framework that provides the magnitudes, or *dimensions*, on which the formulation of the theory depends. While any conceptual framework can be modelled as a conceptual space, the frameworks presupposed by scientific theories are particularly amenable to such modelling (Gärdenfors and Zenker 2011, 2013; Petersen and Zenker 2014; Zenker 2014; Zenker and Gärdenfors 2014, 2015).

Apart from the concepts familiar to us in sensory perception and covered in psychological research—like *colour*, *weight* or *length*—physics introduces theoretical magnitudes such as *mass*, *force*, and *energy*. In the SI units system, magnitudes such as mass are taken to be basic, while others such as energy and force are derived. The basic magnitudes have stipulative definitions of various kinds. The derived dimensions are defined in terms of the basic dimensions. Whether basic or derived, any such magnitude can be modelled geometrically. Newtonian mass, for instance, can be represented as a one-dimensional space isomorphic to the non-negative part of the real number line, and Newtonian space as a three-dimensional Euclidian (vector) space. We give further examples below.

As indicated, some magnitudes/qualities of an object require only a one-dimensional space to represent them, while others require multi-dimensional spaces. Recall that dimensions are said to be *integral* if, to fully describe an object with respect to a given quality, a value must be assigned to each dimension (see Section 2). For instance, the three spatial dimensions length, width and height are integral because an object cannot have physical extension without values for all three dimensions being assigned. Similarly, one has not specified the force acting on an object in CM, until one has chosen a spatial frame and specified the component forces acting parallel to the three axes of that frame. Dimensions that are not integral are said to be *separable*; for example, the shape and hue dimensions, or the mass and charge dimensions in Classical Electro-Dynamics.

The choice of domains, and their constitution, is not uniquely determined because the organization of dimensions into domains depends on conventions regarding which of a theory's dimensions are taken as *basic* and which as *derived*. Outside of physics, conceptual architecture is generally not laid out with sufficient precision to easily resolve which subset of our various dimensions shall count as basic and which shall be defined in relation to them. But in the sciences, and particularly in physics, this division is made on conventional or pragmatic grounds, often in theory formulation-specific ways. Where system constraints are non-holonomic,[4] for instance, one may prefer the Newtonian formulation of classical mechanics (with position, time, force and mass as the basic domains), while the Lagrangian formulation (with position, mass and time as basic domains) is generally easier to work with for systems with holonomic constraints.[5]

As scientists develop new theories they invariably do so *from within* a conceptual framework and these can be modelled as conceptual spaces. It follows that tracing

[4] Holonomic system constraints can be expressed by a zero-valued function like $f(x_1, x_2, x_3, \ldots, x_n, t) = 0$. It is characteristic of holonomic constraints, moreover, that they do not have velocities as arguments. For instance, the constraint $x^2 + y^2 - L^2 = 0$ on the position of a 2D pendulum of fixed length L is holonomic. A constraint that cannot be expressed in the above form is a non-holonomic constraint.

[5] Those inclined to grapple with scientific realism, of course, take the question 'Which dimensions are basic, which derivative?' to be one of principle. When Isaac Newton introduced the force domain into mechanics, for instance, many natural philosophers advocated—on grounds of ontological suspicion—a reformulation that made force a defined domain. Similarly, those following Niels Bohr in claiming that all measurements are ultimately spatio-temporal determinations, from which the values along other dimensions are inferred, view space-time as the basic domain. Whether conventional or principled, the division itself is (with the possible exception of such "Babylonian" physicists as Richard Feynman) present, and important to observe.

continuity between the conceptual spaces that are presupposed by two, or more, theories is a tracing of the continuity in the conceptual frameworks of those theories. As the next section exhibits, accepting this approach offers a detailed rebuttal of the "revolutionary" account, famously proposed by Kuhn (1970),[6] of scientific change as a *ruptured* historically process.

We are not the only parties interested in 'demonstrating continuity through theory change'. In particular, structural realists share this wider aim—though their primary interest is in demonstrating *structural* continuity. For the purposes of comparison, structural realists come in two flavours: those who describe theory structure syntactically in a formal language (e.g., Schurz 2009; Maxwell 1970) and those who describe this structure algebraically in order to ascertain the theory invariants (e.g., Suppes 2002; Thébault 2014). Though either flavour faces its own issues, the latter is closest to our approach. From our perspective, the "algebraists" describe, among other things, the structure of the conceptual spaces we are interested in, and so offer part of what we seek to provide when extending the remit of conceptual spaces to the sciences.

Indeed, there is significant overlap between Thébault's (2014) 'state space and observables algebra'-framework for ontological structural realism with respect to scientific theories, on one hand, and our 'phase space is conceptual space'-approach to the conceptual frameworks presupposed by those theories, on the other. The crucial difference lies in the goals being pursued. Ontological structural realists seek to single out those parts of our *theories* that can be safely reified despite the pessimistic meta-induction and under-determination arguments. By contrast, our goal is to demonstrate that there is *conceptual* continuity through "revolutionary" theory change despite semantic holism.

4 Modelling continuity in conceptual change

Evaluating conceptual continuity through theory change, of course, presupposes a way of describing framework change. Gärdenfors and Zenker (2011, 2013) have provided a taxonomy of conceptual space changes, including a ranking by severity of these changes, and a selection of historical examples (also see Petersen and Zenker 2014). Here, we merely give a brief overview.

In order of increasing severity, or of diminishing conceptual continuity, Gärdenfors and Zenker (2011, 2013) distinguished five types of change: addition and deletion of special laws; change in a dimension's importance; change of geometry; change in separability; addition and deletion of dimensions. The first two of the five types address *intra*-framework changes; the other three types describe *inter*-framework changes. Recall that our concern is with conceptual frameworks rather than theories, and that two or more distinct theories can share the same framework.

Firstly, in our model the special laws of a theory provide constraints on the distribution of points over a conceptual space. Newton's second axiom, for instance, brings about a general restriction to points on the hyper-surface described by $F=ma$ in the conceptual space consisting of the domains *time*, *mass*, (physical) *space*, and *force*

[6] We have previously argued that Kuhn's account unduly assumes the primacy of the symbolic level of representation (Zenker and Gärdenfors 2015).

(Gärdenfors 2000). The law of gravitation, which is a special-law, here achieves a more specific restriction to regions where $F = GMm/r^2$ holds. Any further restriction to singular points in this space comes about exclusively in virtue of assuming particular antecedent and boundary conditions. We view the addition, or deletion, of axioms and special laws as the mildest type of change because, unlike the other four types, it leaves the conceptual framework intact. That is, such changes are inter-theoretical while being intra-framework. For instance, a somewhat "deranged" Classical Mechanics with $F = GMm/r^3$ as its law of gravitation still makes for a theory of Classical Mechanics, it just is not Newtonian Mechanics. Often this kind of change is nothing more than the specification/definition of further derivative dimensions on the basic dimensions and this, obviously, one can do without altering the basic dimensionality of the theory.

Secondly, the dimensions, or domains, that are presupposed by a scientific framework may change in importance; in particular, dimensions can go from basic to derivative and vice versa. For instance, until the work of Lagrange and Hamilton, energy remained of little significance to Classical Mechanics (Gärdenfors and Zenker 2013). In contrast, energy became ever more important in the development of 19th century fluid dynamics as more fine-grained versions thereof were developed (Petersen and Zenker 2014). Although the importance of dimensions may differ between two distinct versions of the same theory, the predictive content of that theory need not change. For instance, a version of Newtonian mechanics featuring the law of preservation of kinetic energy—which Huygens, calling it '*vis viva*', established in around 1663 (Rothman 1972), well before Newton's own efforts—is empirically equivalent to the Hamiltonian formulation of CM, even though the role of energy in the respective variants of that theory had changed drastically. For this reason, such changes must be viewed as intra-framework.

Thirdly, each domain of a conceptual space is endowed with a geometry that determines a measure of distance in the domain. This geometry may be replaced by another without changing the rest of the conceptual framework.[7] For instance, a circular dimension with 24 equally-sized intervals called 'hours' models daily time. Before mechanical clocks had been invented, the two points on this dimension that separate twelve night-time hours from twelve daytime ones were locally coordinated to sunrise and sunset. As, again locally, these points shift throughout the year, their distance changes. So we had a variable temporal metric that has since given way to constant clock intervals. The same, as it were, occurred in reverse with physical space when the constant and invariant geometry of Euclid, that was assumed by Newton, Kant and others, gave way to a relativistic geometry that varies with mass distribution. As such examples indicate, a change in the geometry of the conceptual space of the basic dimensions of a theory is plausibly an inter-framework change.

Fourthly, even though the dimensions of two frameworks may be the same, they may differ with respect to the division of these dimensions into domains. For instance, 3D space and 1D time are separable domains in Classical Mechanics, but they form an integral 4D space-time domain in relativity theory. In this article we initially include

[7] This version of the criterion is a strengthening of the criterion proposed in Gärdenfors and Zenker (2011, 2013) where only changes of scales; e.g., the change from Celsius to Kelvin, were considered. This strengthening corresponds to a more severe change of a theory, so we now put it after changes in the importance of dimensions.

changes in interference between dimensions in the fourth category of changes due to the traditionally assumed association between dimension interference and their non-separability. As we will see more clearly below, changes in the interference of dimensions go along with changes in the structure of the conceptual space that are in a relevant sense "deep" or far-reaching. This will ultimately lead us to suggest that change in the interference between dimensions should form its own category of conceptual space change, one that is more severe than any changes in the separability of dimensions.

Fifthly, the most radical type of change is the addition of a basic dimension/domain to a scientific framework, or its deletion from it, for instance, the addition of the dimension *charge* to yield electrodynamics, or the deletion of the dimension representing the quantity referred to as *caloric* in early versions of thermo-dynamics, which proved to be coextensive with energy. A further example is the addition of rest mass energy in SRT as a component of the relativistic energy, to which we will return below.

Categorizing changes of theories into these five (possibly six) types clearly provides a finer grain than Kuhn's distinction between normal and revolutionary change, and so provides a richer toolbox for studying concept evolution in science. Against this background, we now turn to a presentation of phase-space formulations for classical mechanics, quantum mechanics, and relativity theory.

5 Phase-space as conceptual space

A physical phase-space is a space of physical states: the state of a physical system can be represented as a single point in the appropriate phase-space, the points of the space together covering all possible system states.[8] The similarity with conceptual spaces should be apparent. Indeed, our thesis is that phase-spaces *are* conceptual spaces.

As a matter of contingent historical fact, theories are not born in their phase-space formulations; rather, there is some time lag between the original presentation of a theory and the casting of that theory in a phase-space formulation. That much effort is spent in developing these formulations irrespective of their practical utility[9] and despite the availability of alternatives stands as a strong testament to their important role in the foundations of physics. In particular, it is a widely held view that important insights into how physical theories are related can be gained from comparing their phase-space formulations. This is entirely in line with our thesis that phase spaces are conceptual spaces. It is perhaps for this reasons that Gleick (1987, 134) calls phase space "one of the most powerful inventions of modern science".

Although the three major physical theories dealt with here differ significantly from one another when viewed in their standard guises, a great deal of continuity between them readily appears when viewing them in their phase-space formulations. We start

[8] Phase-spaces have been predominantly used to represent the degrees of freedom of complex systems and to model chaotic behavior. This use is related to their use in formulating physical theories in the tradition of Boltzmann (that we are interested in here) but often obscures that original purpose. For a historical introduction see Nolte (2010). For a brief introduction to classical and quantum phase-space see Tao (2007).

[9] The phase space formulation of quantum mechanics is particularly difficult to work in relative to the other formulations of that theory.

with the phase-space of Classical Mechanics (5.1), move on to Quantum Mechanics (5.2), then compare the algebras of observables in Classical Mechanics and Quantum Mechanics (5.3), to finally reach Relativity Theory (5.4).

5.1 Classical Mechanics (CM)

In CM the state of a system at time *t* is given by the momenta $\mathbb{p}$ and positions $\mathbb{q}$ of its constituent parts. Given these positions and momenta at *t*, the positions and momenta of these constituent parts at all times other than *t* are determined by Hamilton's equations of motion:

$$\frac{d}{dt}\,\mathbb{q}(t) = \frac{\partial}{\partial \mathbb{p}} H(\mathbb{q}(t), \mathbb{p}(t)),\ \frac{d}{dt}\,\mathbb{p}(t) = -\frac{\partial}{\partial \mathbb{q}} H\big(\mathbb{q}(t), \mathbb{p}(t)\big) \tag{1}$$

These equations describe the motion of a system through phase-space as a function of time: that is, $(\mathbb{q}(t), \mathbb{p}(t))$ is the trajectory of a system in phase-space. The function H takes a phase-space trajectory as its argument and has a real number as its value. This function is commonly referred to as the *Hamiltonian* of the system and its value at *t* is interpreted as the energy of the system at *t*.

Geometrically, the configuration/physical space of a system can be defined as a *manifold M*, such that for any position $\mathbb{q} \in M$, the momentum of a system at that position is in the cotangent space $T^*_{\mathbb{q}}M$ at $\mathbb{q}$. Thus phase-space can be defined as the *cotangent bundle* $T^*M := \{(\mathbb{q}, \mathbb{p}) : \mathbb{q} \in M, \mathbb{p} \in T^*_{\mathbb{q}}M\}$ with attendant *symplectic form* $\omega := d\mathbb{p} \wedge d\mathbb{q}$. In terms of the symplectic form—with $x(t) := \big(\mathbb{q}(t), \mathbb{p}(t)\big)$ as the phase-space trajectory of the system, and ∇_ω as the symplectic gradient—Hamilton's equations of motion may be combined into a single equation:

$$x(t) = \nabla_\omega H(x(t)) \tag{2}$$

For a single body system, phase-space will have six dimensions; three of these will be spatial dimensions and three will be momenta dimensions. The geometry of this space is Euclidean, and so the Lie algebra of the phase-space transformation operations (spatial translations, temporal translations, boosts, and rotations) is that of the Galilean group. Add the dimension of time, and we have the seven basic dimensions of classical phase-space. A single bodied system is fully described by a trajectory in that space, with all other qualities of that system being defined in terms of these basic dimensions; for example, force at time *t* is the time derivative of system momentum at *t*. In a *n*-body system the phase-space dimensionality is 6*n*, but again these are evenly split into spatial and momenta dimensions. That is, classical phase-space (indeed, any phase-space) can be used to represent "the complicated motions of multiple particles in a single three dimensional space as a single point moving in a multidimensional space" (Nolte 2010); see Fig. 2.[10]

[10] There is a lesson for psychology, and other disciplines using conceptual spaces, here. In the literature on conceptual spaces one does not see the multiplying of dimensions with the multiplying of objects being located in the conceptual space, but if our thesis—that phase-spaces are just an example of a conceptual space—is correct, then this aspect of conceptual spaces should be recognized more widely.

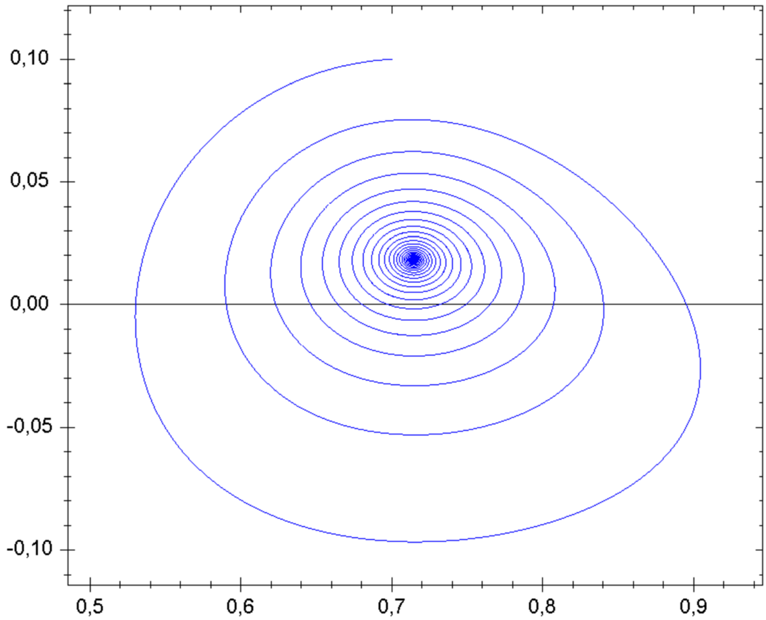

Fig. 2 A phase-space plot of a system with focal instability. Here, the horizontal axis gives the position and the vertical axis the momentum. As the system evolves, its state follows the trajectory line

There is an alternative description of the Hamiltonian equations in terms of *observables*, which are functions $A: T^*M \to \mathbb{R}$ such that:

$$\frac{d}{dt}A(x(t)) = \{A, H\}(x(t)) \tag{3}$$

Here, $\{A, H\} := \omega(\nabla_\omega A, \nabla_\omega H)$ is the Poisson bracket. The term 'observables' is used in a wide sense, to denote qualities of a physical system that can be ascertained by means of experiment rather than in the restricted sense of qualities that are directly observable by unaided human agents. This usage is in accord with modern physics. The Poisson bracket in CM codifies a commutative Lie algebra—a Lie algebra where the Lie Bracket is uniformly vanishing—on the space of observable functions. Thus for all classical observable functions $A: T^*M \to \mathbb{R}$ and $B: T^*M \to \mathbb{R}$, $[A, B] = 0$.

Finally, where there is uncertainty in the position/momentum of a system, the state of the system can be represented by a density distribution $\rho: T^*M \to \mathbb{R}^+$ over Euclidean phase-space. This distribution may be time-dependent, with the attendant equation of motion for such a system being given by the *Liouville equation*:

$$\frac{\partial}{\partial t}\rho(\mathbb{q}, \mathbb{p}, t) = -\frac{p}{m}\nabla_{\mathbb{q}}\rho(\mathbb{q}, \mathbb{p}, t) + \nabla_{\mathbb{q}}V\,\nabla_{\mathbb{p}}\rho(\mathbb{q}, \mathbb{p}, t) \tag{4}$$

In a 2D phase-space, the Liouville equation can be written succinctly in terms of the phase-space density and Hamiltonian functions using the *Poisson bracket*:

$$\rho = -\{\rho, H\} \tag{5}$$

The expectation of a classical observable $A(\mathbb{q}, \mathbb{p})$ at t is given by the so-called "phase-space average":

$$A = \iint A\rho \, d^3p d^3q \tag{6}$$

5.2 Quantum Mechanics (QM)

Just as CM has a phase-space formulation, so does QM. One might think that because no system can have a precise momentum *and* position due to the Heisenberg uncertainty relation, this would preclude a phase-space formulation of QM. Recall, however, that where there was uncertainty in the position/momentum of a system in CM, its state could nonetheless be represented by a density distribution $\rho : T^*M \to \mathbb{R}^+$ over Euclidean phase-space. A similar representation is available for quantum states in a phase-space T^*M with a Hamiltonian $H: T^*M \to \mathbb{R}$. Here, the (pure) state of the system is represented by a complex valued wave function $\psi := M \to \mathbb{C}$, which can be thought of as the quantum analogue of a classical density over configuration/physical space.

The analogy with CM can now be extended by introducing the *Wigner distribution* $W_\psi : T^*M \to \mathbb{R}$, which is a quasi-density distribution, determined by the wave function ψ, over Euclidean phase-space, and otherwise similar to the classical ρ except that it can take negative values (Fig. 3).

That the Wigner distribution of a system can take negative values might appear problematic, but it can be shown that the regions in which this is so are of the order of $\hbar$ in volume. As the uncertainty relations preclude localization of a system to regions of phase-space smaller than $\hbar$, the negative probability for location in such regions can be interpreted as a sign of physical impossibility. That is, negative probabilities are, for all practical purposes, treated as zero-probabilities.

The QM equation of motion for a system represented by a Wigner distribution over Euclidean phase-space is the *quantum Liouville equation* (compare with Eq. 4), which can be written as:

$$\frac{\partial}{\partial t} W_\psi(\mathbb{q}, \mathbb{p}, t) = -\frac{p}{m} \nabla_{\mathbb{q}} W_\psi(\mathbb{q}, \mathbb{p}, t) + \nabla_{\mathbb{q}} V \, \nabla_{\mathbb{p}} W_\psi(\mathbb{q}, \mathbb{p}, t) + \hbar K \tag{7}$$

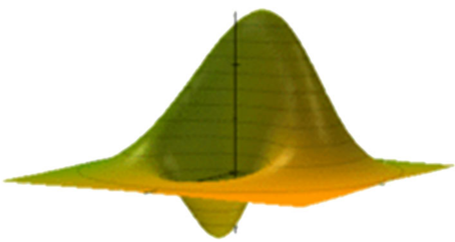

Fig. 3 The Wigner function on a phase plane for the simple harmonic oscillator in a combined ground and 1st excited state. Note the portion of the distribution that is negative

In a 2D phase-space Eq. 7 can again be re-written in terms of the *Moyal Bracket*, which is a deformation (extension) of the *Poisson Bracket* with deformation parameter $\hbar$ (compare with Eq. 5).

$$W_\psi = -\{\{W_\psi, H\}\} = -\{W_\psi, H\} + \hbar K \tag{8}$$

Here, K is a sum of higher order derivatives, greater than third order, of the potential. For systems—such as harmonic oscillators—where higher order derivatives of the potential are zero, the quantum equation of motion for W_ψ is precisely the same as for the classical ρ. The quantum correction terms also vanish in the classical limit that $\hbar \to 0$.

Classical observable functions $A : T^*M \to \mathbb{R}$ are *quantized* to become self-adjoint Hermitian operators $\hat{A} : M \to M$ in the standard formulation of QM. However, in the phase-space formulation of QM, those operators are Wigner-transformed back into functions on phase-space such that the expectation of the observable corresponding to the operator $\hat{A}$, in the phase-space formulation where $a(\mathbb{q}, \mathbb{p})$ is the Wigner transform of $\hat{A}$, is the phase-space average (compare with Eq. 6):

$$\hat{A} = \iint a W_\psi \, d^3p d^3q \tag{9}$$

The Moyal bracket in QM codifies a non-commutative Lie algebra—a Lie algebra where the Lie bracket is not uniformly vanishing—on the space of observable functions. Thus there are pairs of quantum observable functions $A : T^*M \to \mathbb{R}$ and $B : T^*M \to \mathbb{R}$ such that, $[A, B]_M \neq 0$.

5.3 Change in algebras of observables

As the preceding two subsections should have made clear, in the phase-space formulations of both CM and QM the physics of a single body system is fully described by locating, as it were, that system in a six dimensional Euclidean phase-space at each moment in time. All that is required to give a complete description of a physical system in both theories, then, are three spatial dimensions, three momenta dimensions and one temporal dimension.

Where QM differs fundamentally from CM is in the space of its functions representing observables; in particular, while all observable functions commute in CM, this is not the case in QM. *That is, the change from CM to QM is a change in the space of functions representing observables of the system, and not a change in the phase-space on which those functions are defined*. While there is a change in the functions themselves, and so a change of type 1, this is not what leads to a non-abelian Lie algebra of the observable functions in quantum mechanics. The non-commutativity comes rather from the Wigner transform of the standard QM operators into functions on phase-space. In order for this transform to uphold the commutation relations of standard quantum mechanics, the standard product of the classical Lie bracket has to be replaced with what is called the *twisted product* of the quantum Lie bracket, which is associative but non-commutative. In effect, what this change in the Lie bracket product does is replace the standard product with a pair of operators "sandwiched" between two standard products.

The change in the product of the observable's Lie algebra is difficult to place among the conceptual space changes considered earlier. Its effect is to make the algebra non-abelian and so it renders some dimensions interfering (see Section 6.1). For the present, we will classify this as a change of type 4 (but see below). The change clearly is to the structure of the conceptual space—that is, to the relations between the dimensions—and therefore it is a profound inter-framework change.

5.4 Relativity theory

Both special and general relativity theory (SRT and GRT, respectively) differ from QM in that their phase-space is not the classical Euclidean phase-space. For a single body system, the mathematical entity of relativistic phase-space is rather eight-dimensional (8*n*-dimensional for *n*-body systems), with time no longer a separable dimension. The relativistic configuration space of such a system is space-time, which can be defined as a *manifold* M^4 such that, for any four-position $\mathbb{q}^4 = \{q_x, q_y, q_z, ict\} \in M^4$, the four-momentum $\mathbb{p}^4 = \{p_x, p_y, p_z, i\, {}^{E}\!/_{c}\}$ of the system at that four-position is in the *cotangent space* $T^*_{\mathbb{q}^4}M^4$. So, again, phase-space can be defined as the *cotangent bundle* $T^*M^4 := \left\{(\mathbb{q}^4, \mathbb{p}^4) : \mathbb{q}^4 \in M^4, \mathbb{p}^4 \in T^*_{\mathbb{q}^4}M^4\right\}$, only now there is a temporal dimension to the configuration domain and an energy dimension to the momentum domain. Kinetic energy E can be defined in terms of $\left\{p_x, p_y, p_z\right\}$ in CM, but in SRT E is a new basic dimension known as *relativistic energy*. Relativistic energy is the sum of kinetic energy and rest mass energy, the latter being an entirely novel dimension not present in either CM or QM.

The states occupied by the system through time, then, are represented by a *world line* through this phase-space (see Fig. 4). The geometry of the phase-space is no longer Euclidean but Minkowskian in SRT, and pseudo-Riemannian in GRT.

The pursuit of a phase-space formulation of GRT is currently cutting-edge theoretical physics; indeed, it is central to one of the hopes for a grand unified theory (Giovanni et al. 2011). For this reason, our focus in this section is on the phase-space of SRT, which serves as a local approximation of general relativistic phase-space. This space is subject to the Lorentz group of transformations; for example, for velocity boosts in the *x*-direction (with a Cartesian coordinisation of configuration space) the transform is as follows:

$$
\begin{aligned}
dx' &= \frac{dx - vdt}{\sqrt{1-\frac{v^2}{c^2}}} & dp_x' &= \frac{dp_x + \frac{v}{c^2}dE}{\sqrt{1-\frac{v^2}{c^2}}} \\
d\,y' &= dy & d\,p_y' &= d\,p_y \\
d\,z' &= d\,z & d\,p_z' &= d\,p\,z \\
dt' &= \frac{dt - \frac{v}{c^2}dx}{\sqrt{1-\frac{v^2}{c^2}}} & dE' &= \frac{dE + vdp_x}{\sqrt{1-\frac{v^2}{c^2}}}.
\end{aligned}
$$

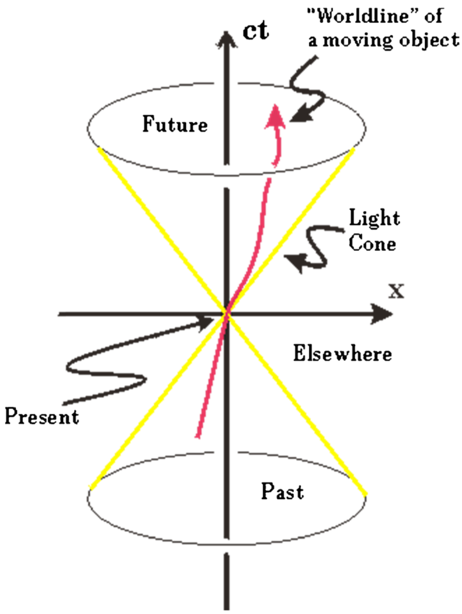

Fig. 4 A relativistic system, represented by a *world line* through Minkowskian phase-space, where momentum can be thought of as the dimension perpendicular to the plane of the page. Points in relativistic phase-space are classified as *events*

Among the invariants under such transforms are the lengths of vectors in both the configuration and momentum spaces as well as the relativistic action $ds = dxdp_x + dydp_y + dzdp_z - dEdt$. From the relativistic action, one can derive the relativistic Lagrangian, and so obtain an equation of motion for special relativistic systems that is invariant under the Lorentz transformations. Furthermore, one can establish that the path taken by any system through relativistic phase-space is that which minimizes the relativistic action ds for that system.

To bring out the similarity of SRT with QM, one can consider cases where there is uncertainty as to the phase-space location of an event. Where there is such *uncertainty* this can be represented by a density distribution $\rho^* : T^*M^4 \to \mathbb{R}^+$ over Minkowskian phase-space such that the probability that the event is in phase-space volume Ω is:

$$P\left((\mathbb{q}^4, \mathbb{p}^4) \in \Omega\right) = \iint_\Omega \rho^* \, d^4p d^4q \qquad (10)$$

Again, observables G are represented by functions $g(\mathbb{q}^4, \mathbb{p}^4)$ on the phase-space such that the expectation of an observable is given by the "phase-space average":

$$g = \iint g\rho^* \, d^4pd^4q \qquad (11)$$

This holds in theory; however, it is generally easier in practice to work solely in the configuration or momentum subspaces of SRT phase-space. Physicists, therefore, tend to implicitly integrate over all of the subspace they are not working in. They can do this because these subspaces are non-interfering domains in CM, SRT and GRT. As we shall see, however, they are interfering in QM.

6 Gauging continuity

We hold that the phase-spaces of CM, QM and SRT as conceptual spaces represent the basic conceptual frameworks of these theories. Using the taxonomy of conceptual space changes (Section 4), we can now classify, and rank, the changes required in order to generate one framework from another.

6.1 From classical mechanics to quantum mechanics

Notice that the transition from CM to QM involves no addition or deletion of basic dimensions (no type 5 change), that there is no change in the relative importance of the dimensions (no type 2 change), and that the phase-space geometry remains Euclidean (no type 3 change). In their phase-space formulations both theories are, at their core, geometrically 6D Euclidean cotangent bundles linearly ordered in time, with the same set of derived dimensions.

There are two related changes of the underlying structure. The first is a change in the observable functions, which is a change of the dependence of derived dimensions on the fundamental dimensions and so a change of type 1. The second change is that in CM, momentum and position, energy and time, etc. are respectively non-interfering domains, while in QM they are interfering. For instance, in CM, one can measure position and momentum simultaneously to an arbitrary degree of accuracy, while in QM the uncertainty relation—$\Delta x \Delta px \geq \hbar 2$—tells us that this is impossible. Indeed, Bohr's *complementarity thesis* infers the lack of non-interfering methods of measurement as the best explanation of the impossibility of precise co-determination in measurement. Murdoch's (1987) summary of the thesis is worth citing in full:

> All observation involves an interaction between two objects, the object of observation and the instrument of observation. The interaction is a finite exchange of energy and momentum between object and instrument. It is, however, a presupposition of classical physics that 'the phenomena may be observed without disturbing them appreciably'. The measurement interaction, according to classical physics, may in principle be made arbitrarily small, and it is in general either negligible, or determinable and controllable. … In quantum mechanics, however, the measurement interaction is in general not negligible, since the energy exchanged is large relative to the total energy of the object, and cannot be made arbitrarily small, owing to the quantum of action. But most important of all, the interaction involved in position measurements is indeterminable … in virtue of the fact that the instrument must be rigidly connected to the

> macroscopic apparatus defining the spatial reference frame [and so] the energy and momentum which the instrument gains or loses in the measurement process disappear irretrievably within the surrounding apparatus. Thus if on the one hand we observe an object, i.e. determine its spatio-temporal location, we interfere with its dynamic state; if on the other hand we determine its dynamic state [by performing a measurement using a device that is loosely connected to a macroscopic device defining the spatial reference frame], we are debarred from observing its position. (Murdoch 1987, 85f).

Thus, from the logical incompatibility of measurements of kinematic and dynamic observables—the former requiring rigidly attached instruments, the latter requiring loosely attached instruments—and the impossibility of arbitrarily small energy/momenta exchanges implied by the quantum of interaction-postulate of QM, Bohr explains the impossibility of precise co-determination of kinematic and dynamic observables through experiment, and hence the uncertainty relations. That is, Bohr explains the uncertainty relations as a material consequence of the absence of non-interfering measurement procedures for kinematic and dynamic observables given the quantum of action, which is to say that Bohr explains the uncertainty principles as a consequence of domain interference.

The formal expression of this change in the transition from CM to QM is the aforementioned change in the Lie algebra of observables; in particular, while all observable functions commute in CM, this not the case in QM. This non-commutation of observable functions is standardly interpreted as the mathematical representation of the fact that the order of measurement matters: i.e., if one measures the momentum of a system in the x-direction *after* having measured its position on the x-dimension one is liable, though not certain, to obtain a different result than if one had measured that momentum *before* measuring that position. Moreover, if one measures x-momentum, x-position and then x-momentum again, then the first momentum measurement is liable to differ from the second. But if one just measures x-momentum twice in a row the second result will certainly agree with the first.

The connection between the change in observable's algebra and the uncertainty relations, and so the interference of the corresponding dimensions, is easily established in the matrix formulation of QM. In this "Hilbert space"-formulation, observables are represented by self-adjoint Hermitian operators acting on vectors $|\psi\rangle$ in a Hilbert space representing the states of systems. Thus, let $\hat{A}$ and $\hat{B}$ be self-adjoint Hermitian operators representing observables A and B, and define ΔA and ΔB as the root mean square deviations of A and B so that:

$$\begin{aligned}(\Delta A)^2 &= \left\langle \psi \middle| \left(\hat{A}-\left\langle \hat{A}\right\rangle\right)^2 \middle| \psi \right\rangle, \text{ where } \left\langle \hat{A}\right\rangle = \left\langle \psi \middle| \hat{A} \middle| \psi \right\rangle, \\ (\Delta B)^2 &= \left\langle \psi \middle| \left(\hat{B}-\left\langle \hat{B}\right\rangle\right)^2 \middle| \psi \right\rangle, \text{where} \left\langle \hat{B}\right\rangle = \left\langle \psi \middle| \hat{B} \middle| \psi \right\rangle\end{aligned} \tag{14}$$

Then one can prove the following inequality:

$$(\Delta A)^2(\Delta B)^2 \geq \frac{1}{4}\left\langle i\left[\hat{A},\hat{B}\right]\right\rangle^2 = \left|\psi\right|\left[\hat{A},\,\hat{B}\right]\left|\psi\right|^2 \tag{15}$$

Substituting $\hat{x} = x$ for $\hat{A}$ and $\hat{p} = -i\hbar\frac{d}{dx}$ for $\hat{B}$, one obtains the uncertainty relation:

$$(\Delta x)^2(\Delta p_x)^2 \geq \left(\frac{\hbar}{2}\right)^2 \tag{16}$$

Hence, when going from CM to QM, there is equivalence between the *novel* impossibility of precise co-determination of observables in measurement and changes in the Lie algebra of observables. As the following quote shows, Bohr (1963, 61) was well aware of this.

> Indeed, it became evident that the formal representation of physical quantities by non-commuting operators directly reflects the relation of mutual exclusion between the operations by which the respective physical quantities are defined and measured.

In QM, then, we have a perfect demonstration of how structural changes in a theory's framework are intimately connected to measurement of the dimensions of that theory. Indeed, we suggest that non-commutativity of observable functions is the theoretical signature of interference between the corresponding dimensions of the conceptual space spanned by those dimensions. Recalling Bohr's complementarity thesis, this interference has its physical basis, in QM, in the quantized interaction between instruments and systems, and the incompatibility of dynamic and kinematic measurement.

In summary, there is a change of type 1, a change in the dependence of dimensions on phase-space trajectory, which in this case also accompanies a change of type 4, in the interference between certain dimensions. The structural change "behind" this type 4 change is the switch from the standard product of the observable's algebra of CM to the twisted product of the observable's algebra of QM, and this is certainly an inter-framework change. But there are no changes of types 2, 3 and 5 in going from CM to QM. That is, no change occurred to the underlying geometry, or dimensionality, of the basic dimensions, and the dimensions retain their original level of importance when moving from CM to QM.

6.2 From classical mechanics to special relativity theory

Having described the two relevant changes to the conceptual space in the transition from CM to QM, we now turn to the transition from CM to SRT. In contrast to the CM to QM-transition, there is a change in both the geometry and the dimensionality of phase-space when going from CM to SRT. As we saw earlier, the time dimension is integrated into the configuration manifold and thereby ceases to be a separable dimension. Furthermore, the relativistic energy-dimension—which is a compound of the kinetic energy dimension, itself a derived dimension in CM, and the rest mass energy, an entirely new dimension in SRT—is integrated into the momentum cotangent space. Thereby, relativistic phase-space gains two extra dimensions over the classical and quantum phase-spaces. Moreover, the geometry goes from being Euclidean to Minkowskian; which is to say that the Lie algebra of the phase-space transformation operations (spatial translations, temporal translations, boosts, and rotations) goes from that of the Galilean Group to that of the Poincaré Group.

In terms of the changes in conceptual space identified earlier, the change in phase-space geometry is a paradigmatic example of the third type of change, while energy becoming one of the basic dimensions in the momentum space is clearly an example of the third, fourth and fifth types of change. It is an example of the third kind of change since kinetic energy goes from being a derived dimension to being a basic dimension, it is change of type 4 since the relativistic energy is integrated into the four-momentum, and it is a change of type 5 since the rest mass component of the relativistic energy is a new dimension unheard of in CM. The integration of time into the position domain is also an example of the fourth type of change. Finally, there are likewise changes of the first type: for instance, the new relation between momentum and kinetic energy in SRT. But these changes, of course, are parasitic on more far-reaching inter-framework changes having occurred.

In contrast to QM, however, none of these changes result in a non-commutative Lie algebra for the observable functions; the product of the Lie algebra of the observable functions remains the standard product here. Indeed, the presence of interference between kinematic and dynamic domains in QM is unique—lacking any precedence in the history of physics. There are nonetheless changes in the separability of domains in the transition from CM to SRT, with time becoming integral to configuration space, and relativistic energy becoming integral to momentum space. Thus, there are changes of all types in going from CM to SRT.

6.3 Degree of conceptual continuity between classical mechanics and its successors

The five kinds of changes that conceptual spaces can undergo (see Section 4) were presented in ascending order of severity: the more severe the change, the less continuity between the predecessor and successor conceptual spaces. This immediately allows us to state that there is greater degree of conceptual continuity between the conceptual spaces of CM and QM, than there is between the conceptual spaces of CM and SRT.

The CM-SRT transition, as we saw, involved the introduction of a new rest mass energy dimension. This is a change of type 5, the most severe type of change, and no such change occurs in the CM-QM transition. Moreover, we also saw that a greater variety of change operations is required to go from CM to SRT, than is required to go from CM to QM: only two types of change are required in the latter case, whereas all five types of change are required in the former.

That being said, a great deal of continuity remains between the conceptual spaces of CM and SRT/QM. In all three theories roughly the same dimensions are defined on a state-space composed of a configuration-space/momentum-space cotangent bundle. Just what those functional dependencies are, how they are structured, and what the geometry and dimensionality of that phase-space is, differs from framework to framework in clearly defined ways—allowing us to track how each conceptual space can be altered to yield either of the others.

This remains the case despite any potential disagreement regarding the severity ordering of the types of changes identified. For instance, one may argue on the basis of the above analysis that an alternative classification of the types of change in conceptual spaces is superior to that proposed in Gärdenfors and Zenker (2011, 2013), given the strengthened condition of change in geometry and the new distinction between separability and non-interference. The attendant ranking in order of severity, then, might more plausibly be:

1. Change in the special laws (intra-framework)
2. Change in dimension importance (intra-framework)
3. Change in the clustering of dimensions into domains (inter-framework)
4. Change in the interference of dimensions (inter-framework)
5. Change in basic dimension geometry (inter-framework)
6. Addition/deletion of basic dimensions (inter-framework)

But even on this severity ordering of the changes, there is greater conceptual continuity between CM and QM than there is between CM and SRT. On this classification the CM to QM transition would involve changes of the 1st and 4th kind, whereas the CM to SRT transition would involve changes of the all kinds save the 4th. Of course, this result goes against common opinion as to the similarities between CM, QM, and SRT. In this instance, however, we would advise revising common opinion rather than conceding to it.

Are there further cases where we can definitively state that a theory has greater conceptual continuity with one theory than it does with another? Indeed, there are many such cases. For instance, notice that on either ranking of the severity of changes, the conceptual space of classical electrodynamics (CED) displays less continuity with that of CM than with that of quantum electrodynamics (QED). This is because the addition of the charge dimension is an example of the most severe type of change, and while QED and CED differ in many ways, they do not differ in their basic dimensionality (ignoring intrinsic spin), whereas CED and CM do. So the relations between CM, CED and QED is another case where the classification of changes in underlying conceptual space provides an answer as to how conceptually similar theories are.

7 Discussion

In Section 1, we had raised two theses: (1) Conceptual space models of conceptual frameworks bear a striking resemblance to the phase-spaces of physical theories and that this resemblance is strong enough to warrant the thesis that phase spaces *are* conceptual spaces. (2) By identifying the types of changes in the underlying conceptual space that occur when one theory supplants another, the degree of conceptual continuity between physical theories and their successors becomes apparent.

With respect to the first thesis, the three examples we have given of physical theories in their phase space formulations have, hopefully, sufficed to convince the reader that phase-space *is* conceptual space. This would explain the preoccupation of physicists with these formulations of our theories as contiguous with the preoccupation of psychologists, computer scientists and others with conceptual space models of conceptual systems in general. Since such models represent crucial aspects of how the world is cognized, the desire to find phase-space formulations of physical theories can be viewed as a desire to represent the physical world in manners attuned to human cognitive faculties. Indeed, should humans be *geometric thinkers*, then a geometric physics would naturally be far more accessible to us than one presented, for example, in an axiomatised formal system.

But the rigor of the physics community in developing their phase-space formulations, and the analysis of their algebraic structure that occurs in the discipline, may also hold lessons for other disciplines that are interested in conceptual space models. In most such disciplines the deeper structure of the conceptual spaces they develop is rarely acknowledged, let alone explored, nor is the multiplying of dimensions as the number of objects in one's field of interest increases acknowledged.[11] But if phase spaces are conceptual spaces, then what holds true of the former *may* hold true of latter. So one should explore what the physics of phase-spaces may have to teach us about conceptual spaces in other disciplines.

With respect to the second thesis, we hope to see corresponding analyses of further examples of theory changes to test the viability of conceptual spaces as a tool for exhibiting conceptual continuity between theories and their successors, as well as debate on which classification and ranking of conceptual space changes might be the most plausible.[12] That being said, we judge the claim 'that SRT is conceptually less similar to CM than is QM' to be relatively immune to disagreements over the severity ranking of changes. It is, at any rate, difficult to see how *any* change could be more severe than the addition or deletion of a basic dimension from a conceptual system, or a change in the geometry of those basic dimensions. And as we saw, while CM and QM share the same basic dimensionality and geometry, CM and SRT do not.

That very claim, of course, might strike some readers as a *reductio ad absurdum* of our position, for QM is often felt to be the most bizarre and unique of all physical theories, an attitude that has been propagated by many notable physicists. Richard Feynman (1965) even held that "no one has ever understood Quantum Mechanics" and that, in this respect, QM differed fundamentally from SRT and GRT. Similar comments, by him as well as others, have led to a kind of mysticism about QM and its interpretation. But this now appears far less warranted. In its standard Hilbert space formulation, no doubt, QM *seems* to be conceptually very different from the other physical theories. With the development of its phase-space formulation in the 1970's, however, it should have been clear that the alien nature of QM had more to do with the vagaries of how it is traditionally formulated, and taught, than with anything intrinsic to the theory and its framework. So when considered carefully, our main conclusion—that the conceptual framework of SRT is more dissimilar from that of CM than QM's is—is rather less shocking than the opposing view.

All the same, it is undeniable that a substantive conceptual difference separates QM from other physical theories, which at the same time calls into question standard tests determining whether two dimensions of a conceptual space are integral or separable. In

[11] Besides our own, for instance, we know of no other work on non-Euclidean conceptual spaces, which is obviously needed if one admits that the conceptual space of relativity theory is non-Euclidean.

[12] One intriguing possibility that we cannot explore further here is to motivate a ranking—hopefully, but not necessarily, our ranking—of the severity of conceptual space changes by entailment relations, in a manner similar to how topological morphisms are ranked.

both CM and SRT there is an implicit assumption that the values of distinct dimensions can, in principle, invariably be determined with arbitrary precision. As is clear from Bohr's quote, in QM one has to give up this assumption of co-measurement precision. This change is not a change to the underlying dimensionality or geometry in the conceptual space of these theories, but is rather a change in the geometry of the space of observable functions. Despite the change's slightly peripheral nature, however, it is almost certainly the root of why the transition from CM to QM is often seen as more radical. In any case, it seems implausible to conclude from the interference between domains in QM that those domains are integral, which is precisely the conclusion one should draw on the basis of the psychology test for integrality.

Another example where physics challenges traditional tests for dimension separability is the metric test. Consider that Minkowskian space-time is a domain comprised of four integral dimensions—and yet, its metric is explicitly non-Euclidean. [13] All of this suggests that empirical results on the metric and interference tests for domain separability used in psychology and cognitive science, as well as the conclusions drawn from them, cannot be taken to unproblematically transfer to the physical sciences.

Particularly in Section 6, moreover, judgments regarding conceptual similarity and continuity were seen to depend on assumptions that remain unaffected by, and so are methodologically prior to, applying conceptual spaces. Such assumptions must therefore be established, or criticized, on independent grounds. The hermeneutic fact that the scholarly discussion has so far not been able to "fix" such assumptions, and particularly the related fact that a principled distinction between a theory's conceptual framework and itself is conspicuous by its absence, can go some way towards explaining the varying intuitions among scholars as to whether CM is conceptually less similar to SRT, than CM is to QM.

Phrased positively, with the above severity ordering of conceptual changes in hand, 'whether a given historical transition constitutes a mild ("conservative") or a more radical ("revolutionary") change' becomes a question admitting of more nuanced answers than traditionally supposed, for we have seen that changes to a conceptual framework are *differentially severe* processes rather than all-or-nothing ruptures.

Importantly, inter-framework comparisons, such as those we have conducted herein, depend upon the historically contingent phase-space formulations of our present physics, on one hand, and on some plausible severity ordering of conceptual changes, on the other. In this context, it is vital to recognize both that our physics might have been supplied with different state-space formulations and that, so far, nothing beyond an appeal to *prima facie* plausibility can support either our own, or someone else's, severity ordering of conceptual space transformations. In particular, a phase space formulation of a theory can hardly be appealed to in accounts of that theory's initial development since, as previously noted, the phase-space formulation of a theory typically considerably lags that theory's initial development. Whatever insight might be gained from such an account could at best make for a (highly questionable) episode of counterfactual history.

[13] Perhaps the metric test can be reformulated to account for such examples in the following manner: Dimensions are treated as separable if the data supports the city-block metric as the best measure of "distance" in the space formed by those dimensions and are otherwise treated as integral.

That having been said, contemporary debates can stand to profit from placing conceptual frameworks *next to* the theories that are embedded in them. For instance, doing so may help address concerns—realist (e.g., Sankey 2009) or otherwise—regarding incommensurability: Conceptual spaces and their transformations provide a scaffold allowing the meaning of terms to be successfully coordinated across frameworks.

As another example, our approach may bring additional clarity to Batterman's (2001, ch 7) discussion of the asymptotic relations (as parameter(s) ε are taken to zero) between a theory T and its historical successor T^*.[14] In particular, his main conclusion—that since CM can be seen as a limiting case of SRT as $v \longrightarrow 0$, while the same is not true relative to QM as $\hbar \longrightarrow 0$ (p. 109), so it follows that SRT reduces CM but QM does not (see Nickles 1973; Batterman 2006)—may seem to run contrary to our own claim, namely that SRT is conceptually less similar to CM than QM is. But the fact that one can view CM as a "degenerate idealization" (Hooker 1992) of SRT, on the one hand, while one cannot do the same for CM with respect to QM, on the other, *is* compatible with our claim that SRT is *conceptually* less similar to CM than QM is. After all, ours is a thesis about the conceptual frameworks of these theories, while Batterman has raised a thesis about the theories themselves. Prima facie, and most interestingly, this suggests that relative theoretical reducibility does not imply relative conceptual proximity, which is a conclusion that is only possible given that one distinguishes scientific theories from the conceptual frameworks in which they are embedded.

8 Conclusion

On the basis of an analysis of three physical theories in terms of their phase-space formulations, we have argued for two theses: (1) Conceptual space models of conceptual frameworks bear a striking resemblance to the phase-spaces of physical theories and that this resemblance is strong enough to warrant the thesis that phase spaces *are* conceptual spaces. (2) By identifying the types of changes in the underlying conceptual space that occur when one theory supplants another, the degree of conceptual continuity between physical theories' and their successors becomes apparent.

With respect to the first thesis, the three examples we have given of physical theories in their phase space formulations may have sufficed to convince the reader that phase-space *is* conceptual space. With respect to the second thesis, we hope to see corresponding analyses of further examples and additional debate of the severity ordering of change operations on conceptual spaces. However, we judge our main conclusion—that the conceptual framework of CM is more similar to that of QM than to that of SRT—to be relatively immune to disagreements over the severity ordering of changes. Indeed, when considered carefully, we judge that conclusion to be rather less shocking than its contrary.

Finally, given the (revised) severity ordering of conceptual changes here proposed, 'whether a given historical transition constitutes a mild ("conservative") or a more radical ("revolutionary") change' becomes a question that admits of more nuanced answers than are typically found in the extant literature.

[14] The general schema is: $T = \lim_{\varepsilon \to 0} T^*$

Acknowledgments Parts of this paper were presented at GAP.9, held September 14–17 September 2015 in Osnabrück, Germany, and at EPSA15, held 23–26 September 2015 in Düsseldorf, Germany. We would like to thank audiences at these events, two anonymous reviewers for this journal, Keizo Matsubara and Lars-Göran Johansson for comments and criticisms that helped to improve earlier versions of this manuscript. The authors acknowledge funding from the Swedish Research Council (G.M., F.Z., P.G.), the European Union's FP7 program as well as the Volkswagen Foundation (F.Z.)

References

Batterman, R. W. (2001). *The devil in the details: Asymptotic reasoning in explanation, reduction and emergence*. Oxford: Oxford University Press.

Batterman, R. W. (2006). Reduction. In D. Borchert (Ed.), *Encyclopedia of philosophy* (2nd ed.). Detroit: Macmillan Reference.

Bechtel, W., & Abrahamsen, A. (2002). *Connectionism and the mind: Parallel processing, dynamics, and evolution in networks* (2nd ed.). Oxford: Wiley.

Bohr, N. (1963). *Essays 1958–1962 on atomic physics and human knowledge*. Woodbridge: Ox Bow Press.

Douven, I., Decock, L., Dietz, R., & Égré, P. (2011). Vagueness: a conceptual spaces approach. *Journal of Philosophical Logic, 42*(1), 137–160.

Feynman, R. (1965). *The character of physical law.* Cambridge: MIT press.

Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought*. Cambridge: The MIT Press.

Gärdenfors, P., & Zenker, F. (2011). Using conceptual spaces to model the dynamics of empirical theories. In E. J. Olsson & S. Enqvist (Eds.), *Philosophy of science meets belief revision theory* (pp. 137–153). Berlin: Springer.

Gärdenfors, P., & Zenker, F. (2013). Theory change as dimensional change: conceptual spaces applied to the dynamics of empirical theories. *Synthese, 190*(6), 1039–1058.

Giovanni, A.-C., Laurent, F., Jerzy, K.-G., & Lee, S. (2011). The principle of relative locality. *General Relativity and Gravitation, 43*(10), 2547–2553. preprint, URL = **http://arxiv.org/pdf/1101.0931v2.pdf**. Accessed 1 Apr 2016.

Gleick, J. (1987). *Chaos: Making a new science*. New York: Viking.

Hooker, C. A. (1992). Idealization, naturalism and rationality: some lessons from minimal rationality. *Synthese, 99*, 181–231.

Kuhn, T. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: Chicago University Press.

Maddox, W. T. (1992). Perceptual and decisional separability. In G. F. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 147–180). Hillsdale, NJ: Erlbaum.

Maxwell, G. (1970). Structural realism and the meaning of theoretical terms. In S. Winokur & M. Radner (Eds.), *Minnesota studies in the philosophy of science: Volume 4* (pp. 181–192). Minneapolis: University of Minnesota Press.

Murdoch, D. (1987). *Niels Bohr's philosophy of physics*. Cambridge: Cambridge University Press.

Nickles, T. (1973). Two concepts of intertheoretic reduction. *The Journal of Philosophy, 70*(7), 181–201.

Nolte, D. D. (2010). The tangled tale of phase space. *Physics Today, 63*(4), 33–38.

Petersen, G., & Zenker, F. (2014). From Euler to Navier–Stokes: a spatial analysis of conceptual changes in 19th-century fluid dynamics. *International Studies in the Philosophy of Science, 28*(3), 1–19.

Pylyshyn, Z. W. (1986). *Computation and cognition: Towards a foundation for cognitive science*. Massachusetts: MIT Press.

Raubal, M. (2004). Formalizing conceptual spaces. In Varzi, A. & Vieu, L. (Eds.), *Formal Ontology in Information Systems, Proceedings of the Third International Conference*, *(FOIS 2004)* (Vol. 114, pp. 153–164). Amsterdam: IOS Press.

Rothman, M. A. (1972). *Discovering the natural laws: The experimental basis of physics*. Garden City: Anchor Books.

Sankey, H. (2009). Scientific realism and the semantic incommensurability thesis. *Studies in History and Philosophy of Science, 40*(2), 196–202.

Schurz, G. (2009). When empirical success implies theoretical reference: a structural correspondence theorem. *British Journal for the Philosophy of Science, 60*, 101–133.

Suppes, P. (2002). *Representation and invariance of scientific structures*. Stanford: CSLI Publications.

Tao, T. (2007). Phase space. http://www.math.ucla.edu/~tao/preprints/phase_space.pdf. Accessed 1 Apr 2016.

Thébault, K. P. Y. (2014). Quantization as a guide to Ontic structure. *British Journal for the Philosophy of Science, 67*(1), 89–114.

Zenker, F. (2014). From features via frames to spaces. Modeling scientific conceptual change without incommensurability or aprioricity. In T. Gamerschlag, D. Gerland, R. Osswald, & W. Petersen (Eds.), *Concept types and frames. Applications in language and philosophy* (pp. 69–89). Dordrecht: Springer.

Zenker, F., & Gärdenfors, P. (2014). Modeling diachronic changes in structuralism and in conceptual spaces. *Erkenntnis, 79*(8), 1547–1561.

Zenker, F., & Gärdenfors, P. (2015). Communication, rationality, and conceptual changes in scientific theories. In F. Zenker & P. Gärdenfors (Eds.), *Applications of conceptual spaces: The case for geometric knowledge representation* (pp. 259–277). Dordrecht: Springer.