Trends in the History of Science

Vincenzo De Risi
Editor

Mathematizing Space

The Objects of Geometry from Antiquity to the Early Modern Age

Birkhäuser

Birkhäuser

Trends in the History of Science

Trends in the History of Science is a series devoted to the publication of volumes arising from workshops and conferences in all areas of current research in the history of science, primarily with a focus on the history of mathematics, physics, and their applications. Its aim is to make current developments available to the community as rapidly as possible without compromising quality, and to archive those developments for reference purposes. Proposals for volumes can be submitted using the online book project submission form at our website www.birkhauser-science.com.

More information about this series at http://www.springer.com/series/11668

Vincenzo De Risi
Editor

Mathematatizing Space

The Objects of Geometry from Antiquity to the Early Modern Age

Birkhäuser

Editor
Vincenzo De Risi
History of Science
Max Planck Institute
Berlin
Germany

ISSN 2297-2951 ISSN 2297-296X (electronic)
Trends in the History of Science
ISBN 978-3-319-12101-7 ISBN 978-3-319-12102-4 (eBook)
DOI 10.1007/978-3-319-12102-4

Library of Congress Control Number: 2014957136

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.birkhauser-science.com)

Foreword

The present book collects the essays presented at the conference on *Space, Geometry and the Imagination from Antiquity to the Early Modern Age*, jointly organized by the Max Planck Institute for the History of Science (Berlin) and the Centro Matematico 'Ennio De Giorgi' of the Scuola Normale Superiore (Pisa), and hosted by the Max Planck Institute in August 27–29, 2012. The conference was part of the activities of the MPI Research Group on *Modern Geometry and Space*, directed by Vincenzo De Risi. The Group's principal aim is to study the interconnections between the history of geometry and the philosophy of space in the pre-Modern and Early Modern Age. In particular, several investigations of the fellows of the Research Group were directed toward understanding the complex epistemological revolution that transformed the classical geometry of figures into the modern geometry of space. The above-mentioned conference represented one of the high points of this research, and we are grateful to Birkhäuser for accepting to publish the proceedings of the meeting.

We acknowledge the generous support of the two institutions, the Max Planck Institute and the Scuola Normale, for organizing the conference. We would particularly like to thank the Director of the Centro 'Ennio De Giorgi' in Pisa, Mariano Giaquinta, whose early involvement in the organization was crucial for the realization of the project.

We are especially grateful to the distinguished scholars who took part in the conference and accepted to discuss their views with us on space and geometry, and for their willingness to contribute to the present volume. Their intellectual generosity made it possible for a real scientific exchange to take place that enlightened the workshop in Berlin and still shines in their written essays. We also thank all other participants in the conference, fellows of the Max Planck Institute, students, and visiting scholars, whose points of view and objections played an important part in enriching the discussion.

We finally thank Chiara Fabbrizi and Fred Sengmueller for their help in editing the present volume.

Contents

Contributors

Graciela De Pierris Department of Philosophy, Stanford University, Stanford, CA, USA

Vincenzo De Risi Max Planck Institute for the History of Science, Berlin, Germany

Franco Farinelli Dipartimento di Filosofia e Comunicazione, Università di Bologna, Bologna, Italy

Michael Friedman Department of Philosophy, Stanford University, Stanford, CA, USA

Daniel Garber Department of Philosophy, Princeton University, Princeton, USA

Jeremy Gray Department of Mathematics and Statistics, The Open University, Buckinghamshire, UK

Gary Hatfield Department of Philosophy, University of Pennsylvania, Philadelphia, PA, USA

Andrew Janiak Department of Philosophy, Duke University, Durham, NC, USA

Douglas Jesseph Department of Philosophy, University of South Florida, Tampa, FL, USA

Alexander Jones Institute for the Study of the Ancient World, New York University, New York, USA

Henry Mendell Department of Philosophy, California State University, Los Angeles, CA, USA

David Rabouin Université Paris 7—CNRS Laboratoire SPHERE UMR 7219, Paris Cedex 13, France

Introduction

Vincenzo De Risi

Today the definition of geometry as the *science of space* is generally accepted by both epistemologists and mathematicians. The history of modern geometry is entirely built around the mathematical notion of space, and different approaches to this science, from Gauss' studies of intrinsic curvature to the Erlangen Program, from the discovery of General Relativity to the most recent developments in topology (take, for instance, Thurston's geometrization conjecture and its proof) rely on a general understanding of mathematical space that remains constant through different perspectives and offers a common ground for regarding all these developments as parts of a single enterprise. Modern geometry is simply inconceivable without the notion of space.

Nonetheless, the definition of geometry as the science of space, however standard, is properly speaking *modern*. Should we go through the thirteen books of Euclid's *Elements*, or in fact the entire corpus of ancient mathematics, we would find almost no occurrence of spatial concepts or terms. Were we to follow the millennial development of Classical geometry in the Middle Ages or the Renaissance, we would still not find any reference to space. The first (and quite rhetorical) mention of *spatium* in a geometrical essay does not predate the last decades of the sixteenth century. To see spatial notions effectively *employed* in geometrical reasoning, we have to wait for another one hundred years. Leibniz' work on geometry (the *analysis situs*) is probably the first attempt in this direction, and in any case it ushered in a general discussion about the object of geometrical investigation. The eighteenth century debated whether geometry had to be regarded as a science of space, and this new idea initially attracted more opponents than supporters; by the end of the century the spatial backers won their battle, and the nineteenth century

V. De Risi (✉)
Max Planck Institute for the History of Science, Boltzmannstraße 22, 14195 Berlin, Germany
e-mail: vderisi@mpiwg-berlin.mpg.de

V. De Risi (ed.), *Mathematizing Space*, Trends in the History of Science,
DOI 10.1007/978-3-319-12102-4_1

declared that space was indeed the object of geometry. The *mathematization of space* was then complete: classical geometry came to an end and modern geometry was born.

In fact, even though the divide between ancient and modern geometry may be arbitrarily demarcated into several historically distinct episodes (such as the birth of algebraic geometry or the discovery of the infinitesimal calculus), the transformation of geometry into a science of space is probably the most important change the discipline underwent during the course of its development. In accepting space as its object of investigation, geometry began to study relational *structures* instead of single figures or magnitudes (like triangles or conic sections). In this sense, the entire structuralist approach of modern mathematics is grounded in this important shift of perspective of eighteenth-century geometry, which (to use Cassirer's words) turned a classical geometry of substances (i.e. figures) into a geometry of functions (structures). As soon as space came to be regarded as a mathematical structure, in fact, it was declined in the plural: non-Euclidean geometries, projective geometry, and then the theory of manifolds, were all disclosed by this new approach to space, whereas a geometry of figures could not possibly produce the idea of different global structures. Moreover, the very *spatiality* of these structures was of primary importance. While the primary aim of classical geometry was the calculation of lengths, areas and volumes of given figures, in a geometry of space the notions of position, incidence or direction may play a central role. Once again, non-metric geometry, projective geometry or topology were inconceivable without such a transition from a science of figures to one of space. The transformation of the object of geometry, moreover, had several consequences outside the strictest mathematical domain, and the entire philosophical development of the last two centuries is scarcely understandable without the notion of space conceived as a mathematical structure.

This new picture of mathematics, which in many respects is still the one we have of this science, found its beginnings in the assessment of geometry as the science of mathematical space, and it is grounded in the long process that transformed Greek geometry into a spatial endeavor. This evolution occurred for the most part outside the realm of mathematics, and involved a gradual evolution of the metaphysical conceptions of space, the epistemology of modern science, and several other advances in various disciplines, from geography to the theory of perspective, from mechanics to cosmology. I will now try to sketch a very brief history of the transformation of geometry from the Greek science of figures to the nineteenth-century science of space, with special emphasis on the philosophical aspects of the development. This is obviously a very limited perspective on the whole matter, and it seriously risks several oversimplifications; yet, I think that it may still be useful to have a guiding thread to lead us through the complex history of the mathematization of space before engaging in further and more detailed investigations. To this effect, I would distinguish four stages in the relations between geometry and space: a geometry without space, a geometry in material extension, a geometry in space, and finally a geometry of space.

(1) *A geometry without space*. The first stage encompasses Greek geometry in the Classical and Hellenistic Ages. Here Geometry is completely devoid of any

spatial content or reference. The objects of geometrical investigations are individual figures (triangles, circles, and so on), which are conceived as reciprocally unrelated and are not embedded in any spatial background. They may be regarded as Platonic ideas or Aristotelian abstractions from material substances (or in several other ways), but in any case these singular figures are the sole objects of geometrical enquiry. Accordingly, their definitions are considered to be the true principles of demonstration and theorems only deal with the properties of these figures.

In the epistemology of Plato we do not find any general characterization of the object of geometry since mathematics is for him defined through its method (hypothetical and diagrammatic) much more than through its subject matter. Still, one can easily understand that he regarded *form* and *shape* (εἶδος and σχῆμα) as the main features of geometrical objects. In the age of Plato or shortly thereafter, however, a first general theory of magnitudes and proportions was developed. This theory, which is often attributed to Eudoxus and was later embedded (as Book Five) in the Euclidean *Elements*, was general enough to state and prove theorems that uniformly apply to any kind of continuous extension (be it a line, a plane figure or a solid). In this way, the manifold objects of geometrical (or stereometrical) investigation were reduced to a common *genus*, and it comes as no surprise that Aristotle was able to define geometry through its subject matter (rather than its method) stating that the object of geometrical investigation is continuous quantity, or magnitude (μέγεθος). This definition immediately acquired paramount importance, and was unfailingly repeated by mathematicians and philosophers for more than two thousand years. Even though Aristotle and his successors clearly conceived magnitudes as shaped extensions, and thus as classical geometrical figures (since no abstract or numerical notion of a continuous magnitude was available), shape, form or position were simply regarded as accidents and properties of geometrical objects essentially conceived as concrete magnitudes. The Aristotelian definition bent geometry in the direction of a science of measure, whose principal aims were to compute length, areas and volumes of given magnitudes. In this respect, it concealed the possibility of a mathematical development of the notions of geometrical form or shape, or the (purely spatial) concept of reciprocal position (θέσις) of geometrical objects. Even though we actually find several uses of the concepts of shape and position in most ancient Greek geometry (for instance: a treatment of similarity in Thales, a wide employment of the concept of position in Pythagoreanism, the notion of form in Plato, and all of these in the Euclidean treatise on *Data*), we are not allowed to suppose, of course, that without Aristotle's influence this science actually had the possibility of developing toward non-metrical (and possibly spatial) outcomes. The advancements of Classical geometry as a theory of measurement are largely independent of any epistemological claim, and seem to be essential to this science, which was born (at least in words) to measure the Earth. Yet, it is certainly true that the Aristotelian definition of geometry as a science of magnitudes, so widely accepted in the following centuries, acted as a powerful constraint to later attempts (in the Renaissance or the Early Modern Age) to transform geometry into a science of space, positions and configurations, that could supersede a metric theory of measure; and even Kant, at the very end of our story, could not but define geometry through the category of quantity—as Aristotle had done.

In the Classical Age, spatial notions remained mostly extraneous to geometry. While Plato had understood geometrical objects as pure ideas, the later Greek philosophers, who regarded them as magnitudes (something that in principle might be located somewhere), usually strongly denied that they are in space and claimed that mathematical objects are nowhere (τὰ μαθηματικὰ οὐ πού).

On the other hand, the concept of space itself is altogether missing from the cultural background of Antiquity. The Platonic notion of χώρα seems to refer to a material extension rather than to a proper spatial container (although the interpretation is controversial). The void of pre-Socratic atomism seems to mean much more a mere nothingness than a positive spatial extension, while the void of Epicurean atomism may be extended, but has a few features in common with matter (it moves away when atoms arrive and it is not properly filled by them); in sum, the vacuum of Greek atomism is a spaceless void rather than a prefiguration of space. Both Stoics and Epicureans, apparently, also had a notion of space as something extended that may be filled or not, but they did not even have a name for such an entity which they labeled "the intangible nature" or "what can be occupied by a body"; no properties of such a thing were ever spelled out, nor does it seem they were in any way connected to quantity or mathematics. We do have, of course, several important philosophical treatments of the concept of place (τόπος), which is however an ecological notion rather than a geometrical one and is related to orientation in the environment, geographical position, or cosmic localization (to be in a place is to be in the market, on a boat, in Athens, in a vessel, in the outermost heaven). In all these cases, the place of a body is just another body which is regarded as having the role of a place. Accordingly, there is no space as an independent being that may be quantified, or measured, or otherwise mathematized. Even geography, which developed several mathematical techniques to represent the world, was mostly concerned with the measurement of an object, the Earth or perhaps the Ecumene, rather than its place. It cannot be denied, however, that the plurality of the representational techniques employed in geography and astronomy already in Antiquity (produced by the impossibility of isometrically projecting a sphere on a map) trespassed on the simple conception of measuring given magnitudes, and paved the way for understanding how different geometrical structures may be employed to capture the world. The lack of a veritable concept of space, however, still prevented the possibility of regarding these different geometrical representations as realizations of space structures, or of conceiving them as anything more than an application of the geometry of magnitudes to a specific cosmic object.

(2) *A geometry in material extension*. This very general metaphysical picture, however, began to change already in the Neoplatonic philosophy of Late Antiquity. Aristotelian ontology regarded quantity (and thus magnitude) as an accident of a given individual substance; and this was in perfect agreement with the idea of a geometry of individual figures. Neoplatonic metaphysics, on the other hand, admitted different ontological stratifications, and might conceive (in a few authors)

that quantity is not an accident of the substance, but an accident of matter itself. Matter was thus regarded as quantified and extended, or at least endowed with magnitude as its primary property. This notion of an extended matter, that came from Plato's conception of the Receptacle, had nonetheless acquired an Aristotelian bent, as magnitude was now conceived by Neoplatonists as a positive property that may be geometrically investigated rather than (as in Plato) the unknowable material substrate in which Forms appear. The immediate consequence is that extended matter is in fact the substrate of geometrical objects and their ontological condition, and that this matter exists (in some sense) before being formed in the particular shapes of the individual geometrical figures. This insubstantial extension is thus a kind of background in which geometrical objects have their seat. It is not yet space, as it is a material extension and it is not explicitly connected with any local or positional property; moreover, the material extension is itself devoid of any geometrical property, and geometry remains the science of those definite and individual magnitudes that are made up by this extension: the latter is the condition, rather than the object, of geometry. Yet, this Neoplatonic quantified matter in which geometrical magnitudes are first delineated is the ancestor of the modern concept of space.

This conception of an extended matter as the substrate of geometrical magnitudes is quite general, and it received very different treatments by different metaphysicians and schools. There is no need here to go into the many different perspectives that very different authors held on this topic, as we are only interested in this conception of matter as a prehistory of the notion of space. It needs to be said, however, that while several philosophers intended quantified matter as corporeal matter (the matter of the world) and thus connected the ontology of mathematical objects to that of cosmic extension (possibly through an abstractionist theory of mathematical entities inspired by Aristotle), a few authors from Late Antiquity on regarded the matter of extended mathematical objects as "intelligible matter" (another term taken from Aristotle), or imaginative matter (ὕλη φανταστική). In this way, they advanced an ontology of geometrical objects as ideal entities that are constructed by the mathematician in the imagination. Imagination is conceived as a sort of blackboard on which figures are drawn. Even though, once again, the blackboard itself is not thematized as an object of geometrical investigation, its relevance for the history of the concept of mathematical space is enormous. The main supporter of this view in Late Antiquity was Proclus, who developed an important theory of mathematical "projective" imagination in his philosophical commentary to the First Book of Euclid's *Elements*, but the idea that imagination is the proper faculty of geometry (a conception that was absent from the Greek thought of the Classical Age) had a widespread diffusion in the centuries to come.

In any case, the significance of the ontology of quantified matter (be it corporeal or imaginary) went well beyond the strictest Neoplatonic doctrines, and was generally accepted in the Middle Ages. Especially through the Averroistic tradition, the doctrine of the *dimensiones indeterminatae* of matter (a matter endowed with extension and indeterminate magnitude but without a determinate shape) enjoyed diffusion and acquired relevance in the West (think of Aquinas' *materia signata quantitate*) and formed the basic geometric ontology for several centuries. In the Renaissance, most Late Scholastic authors still conceived matter as essentially extended and regarded it as the ontological condition for the existence of the proper mathematical objects (the shaped magnitudes). The connection of geometry with the imagination remained strong during the Middle Ages, even though the original writings of Proclus were unknown and the Scholastics had to rely on secondary sources. The debate on quantified matter in imagination however reached its highpoint in the Renaissance, when Alessandro Piccolomini (in 1547) interpreted the recently rediscovered Proclus and engendered the so-called *quaestio de certitudine mathematicarum*, a quarrel among metaphysicians and epistemologists that continued for more than a century.

A few philosophers were ready to endorse the Averroistic tradition of quantified corporeal matter as the substrate of geometrical entities to the point of rejecting several Euclidean principles and theorems that cannot be accommodated to bodily magnitudes, since the latter are (for instance) necessarily tridimensional (leaving no space for a pure plane geometry), or since they may not be infinitely divisible; a vast number of indivisibilist mathematicians, in fact, ranging from the Middle Ages (Gerard of Odo or Nicholas d'Autrecourt) to the Early Modern Age (Bruno, Arriaga), grounded their rejection of infinite divisibility and (in some cases) a new finitistic mathematics on their views about the features of bodily extension.

Many others, however, followed Proclus and appealed to the imagination in order to ground the possibility of ideal geometrical constructions. This conception was especially widespread among professional mathematicians, from Peletier to Clavius, Barrow or Borelli, but was shared by several philosophers as well.

Descartes was able to encompass both traditions, with a complex mathematical ontology that conceived geometrical objects as parts of the *res extensa*, but which also stressed their connection with the faculty of imagination. It is remarkable how, in the hands of a great mathematician, the metaphysics of geometrical extension yielded important mathematical outcomes. The algebraization of geometry, in fact, required geometrical figures to be conceived as embedded in a larger (indefinite) extension, that can be captured in a system of coordinates. Such a fundamental mathematical development was not possible in the classical ontology of individual figures, and was in need of an ontology of quantified matter. A few further issues and techniques in geometry, like the theory of *loci* or the use of motion in a proof, acquired new meanings in the wake of this ontological transformation.

Yet, in the innumerable disputes that divided both Scholastics and new philosophers at the end of the Renaissance, and produced such different metaphysical systems, the conception of geometry remained firmly chained to these two pillars: it was the science of magnitudes, as Aristotle had said; and these magnitudes were to be conceived as parts and chunks of an originally extended and quantified matter.

Another development, however, was maturing in the Neoplatonic tradition: the ancient conception of place as an ecological notion was substituted by a concept of space as a three-dimensional extension. This conception is generally ascribed to the Aristotelian commentators Philoponus and Simplicius in Late Antiquity, who first worked it out at some length (with a few differences between them) disavowing their master's opinion. During the Middle Ages the notion was further developed, even though its acceptance always remained controversial, given the authority of Aristotle who had opposed it; but in the Renaissance it gained numerous followers among new philosophers and Aristotelians as well. From this perspective, the three-dimensional extension of place was still conceived as an accident dependent on the existence of the (individual) located substance, rather than an all-encompassing space ontologically independent of the bodies in it. This notwithstanding, the extension of place was quantified and could be measured like any other magnitude (which was not the case of the classical "ecological" place), and this was an important step forward toward the mathematization of space. Nevertheless, three-dimensional place had no special link with quantity or geometry, and it remained one magnitude among others. It was considered measurable, in fact, only by virtue of its capacity to receive matter, since matter only (as we have seen) is the ontological substrate of magnitude.

A further development coming from the Middle Ages was the introduction of the concept of an imaginary space (*spatium imaginarium*). This notion also came from the Averroistic tradition, and hinted at the fact that the imagination cannot help but conceive further extension beyond the boundaries of heaven. The debate on the nature of imaginary space was wide ranging, but we need only observe that it had several advantages over the three-dimensional concept of place: it was a veritable *space*, in the sense that it was not limited to the extension of the individual located body, but rather it encompassed an infinite (or, shall we say, indefinite) extension; and it was a creation of the imagination, which again hinted toward a connection of it with the ontology of mathematical objects as proper products of the *phantasia*. However, this possible connection remained unexplored in the Middle Ages and the Renaissance since the imaginary spaces were conceived as *purely local* constructions, which had nothing to do with quantity in the proper sense; indeed, several authors insisted that being devoid of matter, imaginary spaces have no quantity whatsoever. Even Hobbes, well into the seventeenth century, was still claiming that while geometry may have a substrate in imaginary space, this only comes from the corporeal origin of imaginary space itself—since extension and quantity are nothing but bodily features.

(3) *A geometry in space*. By the end of the Renaissance, however, these manifold traditions were ripe enough to merge into something new. In particular, the Philoponian conception of three-dimensional place and the notion of an infinite imaginary space finally met the ontology of quantified matter and geometrical objects. The hint may have come from Ficino, who translating Plotinus' *Enneads* and dealing with the Neoplatonic treatment of quantified matter (ὄγκος, mass) realized that this was just pure extension (primary matter plus quantity) rather than bodily extension, and thus ventured to translate the term as *spatium*. Ficino did not draw any consequence from his own translation, and continued to profess a standard Neoplatonic philosophy of mathematics and an Aristotelian conception of place. Several years later, however, Francesco Patrizi, a Neoplatonic disciple of Ficino and a vigorous renewer of metaphysics, followed the hint and superimposed

the metaphysics of quantified matter onto that of three-dimensional space. The result was a new conception of space as a three-dimensional infinite extension which is independent of matter and bodies, and is pre-exististent to them as a condition of their own existence. Patrizi's new metaphysics was probably more suggestive than well argued, and was developed in the same years that others (Bruno, in particular) were also elaborating similar views. However, Patrizi was alone in stressing a new conception of mathematics. He claimed that space is originally and essentially quantified and extended, whereas matter and bodies are quantified only because they are embedded in space. He was thus turning the preceding conception upside down by claiming that the possibility of mathematizing the natural world depends on the possibility of applying geometry to space itself. To his essay *On Physical Space* Patrizi added a treatise *On Mathematical Space*, and an entire book *On a New Geometry* (*Della nuova geometria*, 1586), in which he explicitly asserted that "the general subject of mathematics is space" and he tried to rewrite a portion of Euclid's *Elements* in such a way as to reflect the change in the object of study. To my knowledge, this is the first occurrence of such a claim in a book of geometry, and almost the birth of the idea of a geometry of space.

We should add, however, that there is not much more here than an idea and a project. Patrizi's skills as a mathematician were very scarce, and in any case he was not able to foster any real improvement in Classical geometry. Patrizi was still discussing straight lines and circles, not space, and he was not conceiving space as a structure that could be investigated. His claim that space is essentially quantity had allowed him to think of it as the substrate of geometry (a big step forward), but had also imprisoned his "new geometry" in the usual cage of a science of magnitudes (rather than spatial relations), and Patrizi did not really question the Aristotelian tenet that magnitudes are the objects of geometrical investigations. He only changed the ontology, stating that these magnitudes are not just chunks of matter, but rather regions of the all-encompassing space. In sum, his is not a geometry *of* space, but a geometry of figures and magnitudes *in* space.

Patrizi's metaphysics of space enjoyed a wide diffusion, and informed the new atomistic philosophy of the seventeenth century (Gassendi was very familiar with his work), as well as further Neoplatonic developments (such as those of the Cambridge Platonists), and directly or indirectly the new metaphysics spread during the course of the century. Newton himself may be regarded as an heir of these Renaissance metaphysical conceptions of space, and in the *De gravitatione*, in particular, he gives a description of space and the figures existing in it which are reminiscent of several passages of Patrizi's. The idea that the objects of geometry are spatial figures gradually spread, and in a few decades it became commonplace in several circles. The two opposing views that geometrical objects are material or rather spatial entities battled with each other over the next hundred years, and if at the beginning of the seventeenth century nearly everyone had advocated the former position, by the end of the century only a few indomitable Cartesians (or very late Aristotelians) were opposing the latter. The battle was purely metaphysical, though, as all of the participants were still working within the framework of the classical geometry of figures, and (almost) no mathematical consequences were drawn from their different geometric ontologies. Newton, for one, considered himself a classical geometer, and his innovations in the field do not at all concern the subject matter of geometrical investigations (even though his conception of space may have influenced his geometrical endeavors).

Among the several developments that contributed to the discovery of space as a structure, we may at least mention here the development of a mathematical geography and cosmography, as well as the tradition of perspectival studies.

In his celebrated essay on *Perspektive als symbolischer Form*, Panofsky claimed that the modern concept of space was born in the Renaissance treatises on perspective. At least from the time of Giotto or Duccio, to be sure, a new form of spatial organization may be found in pictorial representations. But this new way of representing space was only codified much later, in the late fifteenth century, when a *mathematical* reflection on those painting techniques began to be available in the works of Alberti, and then Piero della Francesca, Albrecht Dürer, and many others. In the sixteenth century, we find an explosion of geometrical essays on the theory of perspective that explicitly discuss the mathematical structure of place and space. The humanist Pomponio Gaurico, in his book *De sculptura* (1504), claimed that the painter should first depict the *locus* of the represented things, and only then the things themselves, since place is prior to located objects. He simply meant that the perspectival technique (the *locus*) should be selected in advance; but his choice of words shows a conception of place that is not an object but rather a mathematical structure which is a condition and an order of objects. Different structures (different spaces) with different mathematical properties may be employed to represent the very same things. While ancient paintings may have depicted single objects in perspective, juxtaposing certain objects with others (each object, I mean, with its own perspective and vanishing point), a modern painter might offer what Gaurico calls a perspective which *ad totum opus pertinet* encompassing the whole scene in a single space. In this respect we moved from a perspective of substances (figures) to a perspective of structures and relational systems.

It is true that this use of the word "*locus*" to mean a perspectival technique may seem quite eccentric, and it was at first confined to the treatises on this discipline without acquiring any more general sense. The essays on perspective, however, became more and more mathematical over the years, and by the end of the sixteenth century they had already transformed from painters' handbooks into studies for mathematicians (think of Guidobaldo del Monte's *Perspectivae libri sex* in 1600). When Desargues and then Pascal or La Hire in the seventeenth century applied perspectival techniques to prove classical geometrical theorems, the development was complete: and the perspectivists' conception of place or space as a structure and an order had became a tool for "real" mathematics. For the first time a few purely spatial notions, like those of *position* or *situation* of points and lines, and the study of spatial incidence and configurations, became an object of geometrical enquiry.

(4) *A geometry of space*. The synthesis of all these different trends toward a geometry of space was first realized in Leibniz' work on *analysis situs*. Leibniz had

the opportunity to read Pascal's lost treatise on conic sections in manuscript (where Pascal had probably employed the notions of *space* and *situation* to ground projective techniques in geometry) and devoted himself to the task of developing a new geometry entirely based on the spatial notion of situation (*situs*).

Leibniz' conception of space, which he famously opposed to Newton's absolute space, was that of a system of situational relations (an *ordre de situations*), and thus a structure. The new *analysis situs*, in developing a geometry of situations, had to study the structure of space itself. This was the first time in history that space was conceived as a structure and geometry as its science. In fact, Leibniz' work on *analysis situs* revolves around the definition and interrelation of a few structural properties of space, like uniformity (our isotropy), homogeneity (our manifold structure), dimensionality, continuity, connectedness, flatness, and so forth, that had been never studied before. The possibility and properties of every figure should be reduced to those of the ambient space, which is the only real object of geometrical investigation. Leibniz' aim with his *analysis situs* is in fact to show that only one all-encompassing spatial structure is possible, thus to ground on absolute foundations the existence of a three-dimensional, uniform, homogeneous, continuous, connected, Euclidean space. And if his goals were unreachable, his endeavors were nevertheless quite interesting as they show how a geometer could forsake the classical geometry of figures and ruler-and-compass constructions and attempt a much more abstract study of the structural properties of a system of relations.

The concept of situation is generally reduced to that of distance (a situation between two points being expressed by their distance), and this fact prevented Leibniz from developing a truly non-metrical geometry (a projective geometry, for instance). This notwithstanding, geometry is for Leibniz first a science of position and place, and only derivatively a science of quantity. He attempted, in fact, to develop a wide-ranging theory of similarity that should have been (partially) free from metrical considerations. Leibniz' relational conception of space also offered a new solution to the problem of the composition of the continuum, which stated that whereas a set of unextended points (however numerous) could never build up a continuous extension, a set of points endowed with situational relations (that is, points considered as terms of relations of distance) may nonetheless be extended, since extension and continuity themselves are nothing but structural properties of such a set. This approach, which was very modern and had never before been attempted, opened the way for a general set-theoretical and structural view of space.

However promising, Leibniz' attempts at a geometry of space remained unpublished, and the eighteenth century only caught a few glimpses of this new science. The Leibnizian program in geometry, however, was well known, and it sparked a lasting debate (especially in Germany) on the objects and methods of geometry. It opposed traditional geometers, who still claimed that figures in space are the object of geometry and appealed to the authority of Euclid, to Leibnizians who advocated the advantages of a geometry of space but could not show any significant consequence of this new science (Wolff, for one, stated that geometry is the science of space but was unable to produce any result that went beyond Euclid).

The same seeds that had engendered Leibniz' *analysis situs* in the seventeenth century, however, continued to bear fruit in the eighteenth century. The developments of a mathematical theory of perspective ended up with Monge's and Poncelet's first attempts at a veritable descriptive and projective geometry, which required an array of spatial and local concepts and tools, and could not help but rely on the idea of a science of space. Closely related to these mathematical developments, the theory of visual perception also pushed a few philosophers on the same shore. The geometrization of visual space had already begun (I mean, with full philosophical awareness) in Descartes' *Dioptrique* and theory of perception, which represented a fundamental step forward toward a mathematization of the world accomplished through a (kind of) spatial structure. But this *géométrisation du regard* continued in the eighteenth century in the works of Berkeley, Hume and Reid (among others), who discussed the manifold properties (and alleged differences) of visual and tactile spaces; and while Berkeley seems to aim at a projective conception of visual space that may not be captured by a geometrical theory, Hume's conceptions of perceptual *minima* (also stemming from Berkeley) may tend toward a finitistic geometry, and Thomas Reid's (quasi) axiomatization of spherical geometry as the geometry of sight professes that space is nothing but a perceptual structure, that a plurality of such structures is possible, and that they can be differently treated from a mathematical point of view.

In the same years, other mathematicians were developing their first essays on non-Euclidean geometry. Saccheri in 1733 was still unable to recognize his "obtuse angle hypothesis" as an instance of spherical geometry (the same that was discussed by Reid), since he was a hyper-classicist mathematician who considered his own studies on the Parallel Postulate to be a Euclidean exercise in a geometry of figures. As early as 1766, however, Lambert realized that the Parallel Postulate does not concern parallel or incident lines, but the deeper structure of space itself. Wallis had shown that the Postulate is in fact equivalent to the possibility of transformations in space through similarity, and the latter is clearly a "second order" property about figures rather than the property of a single figure (or a couple of lines). This allowed Lambert (who had also worked on perspective) to recognize spherical geometry as a description of a non-Euclidean space, thus opening the way to conceive an abstract model for hyperbolic space as well. And while Lambert had written a *Theorie der Parallellinien*, János Bolyai, at the culmination of the non-Euclidean revolution, would later write a *Scientia spatii* (1832), or *Raumlehre*, fully realizing that his researches disclosed the science of a different space rather than a new theory of parallel lines.

Yet another line of enquiry and discussion concerned the role of principles in geometry. Definitions had been considered as the true principles of demonstration from Antiquity to the Eighteenth century, reflecting the idea that the objects of geometry are individually defined figures. From this perspective, axioms were generally regarded as immediate consequences of the definitions. In the eighteenth century, a few mathematicians began to claim that axioms should be the true principles of geometry, and definitions should follow from them (an instance of this attitude is once again in Lambert). This new epistemological claim mirrored the growing awareness that only the complete system of figures and objects, that is the space structure as such, is the true subject of investigation; as well as the awareness that a plurality of such structures (defined by different systems of axioms) are in fact possible.

In sum, just as the seventeenth century had witnessed the battle between supporters of a science of material figures against supporters of spatial figures, so the eighteenth century was divided among those who still professed a geometry of figures (in space) and those who looked for a new geometry *of* space.

The century's richest and most developed epistemology of mathematics, that of Immanuel Kant, still reflected these quarrels. Kant's philosophy of mathematics may be easily seen as the culminating point of several philosophical traditions that we have mentioned so far. His theory of productive imagination and its application to geometry, in particular, clearly draws from the Proclean tradition of a projective *phantasia*, while the idea of the applicability of mathematics to the phenomenal world through the mathematization of space comes from the Neoplatonic tradition of the Early Modern Age. And yet all these elements (and many more) are merged together into a new synthesis that seems to push them to their maximum conceptual strength. In many respects, in the epistemology of mathematics Kant appears to be the last philosopher of the classical age, and to be advancing a complex and consistent theory of geometry as a Euclidean science of figures, rather than a modern science of space. Not only did he strongly attack Leibniz' point of view on the essence of space (now conceived as a pure intuition) and the nature of a mathematical proof (based on synthetic *a priori* judgments rather than analytical statements), but his entire positive theory of mathematics is grounded on a constructivist stance that makes use of the synthesis in imagination for the composition (*Zusammensetzung*) of individual finite figures in space. These figures are the objects of investigation in geometry, as in the classical tradition; more than that, they fall under the category of quantity and are called *magnitudes*, as in the Aristotelian tradition. Space itself is regarded as the condition of geometry much more than its proper object: a formal intuition which is needed as the background of the productive activity of the imagination. This space has a unity and may be regarded as the product of a *sui generis* synthetic act, which is an intellectual but pre-categorial "putting together" (*Zusammenfassung*) of the spatial manifold. This act, which gives a unitary structure to geometrical space, is not a synthesis of imagination ruled by the category of quantity (a "composition" in the proper sense), and thus the structure of space as a whole does not fall under the scrutiny of geometry. Yet, the very concept of this act, which is explicitly (albeit peripherally) thematized in Critical philosophy, distances it from the classical Neoplatonic tradition, and reflects the idea that space as a whole may have a unitary logical structure.

In any event, the first generation of Kantian followers no longer had any doubt that space has a given geometrical structure, and in this respect Kantianism (often blended with Leibnizianism) became an important cultural force (perhaps against the will of Kant himself) to foster the idea of a new science of space. It was through these transformations and philosophical quarrels that the modern party eventually won out against old Euclid, and by the dawn of the nineteenth century it was so common to state that geometry is the *science of space* that almost no one realized that a different opinion had been possible in the past.

The history of the mathematization of space was in fact much more complex (and less linear) than the idealized picture given above. I hope that the essays presented in this book, which illustrate a few episodes of this history in great detail, will help the reader to understand the long development that transformed Greek geometry into Modern mathematics. They also show the many and various epistemological contributions that went into this conceptual development, and point to a few of the several disciplines and figures involved in the "spatial turn" of geometry.

What's Location Got to Do with It? Place, Space, and the Infinite in Classical Greek Mathematics

Henry Mendell

Here is a basic question: how much philosophy of mathematics can one pull out of Greek mathematical texts? Obviously, this depends on what we are taking as philosophy of mathematics. We can describe easily enough what goes on in a typical Greek mathematical treatise, but even here, 'typical' is a loaded word. We define the scope to our liking and to some 7 or 8 extant authors, or maybe to more, but the more we extend our list, the less does the result fit, for example, the tidy and austere world that Netz (1999) portrays in his ground breaking study. Even within this group, we should expect difference and variation. Austerity aside, unless they tell us, we have not a clue how Greek mathematicians thought about their work. In discussions of the philosophy of Greek mathematics, it is very common, following in the path of Proclus, for moderns to find a philosophy of Greek mathematics and then to trace the philosophy to Aristotle or Plato. If we distinguish, however, issues that are intrinsic to a mathematical exposition from external questions such as the ontology of mathematics, we can observe that there is very little evidence of any views about ontology expressed in any Hellenistic mathematicians, while later mathematicians tend to come out of neo-Platonism. Did Autolycus, Euclid, Archimedes, or Apollonius believe that mathematical objects were intermediates, ideal impressions in the imagination, physical objects qua mathematical, or something else, or did they not think about the issue? We have not a clue, except that some of these are unlikely in the Hellenistic Age. If anything, a Hellenistic mathematician was more likely to be a Stoic than an Aristotelian or Platonist, to say nothing of an Epicurean. But why any school at all? Even if we did find a fragment of philosophy, it would be a strange induction to infer that the view expressed was

H. Mendell (✉)
Department of Philosophy, California State University, Los Angeles,
5151 State University Drive, Los Angeles, CA 90032-8530, USA
e-mail: hmendell@calstatela.edu

V. De Risi (ed.), *Mathematizing Space*, Trends in the History of Science,
DOI 10.1007/978-3-319-12102-4_2

anything other than the view of that author. It would be merely less peculiar to infer that the author's view told us how we should understand the ontology expressed in the texts. Many a very good mathematician has made a not so good philosopher of mathematics. When we look at issues internal to the structure of mathematical treatises, we are on better ground, but here too we should resist many temptations, e.g., to connect practices to Aristotle, to over interpret, to generalize about conceptual issues and so forth. With this banal warning, I am going to look at one case where such ontology and mathematics might be thought to cross paths, namely in ancient treatments of place, space, infinite, and related notions: how they play out in ancient philosophical treatments of mathematics and how they show up in mathematical texts. It seems obvious that the notion of the infinite in Greek mathematics can have no other origin than the tradition of Anaximander, Zeno, Melissus, but most significantly Anaxagoras and Democritus. The ordinary treatment of place probably emerges from ordinary geometrical thinking, whether in architecture, surveying, or any other activity that might involve measuring and representation. Speculations on these lie outside my present concern.

I begin with a discussion how certain key words that dominate philosophical discussions of location (taken broadly) are used in mathematical texts. I then look at philosophical discussions that indicate how Aristotle and subsequent Hellenistic philosophers apply locational notions to geometry. If any of these discussions impinged on their treatments in mathematical texts, we would expect there to be a notion of a geometrical space or a conception of spatial relations between geometrical objects in mathematical texts that restricts some mathematical practice or that provides a robust avenue for a mathematical technique. Does either occur? My answer will be mostly negative. Whatever the origins of 'infinite in extent' in Greek discussions of the nature of the world, it becomes a small part of the language of mathematics. Since the infinite is an important spot for trying to see if someone has attitudes about spatial limitations on the treatment of mathematical objects, I shall then turn to Aristotle's claim that the geometers do not use the infinite. In this, I shall look at Euclid, Archimedes, Apollonius, Alexander of Aphrodisias, the third century CE professor of Aristotelian philosophy, and finally at Heron of Alexandria. Here, I think the results will have a certain irony.

A fundamental issue in any philosophy of science or mathematics is that of descriptivism (non-revisionism) versus revisionism, i.e., whether the philosophy aims to describe practice and fundamental conceptions or to change them or even to abandon the science altogether (e.g., because they inhibit one from the good life or because they are irredeemably false). In so far as the philosophy is revisionist, if philosophers can convince practitioners to change their ways, then the revision will become practice and the followers of the philosophers will be espousing a new descriptivism. If they fail, then the conflict will remain unresolved, with the two attitudes remaining incompatible, even if ignored by the practitioners. So I shall examine extents to which Aristotle, Stoics or Epicureans are revisionists and the extent to which their views express an incompatibilism with practice.

We need to distinguish three issues in Greek mathematics. First is the question of how Greek mathematicians treat ordinary spatial relations in their practice.

This is largely a philosophical and interpretative question. Of course, if we can extend lines, draw lines between disjoint and separate figures, move a disjoint line or figure to another, there is a sense in which spatial relations play an important role. This involves much more than drawing straight lines or circles. It involves mechanisms of construction such as sliding rotating lines through a loop (conchoids), coordinated motions (spirals and quadratices), even pointwise constructions; adoptions of simple plane techniques to solid geometry, extending cones or planes, slicing solids, etc.; and maybe even infinite lines. Since the mathematics we are primarily looking at is only metrical in using given magnitudes as measures, these extensions will involve either connecting points, arbitrary but adequate extensions, or multiplying lengths (as in the diagrams for numbers and magnitudes in the proportion theory). To do any of this requires a region where this can be done. Let's call this naïve space.

Much Greek geometry involves merely planar relations (say, where two lines intersect, any third line will intersect at least one of the others if they are extended), and the concept of a plane as extended in two directions is fairly clear. One can ask, however, whether having the concept of objects in the same planar relations requires having the concept of a planar space, that is a plane large enough for all plane objects to be placed, or, more minimally (and plausibly), whether any construction takes place in a plane large enough to place the objects under construction, even if in some cases the plane might be infinite. And one can raise a similar question about solids. This is a cognitive and a philosophical question which I admit I do not know how to address. But it is also a historical question whether certain geometers made any such inference about planar relations to having a conscious theory about planar spaces. Let's say that naïve planar space is that the concept that any new objects introduced will have the same planar relations as other objects already introduced. If we are dealing with solid geometry, then objects in the same plane will have the same planar relations and if objects are being sliced by planes, well those planes are objects. One could go on to define a planar relation more carefully (and a linear relation, etc.), but I prefer here to leave the notion as intuitive. If more is needed for cognitive reasons, then let it be added, but keeping things minimal. The historical question will be whether some Greek mathematicians required more and expressed it in their mathematical discourse.

I take it that naïve space is necessary for doing much of Greek mathematics, although it is not clear to me whether it is also necessary for doing Babylonian or Egyptian mathematics where constructions do not explicitly occur.[1] This feature arises from the constructional nature Greek mathematics. By itself, it tells us no more about Greek mathematicians' conceptions of space than the activity of a pâtissière producing an elaborately layered cake would tell us about her conception of space.

[1] Or at least one could say that the constructed buildings are given, e.g. in BM 85194 (Sippur) Problem (v), one is given a wall of certain dimensions and a rate of construction and asked what is the length of the wall that one worker builds. I suppose one could say that the author treats space being filled, but only very weirdly.

1 Ordinary and Technical Location in Hellenistic Greek Mathematics

It is very easy in engaging the treatment of space in Greek mathematics to turn the study into a rich philosophical study of the foundations of ancient mathematics, to produce a rival to ancient philosophical discussion, where the modern result can be no better than the underlying philosophical assumptions, that is, to fill in a cognitive story of linear, planar, and corporeal space. My goal will be more modest (but not naïve—philosophical attitudes will always come in, even in grouping together certain notions), to investigate some ways in which spatial notions play out explicitly in Hellenistic Greek mathematics and to see any relation with ancient philosophical discussions. By space here, I merely mean regions to be filled in with figures, ordinary space (fill in your favourite theory).[2] Of course, figures may be extended, unconnected figures joined, regions identified. Up and down are normally oriented in relation to a part of the diagram related to the construction ('up' indicating away, 'down' indicating towards), and only incidentally to the orientation of the viewer.[3] Direction is one way, the reader is to scan the positional relations between figures in the diagram,[4] so that diagrams are oriented to the user, as Aristotle notes (see § 2 below). Space is not measured, although figures constructed within a space may be.

Construction itself plays oddly different roles even within Euclid's *Elements*. In the Definitions of Book I, only parallel lines involve a construction in their definition (23: when extended they do not meet). Yet, in the definitions of Book XI, the sphere (Definition 14), cone (Definition 18), and cylinder (Definition 21) are

---

[2] So I resist the temptation to do a philosophy of ancient geometry, such as the a priori construction of the history of geometry that one finds in Husserl (1989). It is not that such a philosophy is completely wrong. For example, it is obvious that Hellenistic treatments of geometrical objects are extremely different from Hilbert's conception of a space, and obviously so, in the difference between treating the geometrical space as a set of points and creating a mathematics of geometrical objects (and where points are just one sort of object), whose spatial relations are determined by geometrical objects. My observation concerns what geometers commonly did, and not some ontology or epistemology of some special object, Greek Mathematics.

[3] A small survey of Euclid should make this clear. The direction is indicated by the verbal prefix, ἀνα- (commonly: up-, back-, re-) or by κατα- (sometimes: downward). For 'lead up (ἀνάγειν)' and 'lead down (κατὰγειν)', cf. Mendell (1984, p. 362) and pace Makin (2006, pp. 233–234) on the same issue. For 'draw up (ἀναγρὰφειν)', see Euclid, *Elements* II 11, where one square is drawn up on a line that is at right angles to a line on which another square is drawn (presumably, up for one and either to the right or left for the other). Cf. also I 47. The apparent contrast between 'draw up (ἀναγρὰφειν)' and 'draw down (καταγρὰφειν)' in II 7 is not at all about directions, but is between the square drawn on (from) the line and the figure *being completed* (diagonals and rectangles filled in), where 'let it be drawn down (καταγεγρὰφθω)' is related to the common word for the diagram (καταγραφή), e.g., II 8, VI 27–29, X 91–96, XIII 1–5; *Data* 58, 59. For 'be stood up (ἀνίστασθαι)', cf. *Elements* XII 10, where the common base of the cone and cylinder is drawn, but the erected figures rise, as it were, from the diagram towards the reader. For all three, cf. Mugler (1958, ad verbum). The common verb 'to construct (κατασκευζειν)' indicates nothing about direction.

[4] Netz (1999, Chap. 1).

defined by rotations. In the Postulates 1–3, someone may draw (ἀγαγεῖν: aorist active) or extend (ἐκβαλεῖν: aorist active) lines, and draw (γράφεσθαι: middle?) a circle, and in problems one is normally required (δεῖ) to do something, to construct or to find (only points and lines and magnitudes in a given ratio[5]). Furthermore, the rare uses of superposition or fitting one figure on another seems to be a change of place, i.e., unless the superimposed figure is never in between.[6] Such actions require a 'there' to do the actions. But there is no account of the 'there'.

To us, it seems a remarkable feature of Greek mathematics that one creates a figure by delimiting its limits, and what is bounded becomes the object. How the stuff between the limits comes to be, whether as space or out of nothing or by stipulation or something else is just not a part of the story (something similar happens with units appearing in Euclid's treatment of multiplication). In any case, we should resist seeing Greek geometry as about space or spatial relations. That is incidental to the objects studied. We can see this through features common to much of Hellenistic geometry. Mathematical texts do not come with any such an ontology.

Archimedes never treats of a distance between figures that is not either within a figure or a constructed line. For example, his famous first assumption in *De sphaera et cylindro*, "a straight-line is the smallest of those having the same ends," concerns lines between two points and not distances between them. Euclid only uses 'distance (διάστημα)' in the *Elements* and *Data* when referring to Postulate 3, the construction of a circle.[7]

The distinction between the figure and the space outside the figure is conceptually, but not visually marked. All we see in a diagram are the borders of the figure. It is not typically filled in, even in the fragments of ancient mathematical texts.[8] We know that there is a conceptual difference. We do not need to appeal to Aristotle, who makes this clear in his discussions of the matter of magnitude.[9] The evidence comes from the definitions of figures that we find in Archimedes, Apollonius, Euclid, and Theodosius that the figure is enclosed (περιεχόμενον) or surrounded (περιληφέν) by a boundary, lines or planes. Thus, it is not the enclosing lines or planes.[10] Unless challenged, one would not have thought otherwise. This is not to say that the language sometimes suggests that a circle is a circumference, as when a

---

[5] A point is found, namely the center of a circle at *El*. III 1, but lines in a given ratio in *El*. VI and magnitudes of a certain kind that implies a relation to other magnitudes (commensurable, incommensurable, etc.) are found in *Elements* X. Numbers are also found in VII–X. Outside Euclid, I do not find any such caution.

[6] *Elements* I 4, 8, III 24. But see also Archimedes, *Stomacheion* 2.416.15–18.4 (as filled in by recent editions of the Archimedes Palimpsest), where the figure is transposed to another place and gets another position (εἰς ἕτερον τόπον … μετατιθεμένου … καὶ ἕτεραν θέσιν λαμβάνοντος). Obviously, the more applied the work is, the less this would gently raise an eyebrow. See also Apollonius, *Conics* VI Definition 1, where conic sections, which are lines, are said to be 'equal' by superposition.

[7] In fact, διάστημα only occurs in the dative in the formula given by Postulate 3.

[8] Some of the astronomical diagrams in the *Ars Eudoxi* (Paris Gr. 1) are partially filled in, e.g., to distinguish dark parts and lit parts of a body.

[9] Cf. Aristotle, *Physics* IV 3.209b2–9, *Met*. V 17.1022a4–6

circle has been drawn about a triangle (just as the circle is enclosed by its circumference).[11] But one can just easily understand the triangle becoming a part of the circle as the circle qua circumference enclosing an area of which the triangle is a part, with its vertices on the 'circle'. It is also possible that the distinction gets loosened in discussions about great circles, after the author gets underway. How would Euclid reply to Heath's point (*Euclid* I 185), that III 10 (circles do not cut circles at more than two points) treats circles as circumferences? We would have to know more about the intentional states of each author.[12]

We would expect the issue of explicit distance (often διάστημα) between figures to arise only in issues of convergence of lines[13] and in applied mathematics, e.g., mechanics, optics or astronomy.[14] In the case of convergence, the distance is crucial but will be handled, naturally, by comparing drawn lines (is there another way in ancient mathematics?). For optics especially, one wants to have theorems about objects nearer and further from the eye. Again, this is studied through the drawn light-ray.

We shall see four central 'spatial' terms playing out in philosophical texts, 'position (θέσις)', 'place (τόπος)', 'room (χώρα)', and 'void or vacuum (κενόν)'.[15] It is useful to begin with a survey of how they are used in Greek mathematical texts, where only 'position' and 'place' are at all common terms, 'position' being mostly

[10] E.g. *Elements* I Definitions 14, 15, 18–20 and XII Definitions 9, 10, 12–14, 18, 25–28; Archimedes, *De conoid. et sphaer*. p. 246.16–19, 246.22–248.6, 248.28–250.1, etc., *De lin. spir*. p. 6.22–6, 8.23–5; Apollonius, *Conica* I Definition 1, 11.1–16, etc.; Theodosius, *De sphaera* post. 1. Contrast with Plato, *Meno* 76A, "figure is the end of a solid."

[11] Things will also get messy in later discourse. So, from late antiquity, ps.-Heron, Definition 7 (helix) says that the line that results from a line rotated about a fixed point is a circle, while Definition 27 (circle) includes an awkward definition of 'circle' that is the line that makes distances equal to all its parts.

[12] Mugler (1958, 260–1) cites no example where κύκλος means 'circumference'. For Archimedes a section (τομή) is a line, the segment (τῆμα) an area. So one assumes something similar in Apollonius, e.g., from his description of the contents of Book IV (proem 4.17–19), in how many ways a section (τομή) of a cone or a circumference of a circle meet (cf. esp. IV 25, but also throughout the treatise). Clearly, 'circle' denotes a disk, so that the section should be a line. However, at I 3, the vertical section is a triangle, which should then be a figure, unless he uses 'triangle' loosely. We should be cautious about imposing much semantic order here.

[13] So Apollonius, *Conics* II 14: When extended infinitely the asymptotes and the [hyperbola] section draw nearer to themselves and come to a distance smaller than any given distance. The proof takes the distance smaller than the given to be determined by a line drawn parallel to the tangent of the hyperbola between the asymptotes, namely, the segment between one asymptote and the intersection of the parallel line and the neighboring part of the curve. So the distance derived is actually larger than the nearest distance between the point and the asymptote, which goes unremarked.

[14] Of course, διάστημα has a different meaning in harmonics, which need not concern us here.

[15] In the following word counts, taken from the TLG_E, I take the following authors without distinguishing works, Autolycus, Euclid, Archimedes, Apollonius, Hypsicles, Serenus. For Pappus, I search only the *Collectio*. For Theodosius, I separate out *De sphaera*, but mention counts for *De habit*. and *De dieb. et noct*. For Heron, I restrict searches to *Pneumatica, De automatis, Frag. de horoscopiis, Mechanicorum frag., Catoptrica, Metrica, Dioptra,* and *Belopoeica*.

either 'in position (θέσει)' or, less commonly, 'having a position (θέσιν ἔχειν)', and 'place' having one of two meanings, a region and a line, plane or solid region where every object of a certain kind has a given property.

In looking at mathematical terminology, a good place to start is Mugler (1958). Mugler finds no uses of 'room (χώρα)' in geometry between Plato and Proclus, where, he says, it is a synonym for 'area (χωρίον)'; indeed, the word is very rare in geometers other than Heron.[16] However, Heron uses 'room (χώρα)' mostly in the more applied, mechanical works *Pneumatica* and De *automatis*, where the areas are less classified and rely more on the drawing. In any case, χώρα has little to do with questions about space and location in geometry.[17]

With regard to the related word 'area (χωρίον),' I note one small, perhaps Aristotelian, feature about its very common use in all geometers. If 'figure (σχῆμα)' and words for particular figures denote an object *qua* having a certain quality, namely shape, 'area (χωρίον)' denotes the object *qua* so-much or being so much (quantum). Here the primary referent may be a rectangle (e.g. *Elements* X) because that is the most common sort of area considered (as, in English, 'eggs' normally and in a general context refers to chicken eggs, even though the word never means 'chicken eggs'). In other words, area is not an abstract entity. That said, it is often the case that 'area (χωρίον)' is just a general word for any two dimensional object,[18] ἐμβαδόν being the more common word for an area qua measured (compare στερέον as 'solid' and as 'volume').

'Void (κενόν)' does not appear in Mugler, nor, with one or two unimportant exceptions, in the core mathematical texts.[19] A method for measuring disorderly figures by measuring an absence of body appears in Heron. In *Metrica* II 12–13, he introduces a technique for measuring shells determined by one volume separated from a similar, smaller volume, where the result is the difference between the two volumes, and recommends it as a method for measuring washtubs, conch shells (with an insignificant error), arches, and vaults. He then says, "Given that the surface within is hollow/concave (κοιλῆς), that is, empty/void (κενῆς), again each of them will be the excess of two similar segments." At the end of *Metrica* II (20), he commends Archimedes' method of measuring a disorderly body by placing it in a full rectangular tank and measuring the emptied place (τὸν ἐκκεκενωμένον τόπον)

[16] χώρα occurs: Autolycus (0), Euclid (1), Archimedes (1 ordinary use in *Sand Reckoner*), Apollonius (0), Theodosius (0), Hypsicles (0), Serenus (0), Heron (69), Pappus (1).

[17] We can supplement Mugler's remarks by noting that in ps.-Heron (e.g. *Mensurae* and *Liber geeponicus*), the word is used to mark an area treated as a figure, e.g., *De Mensuris* 54.1.1: isosceles triangular χώρα (χώρα τρίγωνος ἰσοσκελής). It is not quite a synonym for χωρίον (area), except in its pre-Euclidean uses (cf. Mugler 1958, p. 451, Plato, *Meno* 82B, square area [τετράγωνον χωρίον]).

[18] See the citations in Mugler (1958, ad verbum).

[19] κενόν occurs: Autolycus (0), Euclid (0), Archimedes (1, 'empty speech'), Apollonius (0), Theodosius (0), Hypsicles (0), Serenus (0), Heron (59: *Pneumatica* (56), *De automatis* (2), *Metrica* (1)), Pappus (1, 'with empty hands'). In Heron, the verb 'to empy (κενόω)', the noun 'emptying (κένωσις)', and two compounds, 'to empty together (συνκενόω)' and 'to empy out (ἐκκενόω)', occur 71 times.

in the tank when the object is removed, and a second method where the body cannot be moved, of covering the body with wax or clay to form a rectangular figure that can be measured and then measuring the wax or clay, now molded into a rectangular figure. It's fairly obvious that, in general, Heron has a more physical and environmental treatment of geometrical objects anyway. Or rather, none of the extant works is a work of geometry. All are applied.[20] But one might wonder whether the technique brings along a tiny shift in how to conceive of volumes in a practical context.[21] I am somewhat sceptical, since, I think, it at most reflects Heron's own views about applied geometry. First, the use of 'void (κενόν)' looks like the ordinary use, indeed, as does the use of 'hollow/concave (κοιλῆς)', for Heron does not say that the region is empty of everything. One might be led to think that something unusual is going on by the fact that the inner figure in II 12–13 and

[20] Heron's division of methodologies in the *Metrica* is somewhat subtle. There are three levels of discussion, the purely abstract, the metrical (arbitrary numbers in pure units (μονάδες), not in terms of standard measures), and loose applications where physical objects are mentioned. Through most of the treatise and depending on the complexity of the problem, he often gives an analysis abstractly followed by a synthesis with numbers. Furthermore, metrica is throughout an application of geometry. So at III 10, p. 160.14–7 (cf. also I 7, 8, esp. p. 20.6), he contrasts the numerical presentation with a geometrical demonstration (i.e., the abstract presentation), where numbers are not used. Accordingly, he seems to mean in I 6 (p. 16.11–14), "We acted up to now taking into account (or: calculating on/by the) geometrical demonstrations; in what follows we will make measurements according to the analysis through the synthesis of the numbers (Μέχρι μὲν οὖν τούτου ἐπιλογιζόμενοι τὰς γεωμετρικὰς ἀποδείξεις ἐποιησάμεθα, ἑξῆς δὲ κατὰ ἀνάλυσιν διὰ τῆς τῶν ἀριθμῶν συνθέσεως τὰς μετρήσεις ποιησόμεθα.)," that the employment of numbers follows the geometrical demonstration of the theorem, while later he will provide a geometrical analysis followed by a synthesis that uses numbers. For this is somewhat what he has done and somewhat what he will do, although not until I 10. In this regard, the distinction between orderly and disorderly figures is just that between those where the figures have been studied and have been given an abstract treatment for finding area or volume and those where no clear geometrical reduction is known so that one needs to use metrical or even physical means, whether approximations (*Metrica* I 39) or the methods discussed above (II 20), i.e., those to which geometry can be applied directly and those to which it cannot, for whatever reasons. I would be very surprised if Heron, or any ancient mathematician, thought there was a distinction between geometrical figures per se and non-geometrical figures or even whether a construction employing a mechanism is thereby a part of mechanics, an error that famously goes back to Descartes. Hence, given that the context is never purely geometrical, we expect that geometrical methods will be used but blended with more metrical methods, however these are to be distinguished, although Heron does some work on this in the *Metrica*. Nonetheless, we do not expect a distinction between geometrical plane figures and non-geometrical plane figures (pace Tybjerg 2004, esp. pp. 39–43). Furthermore, just because geometrical shapes get their names from physical objects they represent, it does not follow that they are those objects, and just because a mathematician uses vivid language (e.g., 'pierce') it does not follow that he is thinking physically and not geometrically (again, pace Tybjerg). That said, this might become irrelevant if we can discern anything of Heron's views on geometry (see § 6).

[21] The method appears in late antiquity in works which are probably based on Heron. Here, 'vacuum (κένωμα)' occurs in *De mensuris* (5) and *Stereomtrica* (24), sometimes to measure the vacuum and solid and then the vacuum in order to determine the volume of the solid. The word might have shown up in other works of Heron in this context, but he uses it only once, quite differently (*De automatis* 26.2).

the emptied bath are not figures. After all, this sort of determining areas or volumes by subtraction is as old as geometry, but here they are regarded as devoid of at least the figure in question, but with emphasis. Much more significant is the water, wax, and clay that keep their volume in changing shape, revealing that metrica can be more concrete than the mere adding of measures would suggest. In any case, it is very optimistic to read ontological commitments about geometry from a treatment of applied mathematics, as is reading the end of *Metrica* II as providing the level of abstraction for the entire work. Nonetheless, we shall see in § 6 that Heron takes into account practical considerations in apparently purely geometrical constructions.

Thirdly, let us turn briefly to ‘position (θέσις)’.[22] ‘In position’ occurs in the *Data* and the *Porisms*, a lost work of Euclid, as reported by Pappus.[23] Here is how Euclid defines ‘given in position’ in the *Data* (Definition 4):

> Points and lines and angles that always keep the same place are said *to be given in position*.
>
> δʹ. *Τῇ θέσει δεδόσθαι* λέγονται σημεῖά τε καὶ
> γραμμαὶ καὶ γωνίαι, ἃ τὸν αὐτὸν ἀεὶ τόπον ἐπέχει.

‘Place’ is not mentioned again in the *Data*. In other words, Euclid here treats ‘place’ as an intuitive notion from which one may define the technical notion. Objects do not necessarily keep their place in the *Elements*, as already noted, but this does not seem to be the point here. As a practical matter, one may use the position of something that is given in position in showing that other things are given in position. So it establishes the spatial relation between different things so given. This seems to be all Euclid really needs to mean and all that other geometers need in using the expression. So things have a position (θέσιν ἔχει) when they can be located in a configuration, when their position is given. A circle may be given in position without any point on it being given in position. A point on a circle may be given in position without the circle itself being given in position.

---

[22] θέσις occurs: Nominative: Autolycus (0), Euclid (10), Archimedes (1), Apollonius (0), Theodosius (0), Heron (7), Pappus (6). Genitive: Autolycus (1, in *De ort. et occ.*), Euclid (2), Archimedes (0), Apollonius (0), Theodosius (0), Heron (3), Pappus (28). Dative singular: Autolycus (0), Euclid (210), Archimedes (1), Apollonius (50), Theodosius (0), Heron (26), Pappus (185). Dative plural: Autolycus (0), Euclid (0), Archimedes (0), Apollonius (0), Theodosius (0), Heron (0), Pappus (0). Accusative singular with ‘having’: Autolycus (18: 4 in *De sphaera*, 14 in *De ort. et occ.*), Euclid (41), Archimedes (0), Apollonius (0), Theodosius (0), Heron (84), Pappus (20). Other accusatives: Autolycus (0), Euclid (3), Archimedes (1), Apollonius (0), Theodosius (0), Heron (10: 5 with ‘find (εὑρίσκειν)’, 1 with ‘get (λαμβάνειν)’, 1 with ‘I have (ἔχω), and 3 others), Pappus (8: 7 with ‘get (λαμβάνειν)’ and 1 with ‘preserve (φυλάττειν)’). The distribution of datives largely reflects the language of analysis and ‘given’ (esp. Euclid, Apollonius, Heron, and Pappus), while the accusative with ‘have’ or ‘get’ is largely astronomical or at least involves moving figures. For example, if we add the two astronomical works of Theodosius, there would be 29 occurrences of the accusative, 28 with ‘have’. So too all occurrences in Autolycus. The Euclid passages are all from the *Phenomena*, and the three accusatives without ‘having’ from the *Optics*. Additionally, Serenus and Hypsicles do not use the word.

[23] Pappus, 636.18–30, 648.19–20 and following. Presumably, the *Places* (*Loci*) *on a Plane* (*Τόποι ο πρὸς ἐπιφανείᾳ*), also mentioned by Pappus (first passage), used ‘given in position’.

One might presume (perhaps from thinking about basic properties of atoms in Democritus) that whereas position is a relative concept, place is absolute so that this will mark the difference between them. But, while important in astronomy and any science where things are absolutely located, it is hard to see what role absolute place could play in pure geometry, or in any science where things do not need absolute position for the theorems to be stated. Even in Archimedes' statics, all that is needed is up and down. So we should expect, and do find, a different distinction.

Furthermore, different authors might have different tastes. In Autolycus, Euclid's *Phenomena*, and Theodosius, only circles (or circular-arcs in Theodosius) have a position (θέσις). However, in Autolycus the sun may have a place (τόπος), in Euclid places are rising and setting points on the horizon, while Theodosius doesn't use the term at all. Euclid, the only one of these to use 'room (χώρα)', uses it for the region where the North Star turns. So he keeps place to the horizon. Quite differently, in the commentary of Hipparchus on Aratus and Eudoxus, stars and constellations have positions, but, besides places on the earth (e.g., the regions of Greece), 'place' is used only five times, but in ways different from the others here. From Hipparchus' comments, however, it would seem that in his astronomy Eudoxus, like Euclid in his astronomy, treated place as absolute.[24]

Finally, 'place (τόπος)', is also important in mathematics.[25] Here, Mugler identifies two principal uses. The first is a region defined by lines or surfaces, while the second, technical sense is the familiar, locus use, although the two uses can be merged. In the technical sense, in a 'locus (τόπος)' theorem, a place or locus is a given point, line, area or region, solid or region, such that every object of a determined sort within the locus has certain additional properties (συμπτώματα). For example, in Aristotle, *Meteorologica* III 5, every point in a given ratio (not 1:1)

[24] Two of these are quotations from Eudoxus. Hipparchus, *In Arati et Eudoxi phaen*. I 4.1, where the issue is whether "there is a star always remaining at the same place (ἔστι δέ τις ἀστὴρ μένων ἀεὶ κατὰ τὸν αὐτὸν τόπον)," and Eudoxus' claim (I 9.2), "The sun appears as making a difference according to the places of its turnings (τῶν κατὰ τὰς τροπὰς τόπων)." The places in the last are relative to the observer on earth, as in Euclid's *Phaenomena*, while Euclid uses 'room (χωρα)', where Eudoxus used 'place (τόπος)' for the position of the North Star, perhaps with a different theory. Observe that this amounts to treating the fixed stars as changing place but not position (relative to each other), which would encourage one to use 'position' for the stars and not 'place'. The pole (quoting Eudoxus) has place. If this is right, Eudoxus treats place as absolute. Two more passages concern the fact that the constellations of the zodiac do not extend over their proper places, i.e., are more or less than their twelfth-part (II 1.8, 4.4). Finally, it may indicate a region (I 8.6), namely the region between the River (Eridanus) and the rudder of the Argo that is not large.

[25] τόπος occurs, where I restrict Theodosius to *De sphaera,* and Pappus to the *Collectio*: Autolycus (4, in *De ort. et occ.*), Euclid (64), Archimedes (14), Apollonius (27), Theodosius (0), Hypsicles (1, also a degree arc (μοῖρα) is called τοπική (not mentioned again since it is the primary notion, as opposed to a time degree, χρονική (frequent)), Serenus (1 as 'topic'), Heron (134), Pappus (95). Proclus (In Eucl. 194.25–195.5, on Com. Notion 1) quotes Apollonius as holding that figures are equal that occupy the same τόπος to show the transitivity of equality. Additionally, τόπος occurs 31 times in the astronomical works of Theodosius.

of distances from two given points lies on a given circle, where the ratio is given iff the circle is given. The locus here would be the circumference (Aristotle does not use the term in this way).

Proclus defines a ‘local-theorem (*In pr. Eucl. el.* 394.16–395.1, trans. Morrow (1992), with changes):

> I call ‘local-[theorem]s’[26] those in which the same property occurs throughout the whole of a certain locus, and I call ‘locus’ a position of a line or surface producing one and the same property of a line or a surface producing one and the same property. Some local-[theorem]s refer to lines, others to surfaces; and since some lines are plane and others solid—plane lines being those which, like the straight-line, lie in a plane and whose generation is simple, and solid line those which are produced by some sectioning of a solid figure, like the cylindrical helix and the conic lines—I should say further that of local-[theorem]s referring to lines some have a plane and others a solid locus.

> καλῶ δὲ τοπικὰ μέν, ὅσοις ταὐτὸν σύμπτωμα πρὸς ὅλῳ τινὶ τόπῳ συμβέβηκεν, τόπον δὲ γραμμῆς ἢ ἐπιφανείας θέσιν ποιοῦσαν ἓν καὶ ταὐτὸν σύμπτωμα. τῶν γὰρ τοπικῶν τὰ μέν ἐστι πρὸς γραμμαῖς συνιστάμενα, τὰ δὲ πρὸς ἐπιφανείαις. καὶ ἐπειδὴ τῶν γραμμῶν αἱ μέν εἰσιν ἐπίπεδοι, αἱ δὲ στερεαί, – ἐπίπεδοι μέν, ὧν ἐν ἐπιπέδῳ ἁπλῆ ἡ γένεσις, ὡς τῆς εὐθείας, στερεαὶ δέ, ὧν ἡ γένεσις ἔκ τινος τομῆς ἀναφαίνεται στερεοῦ σχήματος, ὡς τῆς κυλινδρικῆς ἕλικος καὶ τῶν κωνικῶν γραμμῶν – φαίην ἂν καὶ τῶν πρὸς γραμμαῖς τοπικῶν τὰ μὲν ἐπίπεδον ἔχειν τόπον, τὰ δὲ στερεόν.
>
> 22 γένεσις ver Eecke Morrow νόησις Friedlein mss

Proclus own classification of local-theorems is probably derived from Apollonius, or at least one might infer that from the more detailed discussion in Pappus.[27]

Proclus (395.3–12, 396.2–9) then goes on to identify three theorems in the *Elements* as local-theorems: *Elements* I 35 (parallelograms on the same base and on the same parallels are equal), where the locus is the region between the parallel lines;[28] as well as III 21 (angles on the same segment of a circle are equal) and 31 (angles in a semicircle are right, in a larger segment smaller, and in a smaller

[26] ‘Theorem’ is understood from context; however, I prefer to translate τοπικά as ‘local’ than ‘locus’, reserving that for the region. However, the usual description is not quite an idiosyncracy of Proclus—Pappus uses it once (*Collectio*, VII 652.2), but normally just calls a local theorem a ‘locus (τόπος)˜, which at least obviates the issue that some are problems.

[27] Cf. Jones (1986, pp. 539–46).

[28] Proclus’ description of *Elements* I 35 as a local-theorem is somewhat awkward, since he takes the theorem as concerning parallel lines drawn from the base, which is only implicit in the theorem. For our purposes, Proclus’ example is fine, namely as showing that within the τόπος all parallelogram areas on the same base have the same salient property.

larger), where the locus is the circle or disk,[29] as well as a theorem on hyperbolae (parallelograms inscribed in the asymptotes and hyperbola are equal), where the locus is the hyperbola.[30]

It is enough for our purposes that a locus need not be a finite figure or line, but can be an infinite line, surface, or solid, so long as it is partially bounded so that there are points, lines, etc. that are not included in the locus, as this marks the distinction between a local theorem and a general theorem.

Proclus reports local-theorems (cf. 67.20–23), or at least this sort of problem as going back at least to Hermotimus of Colophon, who would have been younger than Eudoxus (say, mid-4th cent. BCE), but does not attribute to him the actual discovery of local theorems. So far as I can tell, the terminology is not in Aristotle, at least not in the two places where we find local theorems, *Meteorologica* III 3 and 5. It is important to keep in mind that Euclid's definition of circle (*Elements* I Definition 15) is a local definition that does not use the word 'locus (τόπος)'.

> A circle is a plane figure enclosed by one line [which is called circular arc], such that all the straight lines falling on it from one point among those lying within the figure [to the circular arc of the circle] are equal to one another, …

Κύκλος ἐστὶ σχῆμα ἐπίπεδον ὑπὸ μιᾶς γραμμῆς περιεχόμενον [ἣ καλεῖται περιφέρεια], πρὸς ἣν ἀφ' ἑνὸς σημείου τῶν ἐντὸς τοῦ σχήματος κειμένων πᾶσαι αἱ προσπίπτουσαι εὐθεῖαι [πρὸς τὴν τοῦ κύκλου περιφέρειαν] ἴσαι ἀλλήλαις εἰσίν.

---

[29] Proclus does not actually say this, but we may infer it from his drawing an analogy with I 35.

[30] Morrow (1992, p. 311 n. 70 ad 395.20) follows ver Eecke in identifying the theorem with Apollonius, *Conics* II 12. However, the theorems do not seem to be the same in one way that is important for this discussion. Proclus seems to be saying that if one draws a line from the hyperbola parallel to the asymptote to the opposite asymptote and completes the parallelogram, it will be equal to any other drawn similarly. Apollonius states that if from a point on the hyperbola lines are drawn to the two asymptotes and from another point on the hyperbola parallel lines are drawn to the asymptotes, the rectangles formed by the lines from the one point are equal to those from the others. Except for special cases, the rectangles are not inscribed in the asymptotes at all. The two basic properties of hyperbolae are trivially related, however. Suppose the lines in Apollonius are drawn as in Proclus, and let the respective lines be a_1, a_2, b_1, b_2, so that by *Conics* II 12, Rectangle(a_1, a_2) = Rectangle(b_1, b_2). Since the parallelograms are drawn in equal angles (the angle between the asymptotes), let it be α, the parallelograms will also be equal. In modern terminology, the area of each parallelogram will be: Rectangle (a_1, a_2 Sin(α)) = Rectangle (b_1, b_2 Sin(α)). Nonetheless, the theorem in Apollonius is more apposite to Proclus' point. For it is unclear why Proclus takes the τόπος of the hyperbola theorem to be the line and not the region between the asymptotes and the hyperbola. For *Elements* I 35 and the hyperbola theorem concern parallelograms and not lines. Perhaps this is what he intended, to draw the parallel, but then digressed to say that the line is a solid line (constructed by a section). One could read *Conics* II 12, however, as concerned with points on a hyperbola or as concerned with lines drawn between the hyperbola and the asymptotes, which would be more akin to Proclus' examples from the *Elements*. Clearly, the terminology is a little loose.

Additionally, Knorr argued that the conics were probably also originally defined in this way.[31] For early texts, Euclid wrote a lost treatise, books on τόποι in planes,[32] while local-propositions also occur in Euclid, *Optics* (37, 38, 44–49), but with a slightly different terminology.[33] Is there any philosophical significance in all this for understanding Greek philosophy of mathematics, concepts of location or space? I do not know. Obviously, a place is defined relative to some other object, distinct from it, given points or lines, etc.

The first use of τόπος mentioned by Mugler is more immediately relevant to our purpose. This is a region determined by some lines or surfaces. We find this use in the discussion of the elements in Plato's *Timaeus*, and one could argue that Aristotle's own discussion of the places of the elements is an attempt to fit this in as well. In the *Elements*, it only occurs in III 16 for the region between a circumference and a tangent, but it also occurs in Apollonius' *Conics*, in propositions about hyperbolas, as the region within an angle or between an angle and the hyperbola;[34] in Archimedes, mostly in less purely mathematical contexts, as length;[35] and in Pappus, with various senses perhaps representing the variety of his sources, but sometimes for the area bounded by different figures and once for the region about a point.[36] I have already noted that in Heron it can be a volume.[37] So let's think of this as an ordinary use. Does it have philosophical or ontological commitments?

Consider Aristotle's category of 'Where?' in the *Categories*, where 'in the Lyceum' gives a region within a boundary. Honing the region to the location of the object is also the motivation behind the technical notion of primary place that we find developed in *Physics* Δ 2. So far as anyone in the Academy or Lyceum is concerned, the notion of a bounded region within which an object may be found must be compatible with any reasonable treatment of place. The dialectical issue would be whether it is compatible with one's opponent's theory.

---

[31] Cf. Knorr (1986, pp. 62–6), Jones' response (1986, pp. 572–599). Knorr apparently intended to reply to Jones, but I am only aware of a sketchy draft.

[32] Pappus, *Collectio* VII 636.21. I am assuming that Euclid precedes Aristaeus. The list in Pappus is not chronological. Also, cf. Apollonius, *Conica* I prol. 30–7.

[33] A place is a position of the eye (a point) or of the seen object (a line), e.g. the places (τόποι) in *Optica* 47 are each points on a semicircle where, with the eye placed at any of them, two adjacent lines will appear equal.

[34] Apollonius, *Conica* I 32, 35, 36, II 13, 32, 33, 42, 49, III 24, IV 42, 51.

[35] Cf. Archimedes, *Sand Reckoner* 236.11–12 (length of 25 poppy seeds as a τόπος), where the other occurrences concern the place of the eye (222.22, 224.4, 12, 29, 226.8), and *Floating Bodies* I 4 328.4 (closed region), II 10 386.3, 392.15, 408.13, 412.14 (area: the base will be cut off at a larger place, etc., by the liquid), but also Stomacheion referred to in note 6.

[36] Pappus, *Collectio* IV 224.15, 252.18 (closed areas bounded by different figures), 242.15 (region under cochloid), 244.26 (region between cochloid and base line); V 306.8, 9, 15 (ὁ περὶ τὸν αὐτὸν τόπος, i.e., the region about the same point that can be filled up with identical equilateral figures). On this last, cf. also ps.-Heron, Definition 71.3–9 and 136.45.

[37] Cf. also the late, ps.-Heron, *Geometrica* 4.15.40–57, cf. 4.8, *Stereometrica* 42a12–13.

I note, however, that in the one quasi-mathematical place where you might expect Aristotle to use τόπος in this way, he doesn't. At *De caelo* A 5.271b28–33, he is discussing an infinite ray rotating about a point. From Aristotle's perspective, the argument is physical and not per se mathematical—the infinite ray is a physical line of the rotating heaven from the center of the earth:

> For if the circularly rotating body is infinite, the lines extended from the middle will be infinite. But the interval [or extension] of infinite lines is infinite. I means by 'interval of lines', that where it is not possible to take a magnitude outside touching the lines [that is part of the interval]. And so, this must be infinite, since for finite-lines, [the interval] will always be finite.

271b28 Εἰ γὰρ ἄπειρον τὸ κύκλῳ φερόμε-
νον σῶμα, ἄπειροι ἔσονται αἱ ἀπὸ τοῦ μέσου ἐκβαλλόμε-
271b30 ναι. Τῶν δ' ἀπείρων τὸ διάστημα ἄπειρον· διάστημα δὲ λέγω
τῶν γραμμῶν, οὗ μηδὲν ἔστιν ἔξω λαβεῖν μέγεθος ἁπτόμε-
νον τῶν γραμμῶν. Τοῦτ' οὖν ἀνάγκη ἄπειρον εἶναι· τῶν γὰρ
πεπερασμένων ἀεὶ ἔσται πεπερασμένον.

In effect, the distance is the region within an angle. So a Greek mathematician might well have used the word 'place (τόπος)' instead of 'interval (διάστημα)'. If so, then the word 'interval' might well be doing the work that 'place' conceived as interval would be doing. Given his definition of place as the inner limit of a containing body and his resulting view that the finite universe has no place, is Aristotle queesy here? If so, is he queesy about calling an infinite region a τόπος, especially if it is unbounded in one direction, even in a reductio ad absurdum (cf. his criticism of Anaxagoras at *Physics* III 205b1–b24). Or is he queesy about something else?[38] Or is the terminology just not there? We cannot know.

The technical use of 'place (τόπος)', as presented by Proclus, is really just this first use in a particular technical context, where certain objects in the determined region have certain properties. Of course, the technical use expands the ordinary use, since in the ordinary use, lines are not places.

Besides these uses of 'place (τόπος)' mentioned by Mugler, I might point out that there are other uses in Greek mathematical texts not mentioned by him, but occurring in definitions or in Heron and later texts, some of which have already been indicated.

[38] Cf. Heron, *Dioptra* 6.5–7: "Nevertheless, let us examine the places given in the interval between the points, how they are related to one another and the initial given points (οὐ μὴν ἀλλὰ καὶ τοὺς δοθέντας τόπους ἐν τῷ μεταξὺ διαστήματι τῶν σημείων ἐπισκεψώμεθα, πῶς ἔχουσι πρὸς ἀλλήλους καὶ τὰ ἐξ ἀρχῆς δοθέντα σημεῖα)".

1. a τόπος can be the ordinary location of objects in a configuration, as Euclid, *Data* Definition 9.[39]
2. τόπος is a region (dare I say space) where a construction is to take place, e.g. the figures that can fill up a place (τόπος).
3. A τόπος can be a physical length or an area (in Archimedes), even a volume (Heron).
4. The word τοπικός is used both as a metamathematical term (as in Proclus' definition above) and an adjective for something having a place.

To summarize, of the four terms that become important in philosophical discussion in the Hellenistic Age, only 'position (θέσις)' and 'place (τόπος)' are important notions. A partially determined region outside object is a place, but not the entire region or the entire plane or three dimensional world. Place is not space. But in its many uses, it is not the same as position either. For to be given in position might well be to be given in some place, and certainly lines that are loci should be given in position (whether as initially given or as the goal of a problem). But in local theorems, the place or locus is where things get positioned. So there seems to be a general functional difference between the two terms (no doubt with many exceptions). A place in both primary senses is a region within which things have positions. This holds whether the place is open ended (e.g., the region closed by a line) or closed (e.g., a quadrant, within which the loci that make up the quadratrix will find position). Usually, but far from always, things are just given in position when they are given.

Although such notions take on mathematical significance, regions where objects may be located, objects given in positions, it is unclear that one can say much about any relation to philosophical discussions. This is where we shall now turn.

2 Aristotle on the Location of Mathematical Objects

Let's begin with a quick survey of views of place in Classical and Hellenistic physics. Here our task is threefold, to present views about place, views about mathematical objects, and the sense in which mathematical objects can have location. Since it encompasses all views on place through the Hellenistic Age, I like to start with a tableau that one can derive from Aristotle (Fig. 1).[40]

In the *Physics*, Aristotle argues that the place of a body is the inner limit of a container of the body. He then argues that points do not have place because they cannot be distinct from the inner limit of a container. By the same argument lines

[39] Cf. ps.-Heron, Definition 4 on the definition of a straight line as that which when rotated about its end points always keeps the same place (τὸν αὐτὸν ἀεὶ τόπον ἔχουσα) and 8 on the definition of surface as the limit of body and of place, and 9 on congruence. Definition 11 (solid and place as having three dimensions) obviously is incompatible with Aristotle's *Physics*, but not with his *Categories* or almost any post-Theophrastus view of place.

[40] Cf. Mendell (1987, p. 219) and (2005, p. 355). the diagram is slightly different here, to emphasize the two possible ways an extension might be separate from the contained.

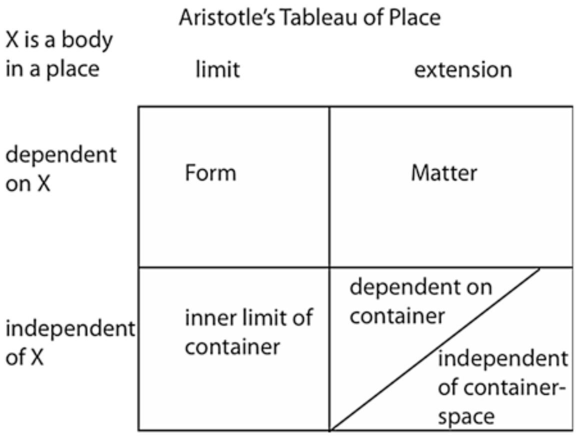


Fig. 1 Aristotle's tableau of place

and surfaces should not have place, but that would be our inference. Later, Aristotle also defines two bodies as being in contact if they have their extremities together, that is in one primary place (*Phys*. E 3.227b21–23). So when two bodies are in contact, the extremities may be distinct but are together. So if two bodies are in contact at a point, it follows that points have primary place. This contradiction, which bothered G.E.L. Owen,[41] for example, requires us to distinguish two approaches, one very strict and one loose. However, the points in question, are associated with physical and not purely mathematical bodies.

Nonetheless, one might think that the official Aristotelian position should be: physical bodies smaller than the universe and below the outer shell of the heaven have place; mathematical objects have position, but not place. Aristotle seems to confirm this in several discussions, some less decisive than others, e.g., at the beginning of his discussion of place, where he is surveying puzzles (he need not be asserting his own beliefs); or in an argument against Academic, intermediate mathematicals, which should not have place on the view of the Academics (*Met*. N 5.1092a17–21):

> But at the same time providing a place to mathematical solids is also absurd (for place is a proprium of particulars, whence they are separate in place, but the mathematicals are not anywhere), that is, saying that they will be somewhere, but not what place is.

> ἄτοπον δὲ καὶ τὸ τόπον ἅμα τοῖς στερεοῖς τοῖς μαθηματικοῖς ποιῆσαι (ὁ μὲν γὰρ τόπος τῶν καθ' ἕκαστον ἴδιος, διὸ χωριστὰ τόπῳ, τὰ δὲ μαθηματικὰ οὐ πού), καὶ τὸ εἰπεῖν μὲν ὅτι πού ἔσται, τί δέ ἐστιν ὁ τόπος μή.

[41] It is remarkable how often Owen returned to the issue, cf. Owen (1961), (1970), (1976), and (1986a).

In *De caelo* Γ 6.305a22–31, Aristotle argues what is unmoveable and mathematical, as lacking weight, cannot be in place.

> In fact, the elements cannot be composed from some body. For it will follow that another body will be prior to the elements. If this will have heaviness or lightness it will be one of the elements, but without having downwards-inclination it will be unmovable and mathematical. Being of this sort, it will not be in a place. For [in the place] in which it rests, it will also be able to move. And if it [can] by force, it will do so contrary to nature; and if not by force, according to nature. And so if it will be in place and somewhere, it will be one of the elements. And if it is not in a place, nothing will be [composed] from it. For what comes to be and that from which it comes to be must be together.

> Ἀλλὰ μὴν οὐδ' ἐκ σώματός τινος ἐγχωρεῖ γίνεσθαι τὰ στοιχεῖα· συμβήσεται γὰρ ἄλλο σῶμα πρότερον εἶναι τῶν στοιχείων. Τοῦτο δ' εἰ μὲν ἕξει βάρος ἢ κουφότητα, τῶν στοιχείων ἔσται τι, μηδεμίαν δ' ἔχον ῥοπὴν ἀκίνητον ἔσται καὶ μαθηματικόν· τοιοῦτον δὲ ὂν οὐκ ἔσται ἐν τόπῳ. Ἐν ᾧ γὰρ ἠρεμεῖ, ἐν τούτῳ καὶ κινεῖσθαι δυνατόν. Καὶ εἰ μὲν βίᾳ, παρὰ φύσιν, εἰ δὲ μὴ βίᾳ, κατὰ φύσιν. Εἰ μὲν οὖν ἔσται ἐν τόπῳ καί που, ἔσται τι τῶν στοιχείων· εἰ δὲ μὴ ἐν τόπῳ, οὐδὲν ἐξ αὐτοῦ ἔσται· τὸ γὰρ γινόμενον, καὶ ἐξ οὗ γίγνεται, ἀνάγκη ἅμα εἶναι.

The argument here is against a physics such as atomism that makes natural bodies have only mathematical properties. However, the argument seems general: mathematical objects cannot be in place in so far as this requires that they be movable.

Thus it may be surprising to some that, in discussing how efficient causes work in *De gen. et corr.* I 6.322b9–3a12, Aristotle denies this:

> And so, just as nearly every other noun is also said in many ways, some homonymously, others from other (uses) and prior (uses), so it holds too in the case of contact. Nevertheless, what is said principally belongs to things having position, and position to just those things that also have place. For one must attribute similarly contact and place also to mathematicals, whether each of them is separate or exists in some other way. And so, if, just as was defined earlier (*Physics* E 3), being in contact is having the limits together, these would be in contact with one another which, by being determinate magnitudes and having position, have their limits together. And since position belongs to those things to which place also belongs, and the first differentia of place is up and down, and these sorts are opposites, all things touching one another would have weight or lightness, either both or one. And these sorts are affective and effective. Thus it is obvious that those things are of a nature to touch one another which, by their being divided magnitudes, have their limits together, as they are causing motion and are moved by one another.

> Σχεδὸν μὲν οὖν, ὥσπερ καὶ τῶν ἄλλων ὀνομάτων ἕκαστον λέγεται πολλαχῶς, καὶ τὰ μὲν ὁμωνύμως τὰ δὲ θάτερα ἀπὸ τῶν ἑτέρων καὶ τῶν προτέρων, οὕτως ἔχει καὶ περὶ ἁφῆς. Ὅμως δὲ τὸ κυρίως λεγόμενον ὑπάρχει τοῖς ἔχουσι θέσιν, θέσις δ′ οἷσπερ καὶ τόπος· καὶ γὰρ τοῖς μαθηματικοῖς ὁμοίως ἀποδοτέον ἁφὴν καὶ τόπον, εἴτ′ ἐστὶ κεχωρισμένον ἕκαστον αὐτῶν εἴτ′ ἄλλον τρόπον. Εἰ οὖν ἐστίν, ὥσπερ διωρίσθη πρότερον, τὸ ἅπτεσθαι τὸ τὰ ἔσχατα ἔχειν ἅμα, ταῦτα ἂν ἅπτοιτο ἀλλήλων ὅσα διωρισμένα μεγέθη καὶ θέσιν ἔχοντα ἅμα ἔχει τὰ ἔσχατα. ’Επεὶ δὲ θέσις μὲν ὅσοις καὶ τόπος ὑπάρχει, τόπου δὲ διαφορὰ πρώτη τὸ ἄνω καὶ τὸ κάτω καὶ τὰ τοιαῦτα τῶν ἀντικειμένων, ἅπαντα τὰ ἀλλήλων ἁπτόμενα βάρος ἂν ἔχοι ἢ κουφότητα, ἢ ἄμφω ἢ θάτερον. Τὰ δὲ τοιαῦτα παθητικὰ καὶ ποιητικά· ὥστε φανερὸν ὅτι ταῦτα ἅπτεσθαι πέφυκεν ἀλλήλων, ὧν διῃρημένων μεγεθῶν ἅμα τὰ ἔσχατά ἐστιν, ὄντων κινητικῶν καὶ κινητῶν ὑπ′ ἀλλήλων.

Here is a summation of the argument that we are interested in.

1. Things that have their limits together are in contact (Phys. E 3).
2. The primary sense of ‘contact’ belongs to things that have position.
3. Things that have position have place.
4. The first differentia of place is up and down.
5. [Things that are up or down have heaviness or lightness or both].
6. ∴ Things in contact have heaviness or lightness or both.
7. Things that have heaviness or lightness are affective or effective (with respect to motion, see step 9).
8. Divided magnitudes have their limits together.
9. ∴ Divided magnitudes are effective and affective with respect to motion.

I assume that the divided magnitudes at the end of the argument are perceptible or physical magnitudes. Yet in the middle of this argument, Aristotle says (a1–3), “For one must give similarly contact and place also to mathematicals, whether each of them is separate or exists in some other way.” Clearly, one must block the inference that mathematical magnitudes are effective and affective. As most commentators agree, this must occur at step 4.[42] At *Physics* Δ 1.208b22–26, in setting out the worthy opinions (endoxa) that place exists, Aristotle argues that we must

[42] Similar analyses may be found in Joachim (1922, ad loc.), de Haas et al. (2004, ad loc.), etc. Buchheim (2010) avoids the interference by switching (2) to its converse (everything in place has position) and altering (1), but at the sacrifice of removing 1–2 as part of an argument. As for our concern, it will still be the case that mathematical objects have place, so that one will still need either to read the rest of the argument very differently (per Buchheim) or to block the inference that mathematical objects are efficiently causal.

distinguish between right/left, up/down relative to us, but that there is also an up/down, left/right, front/back in nature.

> The mathematicals make this clear (that there is a distinction between natural up/down, etc. and relative), since, although they are not in place, nevertheless they have in position right and left relative to us as things merely spoken of due to position, without having each of these by nature.
>
> δηλοῖ δὲ καὶ τὰ μαθηματικά· οὐκ ὄντα γὰρ ἐν τόπῳ ὅμως κατὰ τὴν θέσιν τὴν πρὸς ἡμᾶς ἔχει δεξιὰ καὶ ἀριστερὰ ὡς τὰ μόνον λεγόμενα διὰ θέσιν, οὐκ ἔχοντα φύσει τούτων ἕκαστον.

So, on the reasonable assumption that Aristotle would endorse this endoxon (he might not), mathematicals do have relative but not natural up and down. Presumably, he is referring to the orientation of the diagram to its user, keeping in mind that this issue is distinct from the use of the prefixes 'up- (ἀνα-)' and 'down- (κατα-)' in the verbs discussed in § 1, and where the diagram, whether in sand or on papyrus, will be flat.

We can block the inference then by one of two Aristotelian claims:

1. The sense of 'contact' is clearly different for mathematical magnitudes and physical magnitudes, so that the sense of 'position' and 'place' must also be different.
2. The sense of 'up' and 'down' is clearly different in mathematics and in physics. In mathematics, it is merely relative to the observer of the diagram (or the adjacent text); in physics it is absolute.

To say that natural up and down are the first differentia of place does not entail that there might be a way in which something is in place where neither differentia applies, namely place that is neither up nor down.

In favor of (1), there is a clear difference between physical magnitudes being in contact, where their limits are different, i.e., the end of the pen in contact with the paper is not the same as the point on the paper; and mathematical ends, where it makes no sense to distinguish the end of a line and the point on a plane that it touches. The line is continuous with the plane, or, as Aristotle might say, the point on the plane and the point on the line are one in number but two in being or definition.[43] Putting emphasis on the word 'similarly', one might understand, "To the same extent, mathematicals are in position and in place." Against, Aristotle does not imply that mathematicals are defectively in position, and hence in place. The argument in favor of (1) is basically a rational reconstruction, which could easily be wrong.

---

[43] Cf. *Physics* Δ 13.222a10–17 (cf. 11.220a11–13), where a point is both beginning of one line and end of another.

In favor of (2), ‘place’ is a word that can appear in Greek mathematics. So when Aristotle says that up and down are the first differentia of ‘place’ he might not mean that everything in place is so only in a derivative way. This is particularly so in the case of the heaven and the sublunary world. What is up in the case of the heaven is neither heavy nor light and is effective and affective only in respect of locomotion. So Aristotle’s argument is restricted to natural sublunary bodies.

This may answer the question how general Aristotle’s argument is, but it does not answer our question how he can allow geometrical figures to be in place. For according our standard modern interpretations of Aristotle on mathematical objects, the geometrical objects just are physical objects with physical properties removed from their logical structure.[44] The physical universe seems to be one of the things substracted, and with it place. This reasoning may be too quick and easy. It is perfectly possible that Aristotle thinks of the geometrical universe as the entire physical magnitude of the universe qua magnitude. In this way, mathematical solids might have place. However, I think there is no evidence for this view in Aristotle, although we shall return to the suggestion later.

On the other hand, if we wish to distinguish place and position in Aristotle, we might think that whereas position concerns the parts of a single figure, place concerns the relations between distinct geometrical objects, best represented by *Elements* I 2 (to position a line equal to a given line at a given point). This too is a rational reconstruction, which could easily be wrong. So we have an interpretative puzzle that we can dance around with Aristotelian solutions, but it is not at all clear whether any of them capture Aristotle’s view. At least, it is reasonable to say that in so far as there is a sense of place that does not presuppose motion and rest, Aristotle does not seem to object to mathematical objects as he understands them (as opposed to the views of his Academic opponents) having place, but the question remains what that sense of place is. In any case, his view seems to have little to do with mathematical practice or to affect mathematical terminology.

3 Basic Spatial Notions in the Hellenistic Age

Eudemus and Theophrastus are probably the last people in the Hellenistic age to have endorsed an Aristotelian view of place and his rejection of the void. Starting with Epicurus, the dominant view would be something like the following:[45]

τόπος	place, extension occupied by a body (in Heron, also an extension being emptied)
κενόν	void or vacuum, extension that a body does not occupy or extension that a body does not occupy but that a body could occupy

[44] This is a common point of most modern accounts of Aristotle on mathematical object. Cf. Mendell (2004).

[45] Cf. Sedley (1982), Long and Sedley (1987, pp. 294–7), and Algra (1995).

χώρα	room, extension that a body is moving into or that is partially occupied (as such, it may blend into the next use)
[ἀναφὴς φύσις, or unamed]	intangible nature, the disjunction of the three (Epicureans) or χώρα—room that is indifferently occupied or not (possibly some Stoics[46])

None of the first three terms quite picks out a notion of space, although, as terminology slackens, any of them can. The intangible nature or room as indifferently occupied or not is space.

An issue for the Stoics is whether void is infinite, as Chrysippus and probably most Stoics, or does it just extend to where the universe expands when it burns up, as Posidonius. I take the varying testimonia as evidence of alternative views.[47] Additionally, I should mention the Peripatetic view of Strato of Lampsacus.[48] He too thinks of place as extension where an object is, but has a finite Aristotelian kosmos, with no void outside. However, within the kosmos are small interstices of void. This view influences Heron in his *Pneumatics*.[49]

Now, since our concern is with geometry, we can well ask whether any of this has anything to do with Greek geometry, and my suspicion is, 'not very much'. Although the issue arises in several Hellenistic naturalist contexts, it is hard to see Greek mathematicians thinking of lines and planes as demarcations of space. Of course, reasonable question arises about professional demarcations between philosophers and mathematicians. However, this question is misguided both for the ancient world and ours. The contrast in the Hellenistic Age should be between the study of nature and mathematics. This is, after all, the division that Posidonius attempts to make, between subjects that seek a cause and those that prove things from assumptions.[50] We can use this to mark professional identities, but I am not sure this is necessary. Some we would put down in both ranks, Ptolemy (cf. the

---

[46] The evidence for 'room (χώρα)' being for the Stoics partly occupied and partly unoccupied is primarily Sextus Empiricus, *Adv. math*. 10.3–4. (SVF 2.505). Since void and place are two of the four Stoic incorporeals, one might think that they would not form another sort that is neutral between the two. However, 'room' might well have done that work. Galen, *Qual. inc*. 19.464.10–4 (SVF 2.502) suggests that the Stoics were compelled to agree to extension in three dimensions (τὸ τριχῇ διαστατόν) as something common. Cf. Long and Sedley (1987, p. 296). De Harven (2012, p. 29) takes extension to be a mode of incorporeals and so not something distinct.

[47] For Chrysippus, cf. inter alia Arius Didymus, fr. 25 in *Dox. Graec*. 460.460.18–361.3 (SVF 2.503). For Stoics and Posidonius in contrast on this point, cf. Aetius in *Dox. Graec*. 338a16–19, b16–21. If Cleomedes is relying on Posidonius for the argument of *Caelestia* I 1, that void is unlimited, this report might be incorrect.

[48] Cf. Fr. 54–67 (Werhli) and Algra (1995, pp. 58–70).

[49] One might expect, as a result, some terminological care. Yet, although 'empty (κενόν)' does mean 'empty', Heron is, I suspect, casual about 'place (τόπος)' and 'room (χώρα)'. Cf. '(the air) goes into the emptying place (εἰς τὸν κενούμενον τόπον χωρεῖ)' at I 40.29. However, this needs more careful examination.

[50] From an epitome of Posidonius' *Meteorology* by Geminus and quoted by Alexander, whence Simplicius, *In Phys.* 291.20–292.31.

introduction to the *Tetrabiblos*), Heron; and some we would not, but possibly because we just do not know enough.

I am now also violating my principle of not trying to read ontology/physics into geometrical texts. So let me violate my principle a little more. Aristotle's friend, Eudemus represents no one but himself. Like his mentor, he rejects void, but, in building on Aristotle's argument (*Physics* Z 1) that two points cannot be in succession, he does consider void (fr. 99 = Simplicius, *In Phys*. 10.928.28–9.3):

> In the handbook, Eudemus made use of this [argument]: "For if," he says, "the partless are in succession, it is altogether necessary that there be something between them that is not of the same kind, so that it would not be a point, but a *line or void in a length between points*. And so, the line will not be [composed] from points, since successive points will not be in it. But if there is void, the void will be more in the [class of] continua than [in the class of] things from which it is [composed], that is, [more] than from the points said to be in succession, or there won't be any magnitude at all. For just as two points do not make a length, so too nor does a point and void."

> Ὁ δὲ Εὔδημος τῷ ἐπιχειρήματι οὕτως ἐχρήσατο "εἰ γάρ ἐστι, φησίν,
> ἐφεξῆς τὰ ἀμερῆ, δεῖ πάντως εἶναί τι αὐτῶν μεταξὺ μὴ ὁμογενές, ὥστε
> στιγμὴ μὲν οὐκ ἂν εἴη, γραμμὴ δὲ ἢ κενὸν μεταξὺ στιγμῶν ἐν μήκει. οὐκ
> ἔσται οὖν ἐκ τῶν στιγμῶν ἡ γραμμή· οὐ γὰρ ἐν αὐτῇ αἱ ἐφεξῆς στιγμαί·
> 929.1 εἰ δὲ κενόν, πλέον ἔσται τὸ κενὸν ἐν τοῖς συνεχέσι τῶν ἐξ ὧν, τουτέστι
> τῶν ἐφεξῆς λεγομένων στιγμῶν, ἢ οὐδὲ ἔσται μέγεθος ὅλως. ὥσπερ γὰρ
> ἁπτόμεναι δύο στιγμαὶ μῆκος οὐδὲν ποιοῦσιν, οὕτως οὐδὲ στιγμὴ καὶ κενόν."

Now, the quotation from Eudemus establishes that one could argue that there is no difficulty in thinking of distances between points of the void as lengths. Or rather one might think that lines could represent distances between void-points. However, Eudemus clearly makes a conceptual distinction between void-lengths and lines, where the lines are presumably physical lines. Someone else might make a different move, e.g., Epicureans.[51] So let us turn to the Stoics, who, after all, are the most likely ancient philosophers to have had an influence on Greek mathematicians.

4 Incompatibilism? Compatibalism? Revisionism? Descriptivism?

Certainly, there are many Stoics and Epicureans who were credited with great knowledge in mathematics. What the standard for great knowledge is is another matter. Nonetheless, it is possible that some Epicureans were up in the field,[52]

---

[51] Cf. Verde (2010, esp. pp. 256–61) on minima in time and space among Epicureans. However, even if the Epicureans regard void has having minima, etc., it would not show that they made the additional move of treating distances in the void as lines, as geometrical objects.

[52] Cf. most recently Verde (2010, pp. 213–256). Certainly, there were many Epicureans who are reported to have had some proficiency with geometry: Polienus of Lampsacus, Zeno of Sidon,

maybe Philonides, IF he is the Philonides whom Apollonius commends in the introductory letter to Eudemus that begins *Conics* II.[53] Much less compelling is the suggestion that the skillful correspondent of Hypsicles (the presumed the author of *Elements* XIV) is also the Epicurean Protarchus.[54] In any case, no extant Epicurean text shows a mathematical concern beyond the contents of *Elements* I. Epicureans interested in mathematics, at least according to our sources, were primarily concerned with minima and division and are unabashed revisionists, who seem to have had little influence on extant mathematical practice. Unrestricted division is taken for granted in every standard mathematical text, and without comment. The principle that given two unequal comparable magnitudes one may continuously slice the larger so that the remainder is smaller than the other, the basis of the so-called Eudoxan method of exhaustion, implies that there is no smallest magnitude. Of course, the principle is different from one that says that a magnitude may be infinitely or completely divided, as, no doubt, Anaxagoras had already seen in the 5th cent. (cf. fr. B3). Indeed, the distinction may even be the primary point of the Eudoxan method. The principle is not an isolatable feature of the work of Eudoxus, Euclid, Archimedes, and Apollonius, any more than is the assumption that any line

(Footnote 52 continued)
Demetrius of Laconia, Philonides of Laodicea by the sea. It is another question how proficient they were.

[53] The argument goes back to Crönert (1900). Apollonius recommends that his correspondent, Eudemus in Pergamum, make known Book II to Philonides, whom he had introduced to Eudemus in Ephesus, should he show up in town. PHerc. 1044, a life of Philonides, praises his geometrical skills and says that his first teacher was Eudemus and his second Dionysodorus son of Dionysodorus of Caunus. There is at least one known, post-Archimedean mathematician named Dionysodorus, so that the coincidence of names might seem somewhat compelling, even though Apollonius never suggests that Philonides was a student of Eudemus and indeed may suggest that he was already a geometer when the introduction took place. Nor is there any evidence that the two Dionysodori are the same person. Furthermore, we do not know why Apollonius mentions Philonides; if he is the Epicurean, maybe because he is well connected at the Seleucid court. Hence, the evidence gets thinner the more it is examined. In any case, none of this suggests that Apollonius or any working mathematician, which the Epicurean Philonides seems not to have been, had Epicurean connections, even less that they were influenced. It would at best attest to geometrical skill with one member of the Garden, who, like so many ex-scientists, used his knowledge in a very limited way.

[54] Cf. Crönert (1900) and Verde (2010, pp. 233–4). Hypsicles tells a story about Basileides of Tyre and his father in Alexandria and their common interest in a treatise of Apollonius to explain how he got interested in the topic of dodecahedra and icosohedra, and mentions the friendship of Protarchus and his father, as well as his geometrical prowess. Nowhere does he suggest that Basileides and Protarchus are familiars. The coincidence here is that a Basileides was the fourth head of the Epicurean Garden, and a Protarchus of Bargalia was the teacher of Demetrius of Laconia, whom the Epicureans considered skilled in mathematics. Basileides has no known connection to mathematics, while Protarchus' only connection is that he taught Demetrius, although we do not know what he taught. It is now standard to cite the Epicurean as 'Basilides of Tyre'. I would suggest more caution. Not everyone with the same name is the same person or even from the same town. Netz has informed me of even more compelling reasons to doubt the identifications of the two Philonidai and the two Protarchi.

may be divided at any point. No ancient mathematician asks with which segment the point of a line division goes.

Nor is it typical to think of lines as composed of points. However, Archimedes' *Method* famously treats n-dimensional objects as composed of $n - 1$ dimensional objects, figures of lines, solids of planes. Props. 1–13 also treat each of these as having weight. It has been pointed out that nowhere does Archimedes actually say that the multitude of composing lines is finite or infinite. So "one cannot rule out the possibility that he thought of them as indivisible minimal bodies."[55] I think, on the contrary, that one can rule out the issue as playing any role in the argument, and it is very easy to see why. In 1–13 he uses dimensional-reduction (commonly mis-called 'indivisibiles') and the balance, while in 14 he just uses dimensional-reduction, but in 15 he gives a Eudoxus/Archimedes proof of the previous theorem, which will require unlimited division. So, he has not abandoned the condition of unrestricted divisibility in the treatise and so has not endorsed surreptitiously an Epicurean doctrine as the ideology of the treatise.

Archimedes identifies Props. 1–13 as heuristic because, like the argument of Q*uadrature of the Parabola* 6–17, they use mechanics. Whether he also regards 14 to be heuristic may never be known.[56] In the introduction, Archimedes cites *Conoids and Spheroids* 1,[57] which he will use explicitly in Proposition 14. Anachronistically, if we have two series of magnitudes equal in multitude (τῷ πλήθει ἴσα), a_1, …, and b_1, …, and pairwise for any i and j, $a_i : a_j = b_i : b_j$, and another two series c_1, … and d_1, …, where a_i: $c_i = b_i$: d_i, then Totality(a_i, …) : Totality(c_i, …) = Totality(b_i, …) : Totality(d_i, …). The proof of the lemma in the earlier treatise only deals with the finite case. Indeed, it is hard to see how Archimedes could prove it for the infinite case. Yet all he says in the application is that each of the four groups of magnitudes are 'equal in multitude (ἴσα τὸ πλῆθος)' to each other, a formula one readily recognizes from Euclid and Archimedes as a variant on the other formula. Yet, 'equal in multitude', here without anachronism, in effect means that the groups can be paired up (dare I say, "1:1"). So in a sense, the suggestion is right that Archimedes does not mention either that he is making a leap from the finite case to the infinite nor that he is treating the figure as composed of finite figures. Since, in the proof, all the a's are

---

55 Cambiano (2008, p. 588).

56 That said, it is very plausible that Archimedes also regards *Method* 14 as a heuristic. My point is just that we cannot know this. See most recently Christianidis and Demis (2010), who argue for this plausible thesis, but, much more controversially, that Archimedes believes that one can have mechanical demonstrations of geometrical propositions, so that the heuristic feature of the *Method* is the method of indivisibles. The difficulty is that Archimedes only contrasts: his heuristic method, which he does not describe; theorems investigated through mechanics, which is the topic of the first part of the book; and geometrical proofs which will come at the end. Everything beyond is our guesswork, except that Archimedes makes the same point in the introduction to the *Quadrature of the Parabola*, that the quadrature was discovered through mechanics and that he is sending to Dositheus how through mechanics it was observed and how through geometry it was demonstrated.

57 Whether the actual mention of the book is an interpolation is unimportant, but see Netz et al. (2001, p. 20).

equal (triangles) and all the b's are equal (lines), at least we do not have to worry about the ratios of the adjacent triangles and lines. We know that without a lot of fairly sophisticated mathematics the proof won't work either way. Without restrictions, we know the method will be vitiated by Cavalieri paradoxes. Yet, it works.[58] The theorem also requires that we have the ratio of the diameter of a parabolic segment to a parallel to the diameter and that of a triangle in a prism to a triangle in a cut cylinder. Will we get this if there are minima? There is, however, another paradox, much more ancient, that calls into question the method. Nevertheless, if Archimedes thinks of it merely as an effective heuristic, who's to damn it, either then or in the 17th century?

It is some time since Luria[59] tried to connect Archimedes' *Method* to atomism. So far the only certain connection remains Archimedes' correction of his earlier praise of Eudoxus in *De sphaer. et cyl.*, namely that Democritus first stated the theorem of the ratio of cone to cylinder, and that could have come from anywhere, but most likely from some learned Alexandrian, such as Eratosthenes. If so, it contributes little to our understanding of Archimedes. Whatever Archimedes may have thought, it is clear that either he just assumes that we will read his method one way and not the other, or he doesn't care since the issue of how many lines compose a figure doesn't enter into the argument. It would be just like him to tease the reader with something hanging, but how would we know? After all, he was that kind of guy! Nonetheless, the *Method* presumes propositions whose truth and proofs are incompatible with atomism.

The Stoics seem better candidates for having an influence on mathematical treatments. They can even claim a notable or two, in particular Eratosthenes, who studied with Ariston and even met Zeno once.[60] Well, having studied with Stoics doesn't make one a Stoic any more than having studied geometry makes one a geometer.[61] One would have to find evidence in Eratosthenes of Stoicism at work. There is some evidence that mathematicians at least shared some vocabulary with the Stoa.[62] I shall consider here four issues in the evidence for Stoic views:

[58] These issues are fully explored in Netz et al. (2001). They rightly do not consider the possibility that the method is finitist.

[59] Cf. Luria (1933).

[60] Cf. Strabo, I 2.2, Athenaeus, *Deip*. 7.14–20 (SVF 1.341). Strabo suggests more familiarity (γνώριμος) between the pre-teen and the septuagenarian, Zeno, which we can put down to the ordinary exaggeration that accrues to such stories over time.

[61] ˜I fail see any significance in Chrysippus knowing that great circles bisect one another, the most basic theorem in spherical geometry (a few generations later, it will become Theodosius, *De sphaera* I 1), assuming that the evidence in Cicero, *De fato* 15 does indicate this (so Mansfeld 1983, p. 66). Whether Chrysippus accepts or rejects geometry, he can know this as a theorem. This is not to impugn Chrysippus' knowledge of mathematics, which may well have been extensive, just the way scholars embroider evidence.

[62] 'Property (σύμπτωμα)' is the obvious example, although it already has a logical meaning in Aristotle's *Organon, Topics*. Δ 6.126b35–7a2, less clearly *Cat*. 8.9b19–10a10. However, it is not a central term in Aristotle, nor, for that matter, anywhere else in 4th century BCE discourse. This says nothing about the direction of influence.

treatments of the basic definitions in geometry and whether some Stoics, especially Posidonius, were committed to a conceptualism or to a modified physicalism; Chrysippus on locus theorems; Chrysippus on infinite divisibility (including of void) and the cone paradox; Posidonius on the nature of figures. Whatever they may indicate about Stoic views on Greek mathematics, they will say very little about Greek mathematics itself. Regardless of whether some were Stoics, Greek mathematicians do not wear Stoic badges on their tunics.

This is not the place to sort out the woefully inadequate evidence on the status of mathematical objects in the Stoa and the degree to which views may have differed.[63] Given the Stoic distinction between existing (ὑπάρχον), the status of corporeals, and subsisting (ὑφεστηκός), the status of incorporeals, the standard list being place, void, time, sayables,[64] where do planes, lines, and points go? In addition to the standard list of incorporeals, Cleomedes and probably Posidonius, took them to include the limits of bodies,[65] while Plutarch explicitly takes the Stoics as treating points and planes as incorporeals.[66] Hence, solid figures (σχήματα) are qualia and hence corporeal,[67] while it would follow that lower dimensional figures are incorporeals. Unless the Stoics generally are physicalists, this will not tell us how the objects of mathematics are to be treated. That is, are they incorporeal subsistents dependent on bodies? Or are they conceptual and distinct in some sense from incorporeal limits?

Because Geminus (early 1st. cent. CE) preserved much of Posidonius and was available to Proclus, when combined with a few other fragments, we can get a slightly clearer picture at least of his views on the status of mathematical objects. Diogenes Laertius (*Vitae* VII 134) says that most of the major Stoics (Zeno, Cleanthes, Chrysippus, Archedemus, and Posidonius) distinguish between principles that are eternal, incorporeal, and lacking shape and the elements that have shape but get destroyed in the conflagration of the world.

Diogenes goes on to say (VII 135):

---

[63] Cf. especially Long and Sedley (1987, pp. 297–303), Robertson (2004), Ju (2009), and de Harven (2012).

[64] Cf. Long and Sedley (1987, pp. 162, 164–5).

[65] Cf. Ju (2009, pp. 381 and 385–6), citing Cleomedes, *Caelestia* I 1.139–44 and 3.34–5. I take it that her argument (pp. 381–6) effectively refutes the view that they are corporeals for any Stoic. This is not to say that there might not be some serious concerns, as Paparazzo (2005) raises, but ultimately the issue here concerns Pliny's treatment of patinas, and not Posidonius. Nonetheless, it is of some concern that the evidence for Posidonius, Diogenes Laert., *Vitae* VII 135 (discussed below), comes from book 5 of *De meteora*. One really wants to know the context.

[66] *De comm. not*. 1081B5 for plane, B11–2 for point, 1080EF for the contacts in general of a body with an incorporeal.

[67] So Simplicius, *In cat.* 271.20–22 (SVF 2.383), and compare with 217.32–218.1 (SVF 2.389), that if what's qualified is body the quale is body, and if bodiless bodiless. Cf. Long and Sedley (1987, pp. 169, 172). The context of Simplicius in the first passage is why Aristotle puts figure with qualities. So the Stoics agree with Aristotle in making them qualities but not in making them bodies, where Aristotle classifies bodies as quantities. It is enough for Simplicius' argument that some figures are bodies, namely bodies so qualified.

> A body, as Apollodorus says in the *Physics*, is what's extended in three ways, in length, in width, in depth. This is also called 'solid body'. A surface is a limit of a body or what has length and width alone and not depth. Posidonius in the fifth [book] on *Things Seen Above* [meteorology] admits this both in attentive-thought and in substance. And line is a limit of a surface or widthless length or what has length alone. A point (στιγμή) is a limit of a line or what is a smallest mark (σημεῖον).[68]

7.135.1 Σῶμα δ' ἐστίν, ὥς φησιν Ἀπολλόδωρος ἐν τῇ Φυσικῇ, τὸ
τριχῇ διαστατόν, εἰς μῆκος, εἰς πλάτος, εἰς βάθος· τοῦτο δὲ καὶ
στερεὸν σῶμα καλεῖται. ἐπιφάνεια δ' ἐστὶ σώματος πέρας ἢ τὸ
μῆκος καὶ πλάτος μόνον ἔχον βάθος δ' οὔ· ταύτην δὲ Ποσειδώνιος
7.135.5 ἐν πέμπτῳ Περὶ μετεώρων καὶ κατ' ἐπίνοιαν καὶ καθ' ὑπόστασιν
ἀπολείπει. γραμμὴ δ' ἐστὶν ἐπιφανείας πέρας ἢ μῆκος ἀπλατὲς
ἢ τὸ μῆκος μόνον ἔχον. στιγμὴ δ' ἐστὶ γραμμῆς πέρας, ἥτις ἐστὶ
σημεῖον ἐλάχιστον.

First, the definitions would all seem to be from Apollodorus, but endorsed by Posidonius. What Posidonius adds is that surface is both in thought and in substance. Apollodorus' definition of body as three dimensional is notable in two ways. First, many Stoics defined body by affective and effective capacities. So we might suspect a more geometrical definition is given in order to define a geometrical body, where the definiendum would be solid body (στερεὸν σῶμα), perhaps as opposed to natural body. It is not clear why Diogenes proceeds from talking about eternal principles to specific principles, the definitions of basic mathematical objects. As to the other definitions, the only thing notable is that Stoics should prefer the definitions that provide an $n-1$-dimensional object as the limit of an n-dimensional body. For them, as for Aristotle, this is the ontologically correct version. Yet, we have a bevy of triples of definitions, but mostly unremarkable, as is clear from a comparison with Euclid, *Elements* I Definitions 1–3, 5, 6 (or any perusal of definitions in Aristotle). So the order in Diogenes is from three to no-dimensions and with each (except for the point): the ontologically correct definition, the privative definition, the definition by a list of dimensions. Euclid gives a selection of these in reverse order:

point	1.	different definition	2.	ontologically correct
line	3.	privative definition	4.	ontologically correct
surface	5.	list of dimensions	6.	ontologically correct

There is no definition of solid in Euclid, but Posidonius' is the sort of definition we would expect.[69] So the only thing of note here is that Diogenes gives both the privative and list forms, which Euclid seems not to distinguish. The definition of

[68] Taking στιγμή as the definiendum is a little strange. This common word for 'point' in Aristotle is never used in an any mathematical text between Eudemus and Theon of Smyrna, while σημεῖον is the common word.

[69] Cf. the citing of Euclid's definitions in Philo, *De congressu eruditionis gratia* 147, ending with "what has three dimensions, length, width, depth (ὃ τὰς τρεῖς ἔχει διαστάσεις, μῆκος, πλάτος, βάθος)."

'point' is very strange, but may be an attempt at a substitute for Euclid's definition, "of which there is no part (οὗ οὐ μέρος οὐθέν)," to meet some atomist objection. If the order is significant, it is a reasonable guess that Apollodorus and Posidonius want the ontologically correct version to come first.

This is confirmed by how Posidonius treats the surface, as according to attentive-thought[70] and subsistence, where attentive-thought involves, presumably forming an object of thought where none exists. A similar claim occurs in Proclus, where limits subsist according to attentive-thought.[71] When he finishes his argument that limits can be efficacious, he refers back to this claim and seems to distinguish two Stoic views, one that involves conceptually separating limits from bodies (a physicalist abstractionism) and the other treating them as purely conceptual (a fictionalism).[72] In either case, they would be somethings in a presentational/ representational faculty (φαντασία) that is distinct from the surfaces, etc. that do subsist as incorporeals. If so, is this because they are concerned with the idealization problem, that physical objects are never so round as the geometrical ones?

(Footnote 69 continued)
This definition somewhat appears at Aristotle, *Physics* Δ 1.209a4–6, "And so, [place] has three distances, length, width, depth, by which every body is defined (διαστήματα μὲν οὖν ἔχει τρία, μῆκος καὶ πλάτος καὶ βάθος, οἷς ὁρίζεται σῶμα πᾶν)." The third class of definition is not in Aristotle, but cf. *Met.* Δ 13.1020a11–13, where each kind is given by the number of dimension and the last dimension as the definiendum, e.g., depth is what's continuous in three (τὸ δ' ἐπὶ τρίαβάθος, with συνεχές understood from earlier).

[70] Chrysippus treats ἐπίνοια as the faculty that is able to conceive things that are logically possible, e.g., to divide mentally fire into two bodies. Cf. Galen, *In Hippocr. de nat. hom.* p. 30 (SVF 2.409).

[71] Proclus, *In Eucl. El.* 89.15–7 (SVF 2.488, part), "One should not believe that such limits, I mean 'of bodies', subsist according to a fine attentive-thought, just as the Stoics suppose (ὅτι δὲ οὐ δεῖ νομίζειν κατ' ἐπίνοιαν ψιλὴν ὑφεστάναι τὰ τοιαῦτα πέρατα, λέγω τῶν σωμάτων, ὥσπερ οἱ ἀπὸ τῆς Στοᾶς ὑπέλαβον)."

[72] The full context is important (*In Eucl. El.* 91.19–24), where, in refutation of the Stoic view, Proclus has just argued for the causal efficacy of poles and axes, "In observing things imperfectly subsisting within what are themselves limited the many believe their subsistence to be obscure, that is, some say that these are separated according to attentive-thought alone from perceptibles, while others, I suppose, that they have no being other than in our attentive-thought (οἱ δὲ πολλοὶ τὰ ἐν τοῖς περατουμένοις αὐτοῖς ἀτελῶς ὑφεστηκότα θεωροῦντες ἀμυδρὰν αὐτῶν οἴονται τὴν ὑπόστασιν εἶναι καὶ οἱ μὲν κατ' ἐπίνοιαν μόνην χωρίζεσθαί φασιν αὐτὰ τῶν αἰσθητῶν, οἱ δὲ μηδὲ ἀλλαχοῦ που τὴν οὐσίαν ἔχειν ἢ ἐν ταῖς ἡμετέραις ἐπινοίαις.)." Given the switch from 'subsistence' to 'being', one might think that the Stoics are the first group, but the switch need not be significant given the use of 'subsistence' in the main clause. Cf. Plutarch, *De animae procreatione in Timaeo* 1023B7 on Posidonius, "the existence of the limits about the body (τὴν τῶν περάτων οὐσίαν περὶ τὰ σώματα)." Instead, I think 'the many' just is 'most Stoics'. They are the only group in question here, if Proclus meant most people, even most people who think about these things, the comment would be strange indeed, as most people in his time are his people. And so, we have two Stoic views, one abstractionist and the other fictionalist (putting geometrical objects with centaurs and the like, as neither corporeal nor incorporeal). Since, on either view, there can still be two sorts of limits, as Robertson (2004) and De Harven (2012).

Well, there is paltry evidence that any Hellenistic Stoic ruminated on this.[73] It is, however, at least a nice story that the common interpretation of Aristotle, that the mind contemplates mathematical objects that are derived from sensation but are presented in the imagination, ultimately derives from the Stoa.[74]

The two positions might also reveal different attitudes towards spatial relations and what is acceptable in mathematics. The fictionalist may have more freedom in allowing conflicts with the structure of the physical world, e.g., the decomposition of figures and the extension of lines. It is to these that we now turn. I shall discuss in this section some mathematical issues that arise for local-theorems, Eudoxan division of a figure, Archimedean dimensional reduction, the nature of a figure. In the next section, I shall also look at a small issue on infinite extension in the Stoa.

In his discussion of local-theorems, discussed in § 1, Proclus (*In prim. Eucl. el.* 395.13–21) says that Geminus reports that Chrysippus compares local-theorems to the Ideas (presumably Stoic conceptions[75] and not Platonic Forms). For the simile to be meaningful, it must concern local-theorems and not theorems in general. Just as the Ideas encompass the generation of unlimited entities within determinate limits (ἐν πέρασιν ὡρισμένοις), so the theorems encompass unlimited entities within determinate places (ἐν ὡρισμένοις τόποις). To know more, we have to know more about Chrysippus on universals. On a strong nominalist reading of the remark, it might mean that just as the conception determines the limits within which each thing is so grouped, the locus theorem determines limits, i.e. the locus (τόπος), within which certain geometrical individuals are equal. If Chrysippus had said that ideas are like general theorems, he would merely have indicated that general notions are like general notions in mathematics, a not very profound remark. So a general theorem can easily be nominalized to a statement about individuals ("Every triangle …," to "given a triangle, ABC, …"), just as can any other general claim. By making the simile with locus theorems, he points to the spatial boundary that the locus theorem sets. So the nominalization, "given parallelograms ABDE and ABGC between the parallel lines AB, ECDG," does not dispose of the locus between parallel lines, since that too is an individual as is the region between the parallel lines. If this interpretation is right, the simile says little about mathematics, but much about how Chrysippus wants us to understand general propositions and general notions. Nothing here is revisionist, as it involves merely treating a locus by its boundary. Yet it is important that a locus is not necessarily a figure, indeed, just as is required.

How then would Stoics broach the question raised in the previous section about whether the void is subject to mathematical treatment? Cleomedes, *De motu* 8.10–14, argues that void is incorporeal, intangible, and lacking shape and not being shaped, that it is merely capable of receiving body. Whether or not anyone

[73] Basically, the only evidence is Proclus' discussion and claim that they are imperfectly subsisting (previous note), but this is Proclus' issue. How do we know whether it also bothered the Stoa?

[74] Cf. Mueller (1990), who traces the view to Alexander.

[75] Cf. Long and Sedley (1987, pp. 179–183).

ever reasoned in this way, someone holding this view should probably hold that void is neither divisible nor indivisible and that its infinity outside the cosmos is to be explained counter-factually, by the conceptual possibility of the world expanding infinitely. On the other hand, according to Stobaeus, Chrysippus held that void is infinitely divisible.[76] Does this imply that it is possible to do a geometry of the extra-cosmic void, at least on the abstractionist view?

Yet this testimonium itself is mathematically curious. Stobaeus goes on to say that given a body divided up ad infinitum, a body is not composed of infinite bodies, nor a surface, nor a line nor place. We may infer that he would also apply the principle to the other divisibles that have just been mentioned, time and void. But when would one think of a surface as composed of infinite surfaces (as opposed to, say, lines in the Archimedes' *Method*). If Chrysippus is as knowledgeable about mathematics as his fans believe, then the most natural place would be in a continuation of a typical Eudoxan decomposition, e.g., of a circle into successive inscribed approximating figures. Yet the claim is suspiciously like Aristotle's account of potential infinity. In any case, with the availability of Eudoxan techniques for avoiding exhaustion, the issue is moot.

Nonetheless, Chrysippus pyramid paradox and his discussion of Democritus' cone paradox, as hostilely reported by Plutarch, *De comm. not*. 1078E-1080E, might pose a challenge to Stobaeus' report. The cone paradox of Democritus goes: Suppose we cut a cone into a top cone and frustum. Is the bottom surface of the top cone equal or unequal to the top surface of the frustum. If unequal, the cone will be jagged; if equal, then all such surfaces will be equal and the cone will have been a cylinder.[77] The pyramid paradox is less clear but may have been similar. It is tempting to see the paradox as related not just to Epicureanism but also to any technique of dimensional reduction. If Chrysippus objects to a body composed from an infinite decomposition of a body into bodies, and so forth, then a fortiori he would object to a body composed from an infinite decomposition of a body into planes. Given Stobaeus, the second horn of the dilemma will have to involve decomposing the cone into planes, inferring by transitivity of identity that they are all equal, and then inferring that the original cone was somehow not composed of the planes but that the equality still holds. This is not what Plutarch reports.

Now, if we want to see a dialogue between the philosopher and the mathematician, Archimedes' *Method* may well reflect an older mathematical method that would prompt Chrysippus' looking at the paradox.[78] Alternatively, it is very

---

[76] Stobaeus, *Eclogae* I 14.1e (SVF 2.482). "Chrysippus said that bodies are cut ad infinitum as well as things that are like bodies, e.g., plane, line, place, void, time. And when they are cut ad infinitum, neither are bodies composed from infinite bodies, nor a plane nor a line nor a place nor <a void nor a time> . (Χρύσιππος ἔφασκε τὰ σώματα εἰς ἄπειρον τέμνεσθαι καὶ τὰ τοῖς σώμασι προσεοικότα, οἷον ἐπιφάνειαν, γραμμήν, τόπον, κενόν, χρόνον· εἰς ἄπειρόν τε τούτων τεμνομένων οὔτε σῶμα ἐξ ἀπείρων σωμάτων συνέστηκεν οὔτ' ἐπιφάνεια οὔτε γραμμὴ οὔτε τόπος < οὔτε κενὸν οὔτε χρόνος > .)".

[77] Cf. Hahm (1972), Robertson (2004), inter alios, for a sample of many rival interpretations. My goal here is not to given an interpretation of the text.

[78] Cf. Knorr (1996).

tempting to suppose that just as Hipparchus was to do half a century later,[79] Archimedes is tweaking his nose at his contemporary Chrysippus.[80] Alternatively, Chrysippus might be attacking the *Method*. Of course, we do not know if either is aware of the other.

The problem is not merely that any story we come up with here is at best a nice historical fiction. Other than that it is not the one just suggested, it is not even clear what Chrysippus' solution was and what it says about his views on mathematics. His solution contains four claims:

1. the surfaces are neither equal nor unequal.
2. the bodies are unequal, due to (1).
3. there is a sense in which something can be larger (μεῖζον) without exceeding (ὑπερέχον).
4. bodies are in contact at a limit and not at a part (i.e., we might note, in agreement with Aristotle)

With skill, one can find in these non-revisionist readings of these three claims, and one of these readings might be right.[81] Whatever non-revisionist readings one takes, it must allow one to say that when one runs a plane through two equal cones parallel to the base, the volumes are determined by the diameters of the surfaces and the heights, so that the bottom of the top cone equals the top surface of the frustum. Otherwise, the mathematics will become bizarre, an orphan from a different millennium.[82]

Let us turn now to the fourth issue. In commenting on Euclid's definition of figure (El. I 14):

> A figure is what's contained by some boundary or boundaries (Σχῆμά ἐστι τὸ ὑπό τινος ἤ τινων ὅρων περιεχόμενον.)

Proclus (*In Eucl*. 143.5–21) holds that there is a conceptual difference between the way Euclid and Posidonius conceive 'figure (σχῆμα)':

> And so, in calling what's enfigured 'figure', he reasonably named in addition what's enmattered and existent in quantity 'contained'. But, separating the account of 'figure' from quantity and positing that (the definiens) is a cause of its being determined and limited and of the containing, Posidonius defines 'figure' as an enclosing limit. For the closing is different from what's enclosed and the limit from what's limited. And one (Posidonius) seems perhaps to look towards the boundary that surrounds from without, while the other (Euclid) to a whole, the substrate, so that one will say that a circle is wholly the plane figure

[79] Cf. Plutarch, *De repug. Stoic*. 1047DE and Acerbi (2003).

[80] Chrysippus outlived Archimedes, but the *Method* was probably fairly late in his career.

[81] It may even be that Chrysippus takes the paradox to be about physical bodies and not mathematical objects, Long and Sedley (1987, p. 302). However, the solution is then easily refuted. One constructs a cone equal to the top cone and a frustum equal to the bottom frustum. By transitivity of identity the bottom surface of the cone is equal to the top surface of the frustum. Now repeat ad infinitum.

[82] This is at least how I understand several proposals such as that of White (1992, pp. 284–313).

and the outside containing, while the other according to the circumference. And one indicates that the enfigured as contemplated with the substrate gets defined, while the other that he wants the account itself of ‘figure’ to make clear what limits and encloses the quantity.

ὁ μὲν οὖν Εὐκλείδης τὸ ἐσχηματισμένον σχῆμα καλῶν καὶ τὸ ἔνυλον καὶ τῷ ποσῷ συνυπάρχον περιεχόμενον εἰκότως αὐτὸ προσείρηκεν, ὁ δὲ Ποσειδώνιος πέρας συγκλεῖον ἀφορίζεται τὸ σχῆμα τὸν λόγον τοῦ σχήματος χωρίζων τῆς ποσότητος καὶ αἴτιον αὐτὸν εἶναι τιθέμενος τοῦ ὡρίσθαι καὶ πεπεράσθαι καὶ τῆς περιοχῆς. τὸ γὰρ κλεῖον ἕτερόν ἐστι τοῦ συγκλειομένου καὶ τὸ πέρας τοῦ πεπερασμένου, καὶ δοκεῖ πως ὁ μὲν εἰς τὸν ἔξωθεν περικείμενον ὅρον ἀποβλέπειν, ὁ δὲ εἰς ὅλον τὸ ὑποκείμενον, ὥστε τὸν κύκλον ὁ μὲν ἐρεῖ καθʹ ὅλον τὸ ἐπίπεδον εἶναι σχῆμα καὶ τὴν ἔξω περιοχήν, ὁ δὲ κατὰ τὴν περιφέρειαν. ἐνδείκνυται δὲ ὁ μὲν ὅτι τὸ ἐσχηματισμένον ἀφορίζεται καὶ σὺν τῷ ὑποκειμένῳ θεωρούμενον, ὁ δὲ ὅτι τὸν λόγον τοῦ σχήματος αὐτὸν τὸν περατοῦντα καὶ συγκλείοντα τὸ ποσὸν ἐμφανίζειν ἐθέλει.

The issue relates to a perplexing question for a philosophy of geometry: how does drawing the perimeter of a figure ‘create’ the matter in between (see § 1). As a result one might think that geometry is either a study of demarcated spatial extension or demarcations of a material plenum. Whether Greek geometers are committed to anything beyond the method of constructing a figure by constructing its perimeter is an issue that I have eschewed in my discussion, in as much as there is no evidence from Greek mathematical texts. Posidonius seems to hold the view that describing a figure is just drawing the perimeter. Without the need to add Proclus’ ontology, we have seen that Euclid does treat ‘circle (κύκλος)’ as disk, albeit with some difficulties (§ 1), while Proclus is not claiming that Posidonius treated plane figures merely as perimeters. He does say that the enclosing limit is the cause of the figure being determined and contained. One might well smell a little Aristotelianism here (a geometrical figure as its form or shape enclosing its matter or extension), while worrying about the sense of ‘cause’. Given that such figures are incorporeals anyway, if not mere conceptions, it does not seem to affect how one is to understand geometrical objects or their relations. However, we also see here that the definition of ‘figure’, the quale, of an extension, reverses the ontological ordering we saw earlier in Posidonius. To know better, we would have to have much more evidence, including a sense of how Posidonius defined figures such as circles.[83]

[83] The other principal evidence for Posidonius’ view of geometrical objects is Plutarch, *De an. procr. in Timaeo* 1023BD.

Has the naturalist and metaphysical discourse of the Hellenistic Age that relates to volume and spatial relations been at all reflected in mathematical discourse? Where Epicureans or Stoics were odd or non-standard, they seem to be ignored, or at best teased. Where they were standard, the standards would seem already to have been set by the time of Euclid. Where mathematicians went non-standard, namely Archimedes' *Method*, we can discover many *ben trovata*, but the ones that look to Epicureans don't look very well. Where mathematics becomes more physical, e.g., in the Heronian tradition, the influence will be greater. Where the content is more natural, standard views about nature, such as the sphericity of the (liquid) world in Archimedes' *Floating Bodies* I, will be part of the mathematics. To conclude this part of my discussion, I have found little reason to see any particular view about place or space in mathematical texts, except perhaps in the method of measuring area of a solid by measuring vacuums.

5 Infinite Extensions and Other Infinite Activities

The previous discussions were warm up. If all philosophers except for Epicureans and Skeptics thought that the body of the universe is finite, then what do we do with long lines, large planes, and big spheres. These are all a part of mathematics. In this, it is normal to begin with a notorious claim of Aristotle. He has just argued that in the case of physical bodies the actual infinite in extension does not exist except for time, which does not exist 'together'. The sense in which the infinite exists is that some divisibles may be always divided more and that conversely, an ever smaller amount may be added to a magnitude, so long as it does not exceed every given magnitude. For, the world is finite. It is important to realize that in the entire corpus of Aristotle's work there appears only one argument against actual infinities in mathematics (*Phys*. Γ 5.204b4–10), a 'logical' argument that bodies, by definition, are bounded. However, *Met*. Δ 13.1020a11–14 completely undercuts this argument by defining 'body' as finite magnitude in three dimensions, whence infinite three-dimensional magnitudes could be a distinct species. So, without any argument that infinite extension is impossible in mathematics, Aristotle then says (*Physics* Γ 7.207b27–34):

> The argument does not take contemplation from the mathematicans, although it does take away there being an infinite in such a way that it is in actuality, in increment, and untraversible. For they don't need the infinite, since they don't use it, but merely [need] there to be as much finite [line? increase?] as they want. It is possible for any other sized magnitude to be cut in the same ratio as the largest magnitude, so that the fact of its existence [any size?] in existent magnitudes will make no difference to them for [the purposes] of proving.[84]

[84] It is natural to take the feminine participle at b31 as referring to a line. However, the use of the neuter 'magnitude' in the next line undercuts this. The only feminine noun that could make sense is 'increase'. It is also unclear what exists in or among the existent magnitudes at b34. Whatever it is should to be unneeded in the proof. So Hussey (1983) and the Oxford translation take it that the (potential?) infinite exists in the existent magnitudes. Surely, what must exist in or among the

> οὐκ ἀφαιρεῖται δ' ὁ λόγος οὐδὲ τοὺς μαθηματικοὺς τὴν θεωρίαν, ἀναιρῶν οὕτως εἶναι ἄπειρον ὥστε ἐνεργείᾳ εἶναι ἐπὶ τὴν αὔξησιν ἀδιεξίτητον· οὐδὲ γὰρ νῦν δέονται τοῦ ἀπείρου (οὐ γὰρ χρῶνται), ἀλλὰ μόνον εἶναι ὅσην ἂν βούλωνται πεπερασμένην· τῷ δὲ μεγίστῳ μεγέθει τὸν αὐτὸν ἔστι τετμῆσθαι λόγον ὁπηλικονοῦν μέγεθος ἕτερον. ὥστε πρὸς μὲν τὸ δεῖξαι ἐκείνοις οὐδὲν διοίσει τὸ [δ'] εἶναι ἐν τοῖς οὖσιν μεγέθεσιν.

This passage has perplexed readers. In addition to the few problems with ellipsis, it is unclear what Aristotle has in mind. There are two issues about infinity in the text, both relevant.

1. The actual infinite does not exist, but the potential infinite exists. Mathematicians only need the potential infinite.
2. The universe is finite, so that there is in fact a largest line. So how could one prove things, e.g., about convergent lines that converge outside the universe.

Additionally, there are two general ways of understanding Aristotle's solution, perhaps not very different.

(a) The lines used by the mathematician are the lines in the diagram. These are finite, indeed. So the mathematician does not ever actually use infinite lines. The line cut in the diagram is 'proportional' to any line. This might be seen as a solution to (1) as well as (2).
(b) Since geometrical proofs are about similar configurations, there is a general principle in geometry that the metrical size of the figures is unimportant to the universalization from the particular case.

Of course, Aristotle gives us no hint whether he intends (a), (b), or something else. I do not find (a) very satisfactory, in as much as the diagram is a representation and not an instance of the large figure, especially in reductio arguments. So the convergence of parallel lines in *Elements* I appears as crooked lines, as here from Paris Gr2466(p) folio 9v. Obviously, in a reductio something has to be wrong. My point is merely that the finite depiction of two long lines is not a small representation of long-lines (Fig. 2).

Or take a representation of an infinite line from the same manuscript (4v, Proposition I 12). AB is infinite (Fig. 3).

(Footnote 84 continued)
existent magnitudes should be something that makes it unimportant that one does not have the larger magnitude. My suggestion is that the fact that the smaller figure is used makes no difference to the proof. So what exists in magnitudes is probably their being of any particular size, loosely from line b32.

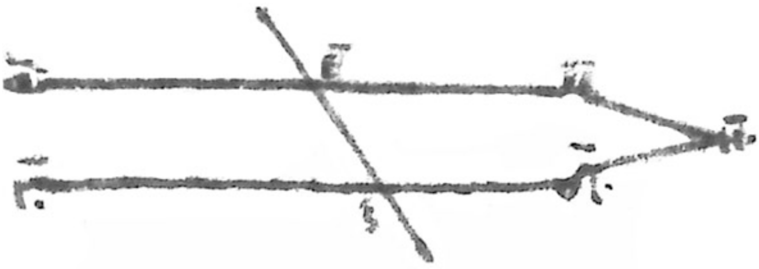


Fig. 2 Reductio arguments

Fig. 3 Representation of an infinite line

So much more is going on than merely reducing the size. I don’t wish here to involve us in questions about diagrams. It is enough that we need some account that deals with (1) and (2). Here, Hussey’s account of how to modify Euclidean mathematics to resolve (1) and (2) is probably fine.[85] Yet, I would like to make a few points about each.

If you adopt a different version of the parallel postulate, e.g., where you substitute for it a principle that interior alternate angles of a line intersecting two lines are equal iff the lines are parallel, you will be able to prove the other theorems needed for parallels. But this is not what Aristotle suggests. His principle concerns ratios. And as we shall see, the parallel theorems in Euclid are not the only problem with the geometry for a small world.

My second point concerns (1). If one allows an arbitrarily large but not actually infinite universe, it doesn’t really matter that one defines parallel lines as lines that will never meet no matter how they are extended. That is, the universe may be

[85] Hussey (1983, ad loc.).

potentially infinite. Proofs that lines are parallel can readily be accomplished by reductios, as is the norm in Euclid.

There is, however, another sort of infinitary proof that employs infinitary sequences. Such proofs normally involve issues distinct from my present concerns. However, Aristotle uses such a technique; so I don't see how he could object to them.[86] Indeed, Aristotle allows himself the luxury of setting up an infinitary, recursive sequence to argue that time is continuous iff motion is, concluding (*Phys*. Z 2.233a7–10):

> For the faster will divide the time, and the slower the length. And so if it is true that they always reciprocate, a division always comes about in their reciprocating.

> διαιρήσει γὰρ τὸ μὲν θᾶττον τὸν χρόνον, τὸ δὲ βραδύτερον τὸ μῆκος. εἰ οὖν αἰεὶ μὲν ἀντιστρέφειν ἀληθές, ἀντιστρεφομένου δὲ αἰεὶ γίγνεται διαίρεσις,

Aristotle does not give us the reductio that we expect. So he can have no serious objection to this type of argument, except that he can object to its being 'completed'. Nonetheless, this sort of argument does not impinge on questions about the infinite in extent.

I suspect it is common, however, at least to say of the standard texts of Greek mathematics that 'infinite' means 'potentially infinite' in Aristotle's sense, so that one can readily eliminate actual infinities from such texts. That is an interesting claim about the logic of Euclid's *Elements*. Since Aristotle probably invented the notion of the potential infinite, although it is implicit in Anaxagoras, if Euclid intends us to understand 'potential infinite' in a few places in *Elements* I, that would show a philosophical dependence of Euclid on Aristotle. There are four groups of texts that mention infinity in Euclid's corpus:

Texts relating to parallel lines (I def 23, post. 5 (cf. 2), props. 29, 44 (both quoting post. 5, but 29 for a reductio);
I 12: to construct a perpendicular from given point to an infinite line;
I 22, *Data* 39 (related to *Elements* I 22): construction of triangle from three given lines;
Text on number (VII 31: the fundamental theorem of arithmetic, but notably not IX 20, the prime number theorem);
Texts on the infinitude of irrationals (X Definition 3, Proposition 115 that there are infinitely different irrationals).

I shall not be concerned here with the parallel postulate, which states the familiar condition under which two lines infinitely extended will intersect. If lines intersect, they intersect at a finite length from a given point. So the fact that they are infinitely extended is convenient overkill. We can think of the parallel postulate as having a

[86] This is a very brief reply to Hussey on this point in a generally excellent discussion (ibid.).

logical form (with P(x,y) = x is parallel to y, and C(x,y) = x and y meet if extended infinitely):

Postulate 5: $F(x,y) \rightarrow C(x,y)$
Definition 23: $P(x,y) \leftrightarrow \sim C(x,y)$
So, the argument for the first proposition on parallels shows of two lines:
Proof of alternate angles (26):$A(a, b)$ & $C(a, b) \rightarrow \perp; \therefore A(a, b) \rightarrow \sim C(a, b)$

So even if in principle the lines could be extended infinitely, it really makes no difference whether we understand the lines as infinitely extended or as extended as much as is needed.

Nor will I be concerned with the fundamental theorem of arithmetic, that every composite number is measured by a prime number, since the infinite appears in a reductio, that if the number is not measured, then it will be measured by infinitely many, ever smaller numbers.

The texts that there are infinite irrationals are more interesting, as they involve infinitely many distinct constructions. Yet, there is some general agreement today that they are late additions to respond to the claim of X Definition 3 (it will be shown that to a given line there are infinitely many commensurable and incommensurable lines). The procedure is to start with a medial A and a rational B and to take the side of $O(A, B) = \Gamma$. Then take the side of $O(\Gamma, B) = \Delta$, and so forth. So, the author gives us a recursive procedure for generating infinitely many irrationals. Is this potential or actual infinity? Well, this has nothing to do with space and place. So we can drop the unanswerable question about a possibly unknown mathematician.

My principal concern then is with I 12, and I 22.

I 12 To draw a straight perpendicular line to a given infinite line from a given point (where 'altitude' is a line drawn to a given line at right angles, and a perpendicular is a line drawn at right angles from a given line. Direction is important!) (Ἐπὶ τὴν δοθεῖσαν εὐθεῖαν ἄπειρον ἀπὸ τοῦ δοθέντος σημείου, ὃ μή ἐστιν ἐπ' αὐτῆς, κάθετον εὐθεῖαν γραμμὴν ἀγαγεῖν.)

I 22 To construct a triangle from three straight lines, which are equal to given straight-lines. … The proof constructs the base of the triangle on a line that's infinite in one direction: Let a line be displayed DE which is finite along D, but infinite along E, … (Ἐκκείσθω τις εὐθεῖα ἡ ΔΕ πεπερασμένη μὲν κατὰ τὸ Δ ἄπειρος δὲ κατὰ τὸ Ε, …)

I 12 is in Book I to complement the drawing of a perpendicular from a given line, but, as Mueller informs us,[87] it is not used until the squaring problem, II 14. Proclus is probably right that Euclid decides to construct the figure on an infinite line, because otherwise there is no guarantee that there will be a altitude to the line, which may fall short. A change in the nature of the problem might deal with the

[87] Mueller (1981, p. 20).

problem. Draw a line from the given point to the given line. Double the length of this line. Draw a circle with the double line as radius and the given point as center, and extend the given line, if necessary, to meet the circle. The altitude will meet the extended line. A similar issue is at work in I 22. One needs a line that is at least as long as the sum of the three lines. So, for convenience, take a partially infinite line. However, this can be very trivially avoided. Instead of cutting off three segments of the partially infinite line, just draw a line with one length and, using circles, extend it twice, for the other two lengths. So why does Euclid ask for infinite lines?

Should we assume that Euclid understands these infinite lines to be really lines that are long enough, or should we just take him to be saying that he wants actual infinite lines? Here the issue is not what Euclid could have done. That is trivial enough. It is what he chose to do. If he had the slightest queesiness about infinite lines, he easily could have avoided them, as he does, in effect, in the case of parallel lines. I should also point out that this rather cavalier introduction of merely convenient infinite lines is somewhat unusual in Greek mathematics.[88]

We can contrast the ambiguity in Euclid with Apollonius. He really does use infinite to mean potential infinite but also has no difficulty in speaking of actual infinities as well. For example, in stating the asymptote theorem for hyperbolas (II 14), he says, "As they are extended ad infinitum, the asymptotes and the section move nearer to themselves and arrive at a distance smaller than any/every given distance (Αἱ ἀσύμπτωτοι καὶ ἡ τομὴ εἰς ἄπειρον ἐκβαλλόμεναι ἔγγιόν τε προσάγουσιν ἑαυταῖς καὶ παντὸς τοῦ δοθέντος διαστήματος εἰς ἔλαττον ἀφικνοῦνται διάστημα)." One might worry about the second clause, but the quantifier is correctly placed and there is nothing more infinitary than the Anaxagorean/Eudoxan principle that it is possible to get a magnitude smaller than any given magnitude, as is clear from the proof, where we are given a magnitude K to get a smaller distance.[89]

One might also, with much caution, see an infinite, spatial plane in Apollonius' marking out four regions (τόποι) by the asymptotes to opposite hyperbolas of the curve at II 33, that is the two regions determined by the 'angle containing the section', i.e., by the asymptotes to the two hyperbolas, and the two regions outside the two angles formed by the asymptotes to the two hyperbolas. In post-Cartesian mathematics, these evolve into, but should not be confused with the regions determined by x and y axes. These regions, as determined by the asymptotes, are infinite if the asymptotes are infinite. It would be anachronistic to see here an

[88] The Heronian Definitions 119 clearly understands actual infinities in talking about infinite magnitudes as, "A magnitude is what increases and is cut ad infinitum. There are 3 species of it: line, surface, solid. A magnitude is infinite than which nothing larger is conceived in existence (ὑπόστασιν) of whatever size so that there is no limit of it." There is no suggestion that such magnitudes are impossible.

[89] So too I Definition 1, Proposition 8 (the infinite extension of a section whose diameter is parallel to the axis of the cone). Here, one might note that Apollonius says that as the surface of the cone and the cutting plane increase ad infinitum, the section also increases. One might wonder about the claim at II 44 that we will find infinite diameters to a section. After all, he could have said 'however many'. Basically, it is as if one said, "it is possible to find infinite parallels to a given line." Is Apollonius speaking hyperbolically?

infinite space divided up into four sections within which the asymptotes lie, as the spatial regions are determined by the asymptotes, and, as sector of the cone, the plane is just another, perhaps infinite geometrical object, that is, if the question what it is ever gets asked. Yet there is no evidence anyone put these two together into a notion of geometrical space (as opposed to geometrical object).

In fact, in Book VI, Apollonius clearly treats some conic lines as infinite.[90] Props. 1–10 concern equal conic sections, those that can be superimposed (Definition 1). Consider Proposition 1, "If the upright sides of parabolic sections (the parameter or *latus rectus*), applied to which the perpendiculars drawn down to the axes equal in power (the ordinates), are equal, then the sections are equal, and if the sections are equal the upright sides are equal (my trans.)." This is only true if the parabolic sections are infinite; otherwise, two parabolic sections with the same parameter but axes of different lengths will not superimpose and be equal. Recall that the square on the ordinate equals the rectangle of the parameter and the cut off part of the axis (the abscissa). And the proof of the proposition bears this out. So too, in proving Proposition 3, that none of the three sorts of sections, ellipse, hyperbola, parabola, is equal to any section of the other sorts, Apollonius argues that it is obvious that an ellipse is not equal to a parabola or hyperbola, because it is finite while a parabola and hyperbola is infinite (Toomer 1985, p. 275.8; Rashed 2009, p. 101.16).

When it comes to infinite lines, I conclude then that Euclid and Apollonius were not bothered.[91]

6 Alexander Was Bothered; Was Anyone Else? Well, Maybe by the Practical?

In his commentary on *Physics* III 8, Simplicius reports an objection of Alexander of Aphrodisias to Euclid, *Elements* I 1. What if we want to construct an equilateral triangle on the diameter of the universe? The construction requires that we go beyond the rim of heaven, so that Euclid's construction is not universal (Simplicius, *In Arist. phys*. 511.30–2.9)

> Alexander inquired how the first theorem of the *Elements* of Euclid is not destroyed, if it is not possible also to extend a straight line outside the universe or to draw a circle (for [this would be the case] if the given finite straight-line on which it is required to construct the equilateral triangle can be the diameter of the kosmos, and it is impossible to construct an equilateral triangle on this, if there be nothing outside the kosmos, since the diameter of the universe becomes the radius of the circles where the [lines] joining their common section to the end-points of the given [line] with it produce the equilateral triangle). After inquiring, he solves it, saying, "Since [that magnitude] is infinite [i.e., potentially] where it is always possible for those who take it in quantity to take something outside, as was shown, it is

[90] I am very thankful to Vincenzo De Risi for pointing this out in a workshop on Apollonius at the Humboldt University, July, 2014. Some comments of Sabetai Unguru on Book III at the same workshop were also particularly helpful for my discussion of Apollonius.

[91] The British reader is permitted to see an allusion to Catherine Tate.

clear that the mathematicians also suppose such lines which they suppose as infinite, so that it is possible that they increase. For these [lines] are infinite where there is something outside. But it is not possible for the diameter of the kosmos to grow. And so they suppose a [line] smaller than the diameter, if they suppose a finite-line, since those [lines] are also infinite which they can add to and which they can extend."

511.30 Ζητήσας δὲ ὁ Ἀλέξανδρος, πῶς οὐκ ἀναιρεῖται τὸ πρῶτον θεώρημα
τῶν Εὐκλείδου Στοιχείων, εἴπερ μὴ καὶ ἔξω τοῦ παντὸς δυνατὸν εὐθεῖαν
ἐκβάλλειν ἢ κύκλον γράφειν (εἰ γὰρ δύναται μὲν ἡ δοθεῖσα εὐθεῖα πεπε-
ρασμένη, ἐφ' ἧς δεῖ τὸ ἰσόπλευρον τρίγωνον συστήσασθαι, ἡ διάμετρος εἶναι
τοῦ κόσμου, ἀδύνατον δὲ ἐπὶ ταύτης τρίγωνον ἰσόπλευρον συστήσασθαι, εἰ
511.35 μηδὲν εἴη τοῦ κόσμου ἐκτός· ἡ γὰρ διάμετρος τοῦ παντὸς ἐκ τοῦ κέντρου
512.1 γίνεται τῶν κύκλων, ὧν ἀπὸ τῆς κοινῆς τομῆς αἱ ἐπὶ τὰ πέρατα τῆς δοθεί-
σης ἐπιζευγνύμεναι τὸ ἰσόπλευρον τρίγωνον μετ' αὐτῆς ποιοῦσι), τοῦτο οὖν
ζητήσας λύει λέγων· "ἐπειδὴ ἄπειρόν ἐστιν, οὗ κατὰ ποσὸν λαμβάνουσιν
ἀεί τι λαβεῖν ἔστιν ἔξω, ὡς δέδεικται, δῆλον ὅτι οἱ μαθηματικοὶ καὶ ἃς
512.5 ἀπείρους γραμμὰς ὑποτίθενται, τοιαύτας ὑποτίθενται, ὥστε δυνατὸν αὐτὰς
αὐξῆσαι. ὧν γάρ ἐστιν ἐκτός τι, αὗται ἄπειροι· τὴν δὲ διάμετρον τοῦ
κόσμου οὐκ ἔστιν αὐξῆσαι· ἐλάττονα οὖν ὑποτίθενται τῆς διαμέτρου, εἴπερ
πεπερασμένην ὑποτίθενται, ὁπότε καὶ ἄπειροι αὗταί εἰσιν, αἷς προσθεῖναι
δύνανται καὶ ἃς ἐκβάλλειν."

The professor of Aristotelian studies needs to assume that mathematicians just keep their lines smaller. The neo-Platonic exile has no such difficulty, since he holds that the objects of mathematics reside in the imagination, where the potential infinite is unencumbered by physical limitations, and this, presumably, is why we do not find Alexander's puzzle in Proclus' commentary on Euclid, although he is concerned about actual infinities. Yet it is difficult to know whether it would have been, in fact, a serious problem even for Aristotle, since the maximal line is still the diameter of the physical universe qua diameter of the heaven qua sphere, so that the size does not enter in. Does the limitations on the size of the line qua perceptible line have any impact on constructions on the same line qua line?

If such a puzzle appeared among philosophers in the Hellenistic Age, we would expect, however, it to be a response to Stoic or late Peripatetic and not Aristotle's own views of the heaven and to Stoic or late Peripatetic and not Aristotle's own views of geometrical objects, at least not before the first century BCE Aristotelian revival. Now our Stoic fictionalists need not trouble about the issue. Yet an abstractionist might be concerned whether lines represented in the imagination need to be no larger than some size. There would seem to be at most one Stoic who might have been so concerned, Posidonius, who thought that the void is only large enough to accommodate the exploded cosmos during the cyclical, great conflagration.[92]

[92] Cf. note 47.

Now, according to al-Nairīzī[93] and Proclus (*In Eucl.* 176.5–17),[94] Posidonius attempts a variation on the definition of parallel lines (Proclus, *In Eucl.* 176.5–17):

> And Euclid defines parallel lines in this way, but Posidonius says that parallels are those in one plane that neither converge nor diverge but have all their perpendiculars equal that are drawn from points on one to the other. Those that always make their perpendiculars smaller converge with one another. For the perpendicular is able to determine the heights of areas and the distances between the lines. Hence, when the perpendiculars are equal, the distances between the lines are equal, when larger or smaller the distances too will be made smaller and they will converge on the side where the perpendiculars are smaller.

> Καὶ ὁ μὲν Εὐκλείδης τοῦτον ὁρίζεται τὸν τρόπον τὰς παραλλήλους εὐθείας, ὁ δὲ Ποσειδώνιος, παράλληλοι, φησίν, εἰσὶν αἱ μήτε συνεύουσαι μήτε ἀπονεύουσαι ἐν ἑνὶ ἐπιπέδῳ, ἀλλ' ἴσας ἔχουσαι πάσας τὰς καθέτους τὰς ἀγομένας ἀπὸ τῶν τῆς ἑτέρας σημείων ἐπὶ τὴν λοιπήν. ὅσαι δ' ἂν ἐλάττους ἀεὶ ποιῶσι τὰς καθέτους συνεύουσιν ἀλλήλαις· ἡ γὰρ κάθετος τά τε ὕψη τῶν χωρίων καὶ τὰ διαστήματα τῶν γραμμῶν ὁρίζειν δύναται. διόπερ ἴσων μὲν τῶν καθέτων οὐσῶν ἴσα τὰ διαστήματα τῶν εὐθειῶν, μειζόνων δὲ καὶ ἐλαττόνων γινομένων καὶ ἡ ἀπόστασις ἐλασσοῦται καὶ συνεύουσιν ἀλλήλαις, ἐφ' ἃ μέρη εἰσὶν αἱ κάθετοι ἐλάσσονες.

The type of definition here is a locus definition such as Euclid's definition of 'circle' (see § 1). What does Proclus hope to gain by this definition? First, it does not, in fact, avoid the parallel postulate.[95] One can construct a line with two points equidistant from a given line—draw two perpendiculars equal to each other and connect them, but it will not necessarily be the case that every other perpendicular will be equal nor that the angles at the points from which the perpendiculars are drawn will be right angles. I have no idea if this is significant, since all attempts to prove the parallel postulate must fail. Nonetheless, neither al-Nayrizi nor Proclus suggest that the motivation was the parallel postulate, nor do they mention it again. Furthermore, al-Nayrizi[96] mentions another definition by a contemporary of

---

[93] Cf. al-Nairīzī (2009, pp. 16–17) followed by (2003, p. 88).

[94] Either both are based on Heron's commentary on Euclid, or, as seems more likely, Proclus and Heron (whom al-Nairīzī uses) are each based on Geminus.

[95] Euclid, *Elements* I post. 5: and that if a straight line falls on two lines makes the interior angles on the same parts smaller than two right angles, then if the two lines are extended infinitely on the parts where the angles are smaller than two right angles, they meet. Euclid tacitly assumes the converse (cf. *Elements* I 17), if the lines meet the angles are less than two right angles. Note, per my example, that there is nothing in Posidonius' definition about whether parallel or non-parallel lines do or do not meet when extended.

[96] al-Nairīzī (2003, pp. 88–9). Even if the source is Geminus, modifying Posidonius, the point remains that this definition is the one used by Agapius (see next two notes).

Simplicius, Agapius,[97] that parallel lines are equidistant even if extended ad infinitum on both sides. This definition then forms the basis, after I 27,[98] of attempts to prove the parallel postulate. As to Proclus, when he turns to the parallel postulate after I 29,[99] Posidonius is absent. So Posidonius' divergence from Euclid might not have been due his concern to avoid the parallel postulate.

More to our point, the construction avoids the need for an infinite universe. The Euclidean definition, that the lines when extended infinitely will not converge, presupposes that the lines could be extended infinitely and be seen not to converge. The parallel postulate guarantees that the exercise need not be undertaken, but the definition requires that it could be.

Let's assume that Posidonius has taken care of the parallel postulate. Then one can show that two lines are parallel simply by constructing arbitrary perpendiculars between them or their extensions (finite) and showing that any two are equal, i.e., if the lines can be extended in some direction to where there are perpendiculars. One can similarly show that they 'converge' by showing that they are unequal, where this is all 'converge' will mean.[100] Of course, what else our Aristotelianizing Stoic hopes to gain, we do not know. He may want it possible to extend all converging lines to where they actually meet (and so leave the parallel postulate where it is), or he may have required a strictly finite geometrical world. Nor do we know if a finitism, corresponding his views on void, even was his intent. But, either way, it would seem more reasonable to put Posidonius in the class of abstractionists. Well, another nice story.

So far as I can tell, no other text that is remotely mathematical suggests Alexander's puzzle. Proclus mentions three instances where someone does raise a difficulty for a construction in a problem because there isn't place (τόπος) to do it:

I 2 To put a line equal to a given line at a given point. (*In Eucl*. 225.16, 225.8–227.8)

I 9 To bisect a given finite line. (*In Eucl*. 275.7, 275.7–277.4)

I 12 To draw a straight perpendicular line to a given infinite line from a given point. (*In Eucl*. 289.18–20, 289.16–290.13), where there isn't enough place on the other side of the given line)

[97] On the identity of Agapius, cf. Lo Bello in al-Nairīzī (2003, n. L5: 224–229).

[98] al-Nairīzī (2003, pp. 157–6).

[99] Proclus, *In Eucl*. 365–73, cf. 362–3.

[100] Of course, in a finite spherical universe, all parallel lines could be shown to be parallel in this way, with some condition that will allow that there might be points on one line from which perpendiculars cannot be drawn to the other. But some non-parallel lines cannot be extended so that perpendiculars can be drawn between them. One suspects that the geometry would end up with more postulates to deal with what could easily be proved by treating the spherical world as a subspace of an infinite one.

We should also put in this list:

I 11 To draw a straight line at right angles to a given straight-line from point given on it. (*In Eucl*. 281.6–7)

Here one is asked to construct the line from an end point without extending the given line. One could surely do this if there were place (τόπος). A modern reader, I suspect, will tend to read these discussions as posing an objection to Euclid's solution to the respective problem. Proclus (*In Eucl.* 389.7–15, ad I 12), however, is careful to distinguish between an objection (ἔνστασις) and a case (πτῶσις—different configurations in proving a theorem or doing a problem) and rightly, I think, treats the difficulty as a demand for a case. Proclus puts up a small barrier in the way of our using the four examples, namely that the commentators did not distinguish between cases and objections. So perhaps he and I are wrong, and someone did think that the lack of place (τόπος) was an objection to the constructions.

It turns out that our commentators might well be Heron of Alexandria. In his commentary on Euclid, al-Nayrizi (Anaritius) preserves a construction identical to that of Proclus' commentary on I 11 (2003, 128–9) and mentions Hero as the author. So, Heath[101] suggests that we add several proofs in al-Nayrizi to the list, besides I 11,

I 16 In every triangle, when one side is extended, the outside angle is larger than each of the opposite and interior angles. (*In Eucl*. 305.21–6, with Hero mentioned)

I 20 In every triangle, two sides, taken in every way, are larger than the remaining side. (2003, 140–4) (*In Eucl*. 323.5–326.5, with Hero and Porphyry mentioned)

I 48 converse of the Pythagorean theorem (2003, 202–3, with Heron mentioned) (cf. *In Eucl*. 430.9–431.14)

Since al-Nayrizi mentions the 'objection' to I 16, we should be cautious. I am not sure why Heath includes it in his list, since Proclus merely says that Heron reports an objection of Philip of Mende that triangles qua triangles do not have exterior angles. Whatever Philip's objection may have been, it apparently has nothing to do with lack of place (τόπος) and everything to do with the nature of theorems about triangles.

Whereas Euclid's construction for I 20 involves two lines outside the triangle, Proclus gives three proofs, the second incomplete, of I 20 which he attributes to Heron and Porphyry, where each, but the first especially, is genuinely simpler in its construction than Euclid's proof, as it uses one line inside the triangle. So the lack of a construction exterior to the triangle may have nothing to do with lack of place (τόπος). A similar remark may be made for I 48. Here, it is enough to show that any two triangles with sides a, b such that $T(a) + T(b) = T(\text{the third side})$ are congruent.[102] So one takes a triangle a, b, c where $T(a) + T(b) = T(c)$ and constructs a

[101] Heath (1926, pp. 22–3).

right triangle with legs equal to *a*, *b*. Heron's version builds the right triangle with *a* as one of the legs, and then shows that the other leg is the same line as *b*. I don't know if the proof is simpler, but it fits with an aesthetic that minimizes auxiliary constructions and seeks to show that two lines with different basic properties are the same line. In sum, none of these proofs need have anything to do with issues of adequate places.

If my argument so far is unsound, then one could argue Heron shows a concern similar to Alexander's. Can one do Euclid's *Elements* in such a way that the size of the universe does not matter? For each of these is a theorem, and there is no reason to worry about adequate room in proving a theorem unless you believe:

> The proof of the theorem about some figure either employs a construction that fits every possible location and size of the figure or breaks up into cases that it covers every possible location and size.

If so, we would have a motivation to see in Heron the influence of some naturalist school, perhaps Strato, who, I noted, seems to have influenced Heron's pneumatics and who held that the universe is finite but not surrounded by infinite void in the way that Stoics believed. Presumably, this is something like what Hintikka supposed in suggesting that Heron was troubled to take into account a finite universe.[103]

As delightful as this speculation might be, I fear that there is a much simpler explanation of Heron's concerns about adequate place (τόπος). Suppose that the only propositions where Heron expressed a concern for adequate space were problems. In our texts, these are I 2, 9, 11, and 12. I need to draw a line at right angles to the end of a given line. But the line goes into a wall. I cannot draw it with the construction in Euclid. So I use Euclid's constructions to construct a square, the fourth line of which is my desired line. I need to construct a perpendicular to a line, but there is no room on the other side. Well, one might suspect a bit of a joke here. I have an infinite line, and there isn't room on the other side of the line. For he doesn't drop the condition that the line be infinite. Nonetheless, the situation is real enough. I have a very long wall and need to draw a perpendicular to the wall and cannot get to the other side. If so, Proclus is right to treat these as cases and not as objections. But what sort of cases?

Now there is something quaint in all this, something that might even be familiar to readers of Heron. I have no idea whether Euclid's constructions might be practical in the real world. I'll leave that to historians of architecture, while noting that sometimes they actually are, and at other times aren't. It is not just that we expect this from Heron. In a context of pure geometry where these issues do not

[102] 'T(x)' is 'the square on x'.

[103] Hintikka (1973, pp. 121), "Aristotle's compunctions about geometrical constructions were apparently shared by at least one well-known mathematician of antiquity. Heron mechanicus tried to dispense with the production of particular straight lines as much as possible, motivated by the idea that there might not always be enough space available to carry out such a production. (It does not matter for my purposes whether Heron was himself worried about this or whether he was trying to reassure others)…"

arise, he brings in constructions that look to applications in tight spaces. Or is it rather that in his commentary, Heron is bringing geometry closer to applied geometry? Of course, I don't know the answer to this question. Yet, it is clear that there is nothing wrong with saying, "Here is a way of constructing a perpendicular," and saying, "By the way, when you apply geometry, you might find this construction useful."

The Arabic translation of Pappus, *Collectio* VIII on mechanics has a remarkably similar concern, where the problems are also close to the selection in Proclus and al-Nayrizi.[104] Suppose you wish to do a construction that requires a compass, but your compass just isn't big enough. The first problems (two methods) presumes that the compass can be smaller, where you want to draw a perpendicular from a given point on it. For the rest you use the maximum width of the compass: to bisect a given line, to trisect it, etc.; to add to a line a segment equal to it; to extend another line with an equal to a given line (3 cases); to construct a triangle from given lines, where any pair is larger than the third—this gets checked (2 cases: where two sides are equal, where none is equal). In fact, the second case of the first problem is a construction from the end-point, where there is no place. Here, again, it is plausible that the issue is about ordinary conditions and not cosmic ones. It is very reasonable to guess that the ultimate source for the Arabic translation was either Heron or a treatise in the Heronian tradition.[105]

The most geometrical of Heron's extant treatises is the *Metrica*, in the sense that the ingredients should be shapes that are to be measured (see above § 1). Except as an appendix to Book I (measurement of areas) and II (measurement of volume) on the measurement of disorderly figures, the figures are taken from the rich mathematical tradition. Here, Heron does distinguish between the metrical and the geometrical (see note 20). So the practical concerns expressed in the commentary on Euclid and the Arabic translation of Pappus are metamathematical. They explain why one might be interested in investigating certain constructions—there might be a wall; the compass might not be big enough.

Furthermore, we can look elsewhere for implicit concerns about physical limitations on geometrical constructions that are quite ordinary. As Sidoli and Saito (2009) observe, in *De sphaera*, Theodosius will employ many sorts of constructions involving the interiors of spheres, even in the demonstration of problems, while the distinction between problems and theorems is not rigorous, except in one respect. With the exception of finding the center of the sphere (I 2), necessary for the rest of the treatise, the other six problems themselves can be accomplished by working on

[104] Jackson (1980). These are sometimes described as rusty compass problems, but, as I indicate, this would not seem to be the concern. Cf. Section F (p. 527), "Now that is not easy to do using the construction mentioned by Euclid in his *Elements*, since we have only one small pair of compasses with which to work."

[105] Jackson (1980) makes an excellent argument that the excerpt belongs in Pappus, *Collectio* VII, but it is much less clear how it belongs, except as a completely distinct topic. He reports (p. 524) that it occurs between § 44, which concerns inscribing seven hexagons in a circle, six about the seventh (constructions that make the demonstration obvious) and § 45, about the juxtaposition of gears.

the surface of the sphere or outside it. As far as the abstract stereometry of spheres is concerned, this is irrelevant, but if one is working with practical, bronze or marble spheres, where lines need to be drawn for that high market star globe, one needs to avoid penetration. So, Heron is well within a tradition that treats problems as practical. Indeed, if we look, we will find other examples, such as Eratosthenes criticism of earlier duplications of the cube due to their impracticality and his praise of his own.[106]

Recall that I did not want to make generalizations about philosophies of Greek geometers. Nonetheless, we have seen that Euclid has no difficulty stipulating an infinite line. Of course, what Euclid intended by that stipulation we cannot know. It is enough to note that we have no reason to suppose that Euclid abjured the actual infinite. Secondly, Apollonius happily uses actual infinite lines. Thirdly, we saw that Heron and Theodosius seem to bring applied geometrical considerations into their treatment of problems in texts usually thought to be pure geometry. That dictates restrictions on a solution to a problem, but says little about how they think of geometry or the world, except that geometrical constructions should be practical. Of course, this too could be wrong. Heron might even have concerns about a finite universe, although, if he did, it would yet more curious why he thinks that there might be no place (τόπος) on the other side of an *infinite* line.

7 Conclusion

I began my discussion with concerns about four basic spatial notions in Greek philosophy and mathematics, well two significant, place (τόπος) and position (θέσις), and two relatively insignificant in mathematics, void (κενόν) and room (χώρα). Here I argued that ordinary notions bereft of philosophical analysis were adequate to describe how mathematicians treated objects. Constructions operate in ordinary worlds without rich notions of space. I then turned to Aristotle's views on place and how his views on mathematical objects can hang together with his views about place. I did not come up with a single answer to the question, but noted that for Aristotle there is a sense in which geometrical objects have place, but that it had to be different at least from the way in which physical, locomotive objects have place. This seems more a reflection of mathematical usage than a profound new way of thinking about mathematical objects. Quite the contrary, Aristotle's conception of mathematical objects as perceptible magnitudes qua magnitudes leaves open the question what gets removed by the 'qua' operator, in our case, 'place' or 'absolute direction'. This is what made Alexander's puzzle about the finite world poignant. I then turned to more plausible candidates for philosophers interacting with Greek mathematicians. Not surprisingly, we did not find much, although many philosophers, Epicurean and Stoic, seem to have tried. However, I also tried to illustrate my point that we expect different Greek mathematicians to have their own

[106] Euctocius, *In Arch. de sphaer. et cyl.* 90.4–14.

quirks. In the case of Euclid, he was happy to request an infinite line; Apollonius, an actual infinite line. In the case of Heron (or ps.-Heron), limitations of place (τόπος) came in, but this was not the cosmic limitation that might have bothered Posidonius. It was the limitations that come into consideration from an interest in practical problems, that also occurs in Theodosius and Eratosthenes. But Heron may also have been a physicalist, in strong contrast to the abstractionism of post-Stoic Aristotelians and Platonists. His attitudes certainly influences the Heronian corpus and tradition. I am sorry if I have disappointed my reader in two ways. First, I did not find any grand theory of place in Greek mathematics. I also did not find a view about infinite lines in Greek mathematics. Instead, I found individual mathematicians introducing the infinite in different ways and without comment. I found variety, and I like that.[107]

References

Acerbi, F. (2003). On the shoulders of Hipparchus: a reappraisal of ancient Greek combinatorics. *Archive for History of Exact Sciences, 57*, 465–502.

al-Nairīzī, A. (2003). *The commentary of Al-nayrizi on book I of Euclid's element's of geometry* (A. Lo Bello, Trans.). Boston: Brill.

al-Nairīzī, A. (2009). *The commentary of Al-nayrizi on books II-IV of Euclid's element's of geometry, with a translation of that portion of book I missing from MS Leiden Or. 399.1 but present in the newly discovered Qom manuscript* (Rüdiger Arnzen, Ed., A. Lo Bello, Trans.) Boston: Brill.

Algra, K. (1995). *Concepts of space in Greek thought*. Leiden: Brill.

Buchheim, T. (2010). *Aristoteles: Über werden und vergehen*. Berlin: Akademie Verlag.

Cambiano, G. (2008). Philosophy, science and medicine. In K. Algra, J. Barnes, J. Mansfeld, & M. Schofield (Eds.), *Cambridge history of hellenistic philosophy* (pp. 585–613). Cambridge: Cambridge University Press.

Crönert, W. (1900). Der Epikureer Philonides (König. Preus. Akad. der Wiss. 41).

Christianidis, J., & Demis, A. (2010). Archimedes' quadratures. In S. A. Paipetis & M. Ceccarelli (Eds.), *The genius of Archimedes* (pp. 57–68). Dordrecht: Springer.

de Haas, F., & Mansfeld, J. (Eds.). (2004). *Aristotle's on generation and corruption I*. Oxford: Oxford University Press.

de Harven, V. (2012). The coherence of stoic ontology. Ph.D. Dissertation, University of California, Berkeley.

Hahm, D. (1972). Chrysippus' solution to the Democritean dilemma of the cone. *Isis, 63*, 205–220.

Heath, T. L. (1926). *The thirteen books of Euclid's elements* (2nd ed.). Cambridge: Cambridge University Press.

Hintikka, J. (1973). Aristotelian infinity. In J. Hintikka, *Time and Necessity* (Oxford: Oxford University Press, 114–134). Reprinted with revisions from *Philosophical Review, 75*, 197–218.

Husserl, E. (1989). *Origin of geometry* (J. P. Leavey, Trans.). Lincoln: University of Nebraska Press.

[107] I should like to thank all the participants from the Conference on Space at the Max Planck Institute (Aug. 2012), many of whom made valuable observations, particularly Dan Garber, Gary Hatfield, Alex Jones, and Vincenzo De Risi, as did Michalis Sialaros and David Sedley later on. I also thank Vanessa de Harven and Francesco Verde for sending me their dissertations.

Hussey, E. (1983). *Aristotle's physics III & IV* (Trans. & commentary) Clarendon Aristotle Series. Oxford: Oxford University Press.

Jackson, D. E. P. (1980). Towards a resolution of the problem of τὰ ἑνὶ διαστήματι γραφόμενα in Pappus' *Collection* Book VIII," *Classical Quarterly* n.s. 30 (pp. 523–533).

Joachim, H. H. (1922). *Aristotle: On coming-to-be and passing-away*. Oxford: Oxford University Press.

Jones, A. (1986). *Pappus of Alexandria: Book VII of the collection*. New York: Springer.

Ju, A. (2009). The stoic ontology of geometrical limits. *Phronesis, 54*, 371–389.

Kidd, I. G. (1988). *Posidonius: The commentary*. Cambridge: Cambridge University Press.

Knorr, W. R. (1986). *The ancient tradition of geometric problems*. Boston: Birkhäuser.

Knorr, W. R. (1996). The method of indivisibles in ancient geometry. In R. Calinger (Ed.), *Vita Mathematica: Historical research and integration with teaching* (pp. 67–86). Washington, DC: Mathematical Association of America.

Long, A. A., & Sedley, D. (1987). *The hellenistic philosophers* (Vol. 1). Cambridge: Cambridge University Press.

Luria, S. (1933). Die Infinitesimaltheorie der antiken Atomisten. *Quellen und Studien B, 2*, 85–106.

Makin, S. (2006). *Aristotle, metaphysics book Θ*. Oxford: Oxford University Press.

Mansfeld, J. (1983). Intuitionism and formalism: Zeno's definition of geometry in a fragment of L. Calvenus. *Phronesis, 28*, 59–74.

Mendell, H. (1984). Two geometrical examples from Aristotle's metaphysics. *Classical Quarterly, 34*, 359–372.

Mendell, H. (1987). Topoi on topos: The development of Aristotle's theory of place. *Phronesis, 32*, 206–231.

Mendell, H. (2004). Aristotle and mathematics. *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/aristotle-mathematics/.

Mendell, H. (2005). Putting Aristotle's *physics* in its place: A discussion of B. Morison, *On Location*. *Oxford Studies in Ancient Philosophy, 25*, 329–366.

Morrow, G. R. (1992). *Proclus diadochus: A commentary on the first book of Euclid's elements, with a forward by I. Mueller*. Princeton: Princeton University Press.

Mueller, I. (1981). *Philosophy of mathematics and deductive structure in Euclid's elements*. Cambridge: MIT Press.

Mueller, I. (1990). Aristotle's doctrine of abstraction in the commentators. In R. Sorabji (Ed.), *Aristotle transformed* (pp. 463–480). Ithaca: Cornell University Press.

Mugler, C. (1958). *Dictionnaire historique de la terminologie géométrique des Grecs*. Paris: C. Klincksieck.

Netz, R., Saito, K., & Tchernettska, N. (2001). A new reading of *method* proposition 14: Preliminary evidence from the Archimedes palimpsest (Part 1). *Sciamus, 2*, 9–29.

Netz, R. (1999). *The shaping of deduction in Greek mathematics: A study in cognitive history*. Cambridge: Cambridge University Press.

Owen, G. E. L. (1961). Tithenai ta Phainomena. In S. Mansion, (Ed.), *Aristote et les problèmes de méthode* (pp. 83–103). Louvain: Publication Universitaire de Louvain. Reprinted in Owen (1986b).

Owen, G. E. L. (1970). Aristotle: Method, physics and cosmology. In C. C. Gillespie, (Ed.), *Dictionary of Scientific Biography* (Vol. 1, pp. 250–8) (New York: Charles Scribner's Sons). Reprinted in Owen (1986b).

Owen, G. E. L. (1976). Aristotle on time. In P. Machamer & R. Turnbull (Eds.), *Motion and time, space and matter* (pp. 3–27). Columbus: Ohio State University Press. Reprinted in Owen (1986b).

Owen, G. E. L. (1986a). Aristotelian mechanics. In A. Gotthelf, (Ed.), *Aristotle on nature and living things* (pp. 227–245). Pittsburgh: Mathesis Publications; reprinted in Owen, pp. 315–33 (1986b).

Owen, G. E. L. (1986b). *Logic, science, and dialectic*. Ithaca: Cornell University Press.

Paparazzo, E. (2005). The elder pliny, posidonius, and surfaces. *British Journal for the Philosophy of Science, 56*, 363–376.

Rashed, R. (2009). *Apollonius de Perge, Coniques: Tome 4: Livres VI et VII: Commentaire historique et mathématique, edition et traduction du texte arabe*. Berlin: De Gruyter.

Robertson, D. (2004). "Chrysippus on mathematical objects," *Ancient Philosophy*, pp. 169–191.

Sedley, D. (1982). Two conceptions of vacuum. *Phronesis, 27*, 175–193.

Sidoli, N., & Saito, K. (2009). The role of geometrical constructions in Theodosius's spherics. *Archive for History of the Exact Sciences, 63*, 581–609.

Toomer, G. (1985). *Apollonius of perga, conics books V to VII: The arabic translation of the lost Greek original in the version of Banu Musa*. New York: Springer. Vols. 2.

Tybjerg, K. (2004). Heron of Alexandria's mechanical geometry. *Apeiron, 37*, 29–56.

Verde, F. (2010). *Elachista: la dottrina dei minimi nella tradizione epicurea*. Tesi di dott.: Università di Roma.

White, M. J. (1992). *The continuous and the discrete*. Oxford: Oxford University Press.

A Note on Lines and Planes in Euclid's Geometry

Jeremy Gray

The purpose of this note is to remind readers of information at times well-known and at times almost forgotten, namely that for several centuries in the modern West Euclid's *Elements* was simultaneously regarded as the epitome of knowledge and as flawed and confused. It is well known that many mathematicians brought up on Euclid and other Greek geometers complained that they found themselves compelled to accept the conclusions but not instructed in how to do geometry, and the long struggle with the parallel postulate has also been frequently discussed. The confusion discussed here is different, and relates to the concepts of straightness and shortest distance. It will also be suggested that the growing awareness of the defects in Euclid's presentation by the end of the 18th century contributed to the creation of the new geometries of the 19th century: projective geometry and non-Euclidean geometry.

It is, of course, notoriously difficult to give good definitions of fundamental concepts, and in any system of ideas some are going to have to be left undefined. The view to be taken here, in line with Lambert's opinion in his (1786), is that a careful reading of Euclid's *Elements* should equip the student to understand it and to use its ideas correctly, and that it does not adequately do so. The confusion can be introduced by comparing two early theorems in the *Elements*. Proposition I.2 asserts that any given line segment in a plane may be copied exactly with one of its end points at any prescribed point in the plane. The proof is quite long, it is scrupulous, and far from obvious. Proposition I.4, in Heath's translation, asserts that:

J. Gray (✉)
Department of Mathematics and Statistics, The Open University, Walton Hall,
Milton Keynes, Buckinghamshire MK7 6AA, UK
e-mail: jeremy.gray@open.ac.uk

V. De Risi (ed.), *Mathematizing Space*, Trends in the History of Science,
DOI 10.1007/978-3-319-12102-4_3

> If two triangles have the two sides equal to two sides respectively, and have the angles contained by the equal straight lines equal, they will also have the base equal to the base, the triangle will be equal to the triangle, and the remaining angles will be equal to the remaining angles respectively, namely those which the equal sides subtend. (Heath 1956, vol. 1, p. 155)

In the proof, two triangles *ABC* and *DEF* are produced meeting the specified conditions, so $AB = DE$, $AC = DF$ and angle *BAC* equals angle *EDF*. Then the triangle *ABC* is "applied to the triangle *DEF*" and it is proved that the points *B* and *E* coincide, as do the points *C* and *F*. Thus applying one triangle to another has the effect of copying a given angle *BAC* correctly in another place. This claim is stated and proved as a theorem much later, in I.23, which, however, builds on these earlier results. The application of figures, as used in I.4, caused numerous commentators to complain that if it is a valid principle then there was no need to expend such care on I.2, and the whole of geometry could be full of such proofs.[1]

The difficulty with I.4 may be traced back to Common Notion 4, where Euclid states that if two segments coincide then they are equal. Heath carefully explains that coincides means, in the Greek, an exact fit, and the assertion can be read as saying that if two segments can be made to coincide then they are equal. How then are segments made to coincide?

We can imagine at least two ways. One corresponds to a physical motion, and the other to copying the same set of instructions in different places. With Common Notion 4, and with Proposition I.4, the idea would seem to be that of a motion.

But the difficulties continue. In Proposition I.4 Euclid used the converse of Common Notion 4, that if two segments are equal then they can be made to coincide, which should have been made explicitly as an assumption, perhaps as part of Common Notion 4. Heath (1956, vol. 1, p. 225) commented here that modern editions of Euclid's *Elements* remark that when *B* coincides with *E*, and *C* with *F*, the lines *BC* and *EF* have been made to coincide, else two straight lines would enclose an area. Euclid never stated that this is impossible, and Heath speculates that either this remark is an interpolation of a later commentator or Common Notion 4 may be an interpolation.

We are not done with the problems in Proposition I.4, but we need now to look at the proffered definition of a straight line. It is well known that the concept of a straight line receives only a most unsatisfactory definition. Indeed, it has recently been suggested that the first seven definitions in the *Elements* may be a later interpolation by a scribe who was impressed by the careful definitions of many types of figure rushed into give definitions of the point, the straight line, and the plane that were better left unattempted (see Russo 1998). Be that as it may, a line is said to be "a breathless length", and a straight line to be a line "which lies evenly with the points on itself". This may help convince readers that they share a common conception of the straight line, but it is no use if unexpected difficulties arise in the creation of a theory, as we have already seen.

---

[1] See Heath's commentary at this point (1956, vol. 1, pp. 249–250).

The plane is also given an inadequate definition, closely akin to that of the line: "a plane surface is a surface which lies evenly with the straight lines on itself" (a surface "is that which has length and breadth only"). After that, the word 'plane' is not mentioned in the first four Books, although they are solely concerned with plane geometry. Solid geometry enters the Euclid's *Elements* in Book XI, which opens with three theorems that purport to show successively that a straight line cannot lie partly in a plane and partly not, that if two straight lines cut one another they lie in a plane and every triangle lies in a plane, and that if two planes meet then they do so in a line. But Euclid's definition of a plane is too weak to allow any of these arguments to count as a proof. Euclid would have been better off assuming them and thus de facto defining a plane. These 'theorems' do form a suitable basis for the results that follow: XI.4 there is a perpendicular to a plane at any point of the plane, and XI.5 all the lines perpendicular to a given line at a given point form a plane.

But properties of the plane have been used tacitly much earlier in the *Elements*. Indeed, I.4 is again problematic. To show that, when AC and DF are made to coincide and the points B and E lie on the same side of the $AC = DF$, it follows that the points B and E coincide it must be assumed that the triangles ABC and DEF lie in the same plane. A good definition of a plane is required, one that allows this result to be proved.

Problems with the *Elements* are not confined to Proposition I.4, or even with the confusions behind it. Nor, indeed, is the much more famous issue of the parallel postulate all that remains to elucidate. Readers who read the first book and tried to be sure what they could correctly say would be puzzled on a number of counts.

They would find that there is a limited vocabulary in the *Elements*, but that the meaning of the most basic terms is unclear. Surely some understanding of the space around us is being captured in this geometry: there is a largely undiscussed ability to bring some intuitions about lines and planes in space to a study of geometry, and some intuitions about lines and planes in space relate to a study of geometry in only two dimensions. However, there is no proof that the geometry on a plane is the geometry described in the earlier books in the *Elements*.

Points and straight line segments can be moved around in the mathematical plane and in mathematical space, or perhaps it would be better to say that they can be copied exactly in any position. But the concept of motion, in the sense of bringing figures into coincidence, is left unexamined.

There is a concept of equality of line segments, which, as we have seen, is not properly tied to the notion of bringing into coincidence. From it follows the definition of the circle: "a plane figure contained by one line such that all the straight lines falling upon it from one point among those lying within the figure are equal to one another", which point is the centre of the circle. Curiously, the definition of the sphere in Book XI is different in kind, and directly involves motion. It reads: "When the diameter of a semicircle remaining fixed, the semicircle is carried round and restored again to the same position from which it began to be moved, the figure so comprehended is the sphere".

However, the most basic and obscure figures (the straight line and plane) are not defined using the concept of distance. The most basic terms seem to be connected to straightness and flatness. Contrary to the case with what is called Euclidean geometry today, Euclid did not take distance as a fundamental concept. He did not define the straight line segment joining two points as the shortest curve joining them. Rather, in Proposition I.20 Euclid showed that "in any triangle two sides taken together in any manner are greater than the remaining one". This result has become known as the triangle inequality, and it goes a long way to proving that the line segment joining any two distinct points is the shortest curve through those points, although Euclid did not even hint at that consequence. It is also worth noting that there is no theorem in Euclid's *Elements* that depends on the actual size of a figure: any theorem that applies to one figure applies to all similar copies.

A plausible reading of *Elements* Book I is that a straight line can be understood as having a direction, so that at every point there is a straight line in every direction and only one straight line at a given point in a given direction. Thus a plane angle is defined as "the inclination to one another of two lines in a plane which meet one another and do not lie in a straight line", and an angle is called rectilineal when the lines containing the angle are straight. Proposition I.27 implies that if there are lines that cross a given line in equal angles then these lines will never meet, and so could be said to point in the same direction. Once the parallel postulate is assumed it follows there are such pairs of lines.

The earlier attempts at defining a straight line that are discussed by Heath (1956, vol. 1, pp. 165–169) also generally appeal to the idea of direction. Plato (in *Parmenides* 137 E) says "straight is whatever has the middle in front of both its ends", and Aristotle says something equivalent to this. Archimedes seems to have gone boldly for the alternative, that the straight line segment is the curve of shortest length between its endpoints, when he wrote at the start of *On the sphere and the cylinder* that "of all the lines which have the same extremities the straight line is the least". This may be the source for Proclus's definition of the straight line as the only curve that occupies a distance equal to that between the points on it, and also as a line stretched to the utmost.

However, interpreting the *Elements* in terms of direction must be regarded as an interpretation, and one that requires quite some work to make precise. For example, as Gauss pointed out (in a book review published in 1816, see *Werke* IV, 365) if one says that two lines are parallel if they cross a third in such a way as to make included angles with the third line that sum to π, then it must be proved that they cross any other line in the same way. Heath (1956, vol. 1, p. 194) regards such an interpretation as having been decisively refuted by Charles Dodgson in his *Euclid and his modern rivals* (1879).

As is well known, for whatever reason, Euclid held back from assuming the existence of parallel lines until Proposition I.29. Then, once the parallel postulate has been introduced Euclid showed in I.34 that opposite sides of a parallelogram are equal. This is as close as he got to saying that the distance between a pair of parallel lines is a constant.

We can conclude that three fundamental concepts are in play, without there being sufficient clarity concerning any of them. There is straightness (and flatness);

equality of line segments (the surrogate for distance); and the ability to bring certain figures into coincidence, and also to make arbitrary similar copies.

It must be admitted that these are difficult concepts to elucidate, still more so when it is not clear whether to read Euclid's *Elements* as a formal system or as an account of the space around us. Indeed, it is neither, but some mixture of both, and for as long as there was no reason to suspect that things could be otherwise, for as long as it was accepted that there was one space and it was described by the one geometry that can be devised, there was no good reason to force a distinction. As a result, mathematicians on occasion cheerfully ran these concepts together, and on other occasions struggled to keep them apart. For example, Clavius, in his edition of Euclid's *Elements*, simply writes "when two or more lines are said to be parallel, or equidistant, …" (first edition 1589, this quote from the 1607 edition).

In order to keep things tidy, we can make two ad hoc definitions. When straightness and flatness are taken as primitive concepts, equality of segments (perhaps in the sense of being brought into coincidence) is used rather than distance, and the language is reluctant to re-place statements in geometry with statements about numbers (say, in the form of coordinates) although coordinate geometry can be erected upon it, we shall speak of purely synthetic geometry. When distance is a primitive concept, line segments are said to have the same or different lengths, congruent figures to have corresponding sides equal in length, and geometrical trans-formations to preserve lengths we shall speak of a metrical geometry. We presently leave the idea of similarities in an ambiguous position.

Elementary geometry in the modern West moved in a confused way towards making distance the primary primitive concept, while often maintaining the Euclidean emphasis on straightness, thus frequently muddling the implications of the different concepts.

A notable example of this being nonetheless productive was Wallis's argument (see Wallis 1693) in defence of the parallel postulate, which rested, as he realised, on the ability to make arbitrary scale copies of a triangle. No-one had thought to doubt that similar, non-congruent figures exist, and it is to Wallis's credit that saw that he had not proved the parallel postulate but only established the equivalence of these two systems:

1. Euclid's *Elements*
2. Euclid's Elements with the parallel postulate removed and the assumption that arbitrary similar figures exist added.

It seems to be the first time that this equivalence was recognised.

In a mass of unpublished work expertly analysed by De Risi (2007), Leibniz made numerous attempts to present a coherent system of geometry that met his own, shifting, philosophical standards. Unlike Euclid, he started with three-dimensional space and a primitive relation of congruence of segments. The plane is then defined as all the points equidistant from two given points, and the straight line as the set of points equidistant from three given points. Somewhat later in life he produced what De Risi (2007, p. 220) called Leibniz's "finest ever definitions of a plane and a straight line". A point *Y* lies on the straight line through *A* and *B* if it is "unique in situation" with respect to *A* and *B*.[2]

Leibniz's definitions remained in his desk drawers, but others later adopted the first of these positions. Fourier, in a discussion with Monge, also took the concept of distance as fundamental, and began with three-dimensional space. He then defined successively the sphere, the plane (as the points equidistant from two given points) and the line (as the points equidistant from three given points).[3] Almost certainly independently, and to no avail, Lobachevskii did the same in his long presentation of non-Euclidean geometry (1835).

A more Archimedean position was adopted by d'Alembert in the *Encyclopedie Methodique* (1784, p. 538). There he defined geometry as the science that teaches us to know the extent, position, and solidity of bodies. A line (i.e. a curve of any kind) is one-dimensional, and the shortest line joining two points is the straight line. Parallel lines are lines that, however far they are extended will never meet because they are everywhere equidistant.

This put d'Alembert at the risk of over-defining. When a figure is characterised by possessing several properties, it may be that some parts of the definition imply others. In this case the characterisation is harmlessly over-defined. But if some parts of the definition contradict others then there is no figure with all these properties, and the over-definition is fatal. In the present case, d'Alembert needed to show that the curve everywhere equidistant from a straight line is itself straight, which he did not try to do, believing as he did that the principles of geometry are founded on truths so evident that it is not possible to contest them. Gauss, however, in a manuscript that may date from 1805 (in *Werke*, VIII, 163–164) observed that "The parallels to a straight line are those from which perpendiculars to the given line all have the same length. If the parallels are themselves straight remains undecided."

The approach to the difficulties in Euclid's *Elements* that was taken by Lambert (1786) is instructive. He noted the high level of Euclidean rigour, and that many trivial, indubitable propositions are proved with great attention to detail, while the parallel postulate is assumed, observing also that without the postulate much of Euclid's *Elements* lapses. The resolution of some of the problems, he suggested, was to admit that the difficulties in the *Elements* are insuperable unless one is allowed to suppose that the objects of geometry are representations. As he put it, "If we may neither see nor make a representation of the thing itself in Euclid's geometry, then there is a problem: To show that two lines do not enclose a space".[4] If, however, representation is admitted (Sect. 3), "one learns to know the thing itself (of which the axiom speaks) and also to add in thought that which seems to be missing in the axiom and in its representation, even if one cannot express this in

[2] See De Risi's analysis for what this means, but roughly speaking points not on the line cannot be unique in situation with respect to *A* and *B* because they cannot be distinguished from their mirror images in the line.

[3] See Bonola (1906, 54) who cites *Seance de l'Ecole Normale*, 1, pp. 28–33, reprinted in *Mathesis* 9, pp. 139–141 (1883).

[4] Translation from Ewald (1996 vol. 1, 159).

words.” In this way, Lambert continued, one learns that one line may not approach another asymptotically, nor two straight lines enclose an area, not because these results can be proved but because they make claims false to the representation of the thing itself. So, by presupposing the representation of the actual thing, and not demanding only words, Euclid’s procedure can be justified.

However, Lambert went on (Sect. 10), the problem with the parallel postulate concerns neither the truth nor the think ability of the axiom. We learn to think about straight lines by under-standing how they are used in Euclid’s Elements, and the parallel postulate is plainly regarded as true, as the evidence of its consequences shows. Therefore, the task is to derive the parallel postulate from the other assumptions of the Elements, or, if that fails, to find other equally evident postulates that do imply the parallel postulate. Now one must cease to appeal to representations of the things themselves, but work entirely in words, as Euclid did in making his postulates, and “the proof should be carried out entirely symbolically—when this is possible” (in Ewald 1996, vol. 1, p. 166). That is to say, in terms of synthetic geometry.

In this case the approach put forward by Wolff, who defined parallel lines as equidistant, fails precisely because he forgot that “arbitrarily conjoined concepts must be established” (in Ewald 1996, vol. 1, p. 161).

Lambert ultimately abandoned his attempts to do give an entirely symbolic defence of Euclid’s Elements, after establishing several novel theorems in a geometry in which the parallel postulate was replaced by the assumption that the angle sum of every triangle is less than two right angles. His *Theorie der Parallellinien* was published posthumously in 1786.

Legendre was a mathematician sympathetic to the didactic aims of the Elements but not to its original formulations. He wrote several different versions of his *Éléments de géométrie* (1794) with a view to restoring Euclidean rigour in the teaching of geometry, which in his view had been corroded by texts, such as one by Clairaut, that relied on motions of self-evidence. They differ largely, as he had to admit, in their unsuccessful attempts to deduce the parallel postulate. Its chief significance for present purposes is that it exemplifies the attempt to ground elementary geometry on a concept of distance, or rather, and more precisely, on the idea that a straight line is the curve of shortest distance between any of its points.

In all these editions Legendre took a firmly metrical point of view. His opening definition of the first edition proclaimed that “Geometry is a science that has as its object the measure of extent”. Extent, he explained, has three dimensions, length breadth, and height; a line is a length without breadth, its extremities are called points and a point therefore has no extent. A straight line is the shortest path from one point to another; surfaces have length and breadth but no height or depth; and a plane is a surface in which if two arbitrary points are joined by a straight line this line lies entirely in the surface. Distance itself is not defined.

Legendre then set out to prove the theorems of the Elements together with some results Euclid had preferred to assume, such as (Legendre’s first result): any two right angles are equal. His Theorem 3 proved that the line joining two distinct points is unique (its existence having been tacitly assumed to be a consequence of the definition of a straight line). Familiar congruence theorems follow in each edition until the parallel postulate could no longer be ignored. Once the existence of parallel lines was assured Legendre showed that they were equidistant. In fact, Legendre's attempts to restore rigour to the treatment of elementary geometry was no better than Euclid's, and in some ways worse, not only because his attempts to prove the parallel postulate inevitably failed, but because he smuggled more into his account than he realised.

Historians of geometry remember Gauss for many reasons, not only for his investigations of the parallel postulate and his ideas about non-Euclidean geometry (see Gray 2011). But Gauss also continued his investigations into the proper foundations of Euclidean geometry for many years. He kept a mathematical diary, in which he recorded new ideas and results.[5] One entry (number 72, from 28 July 1797) records "I have demonstrated the possibility of the plane." How he did this, and whether his demonstration continued to satisfy him we do not know, but thirty years later he wrote to Bessel about problems in the foundations of geometry and commented that, apart from the well-known problem with the parallel postulate "… there is another omission that to my knowledge no-one has criticised and will in no way be easy (although possible) to put right. This is the definition of the plane as a surface in which a straight line joining any two points in the surface lies in it entirely. This definition contains more than is necessary for the determination of the surface, and tacitly involves a theorem that must first be proved…" (*Werke* VIII, 200). Then, in pages that the editors of Gauss's *Werke* date to March 1831, and in letters to Schumacher in May 1831, Gauss set out some of his ideas. They are synthetic in character, and inconclusive.

L.A. Sohnke, writing in the *Allgemeine Encyclopaedie der Künste und Wissenschaften* on the term 'Parallel' in 1837 found no less than 91 attempts on the theory of parallels, which he surveyed by dividing them into those where 'parallel' means never meeting; where 'parallel' means equidistant; and where 'parallel' means cross a third line in equal angles. A reasonable conclusion for us to draw from all this unsuccessful activity is that metrical geometry needed to put its house in order, and it probably could not do so taking a strictly synthetic route, nor by grafting the concept of distance onto a structure modelled on Euclid's *Elements*. This is an awkward position for traditional geometry to be in, and it may have opened people's minds to the possibilities of alternatives. Certainly, two were to be produced. One, projective geometry, amplified and improved the synthetic side of geometry. The other, non-Euclidean geometry, was a new and challenging metrical geometry.

This note is not the place to explore how these new geometries may have come about.[6] But it interesting to note that there was an abundance of theorems in Euclid's Elements, even in Legendre's presentations, that drew on the straightness of the line, and Poncelet's ideas, as presented in his book (1822) stressed the

[5] In Gauss *Werke*, X.1, 483–574. There is an English translation in Dunnington (2004, 469–496).

[6] The non-Euclidean story is much better known, see e.g. (Gray 2011); the history of projective geometry needs to be written, and a start has been made in (Bioesmat-Martagon 2011).

incidence properties of lines (two lines meet in a point, two points determine a line) in creating what he provocatively called a non-metrical geometry. Equally, in their publications of the 1830s, Bolyai and Lobachevskii emphasised the metrical aspects of geometry, and indeed the capacity to express geometrical results in the language of novel trigonometrical formulae.

References

Bioesmat-Martagon, L. (2011). *Éléments d'une biographie de l'espace projectif* (Vol. 2). Nancy: Presses Universitaires de Nancy, Collection histories de geometries.

Bonola, R. (1906). *La geometria non-Euclidea*. (H. S. Carslaw, English Trans.), preface by F. Enriques. *History of non-Euclidean geometry.* Chicago: Open Court, 1912; reprint New York: Dover 1955.

Clavius, C. (1607). *Euclidis Elementorum Libri XV.* Frankfurt: Rhodius.

d'Alembert, J. (1784). *Encyclopédie Méthodique: Mathématique.* Paris: Panckoucke.

De Risi, V. (2007). *Geometry and Monadology: Leibniz's analysis situs and philosophy of space*. Basel: Birkhäuser.

Dodgson, C. (1879). *Euclid and his modern rivals*. London: Macmillan.

Dunnington, G. W. (2004). Introduction and appendices. In J. J. Gray (Ed.), *Gauss: Titan of science*. Washington: Mathematical Association of America.

Ewald, W. (1996). *From Kant to Hilbert: A source book in the foundations of mathematics* (2 Vols.) Oxford: Clarendon Press.

Gauss, C. F. (Ed.). [Werke], *Werke*. vols. IV (1880), VIII (1900), X.1 (1912). Königlichen Gesellschaft der Wissenschaften zu Göttingen. Göttingen: Dieterich; later Berlin: Springer, Leipzig: Teubner.

Gray, J. J. (2011). *Worlds out of nothing; a course on the history of geometry in the 19th century* (2nd ed.). London: Springer.

Heath Sir T. L. (1956). *Euclid's Elements* (3 Vols.) Cambridge: Cambridge University Press, reprint New York: Dover.

Lambert, J. H. (1786). Theorie der Parallellinien. In F. Engel & P. Stäckel (Eds.), Theorie der Parallellinien von Euklid bis auf Gauss (1895). Leipzig: Teubner.

Legendre, A.-M. (1794). *Éléments de géométrie.* Paris: F. Didot, several editions.

Lobachevskii, N. I. (1835). New Elements of Geometry, with a Complete Theory of Parallels (Russian). *Gelehrten Schriften der Universität Kasan*. German translation in Lobachetschefskij *Zwei geometrische Abhandlungen*, (F. Engel, Trans.), Leipzig: Teubner 1898.

Poncelet, M. (1822). *Traité des propriétés projectives des figures*. Paris: Bachelier.

Russo, L. (1998). The definitions of fundamental geometric entities contained in book I of Euclid's. *Elements, Archive for History of Exact Sciences, 52*(3), 195–219.

Wallis, J. (1693). De postulato quinto et definitione lib. 6 Euclidis deceptatio geometrica. In: *Operum Mathematicorum*, *2*, (pp. 665–678). Oxford: Oxford University Press.

Theon of Smyrna and Ptolemy on Celestial Modelling in Two and Three Dimensions

Alexander Jones

Ptolemy, as is well known, devoted the great part of his *Almagest* (properly, *Mathematical Composition*) to modelling the motions of the Sun, Moon, and planets by means of combinations of circular motions that are represented in his geometrical analyses by two-dimensional assemblages of circles, straight lines, and points. It is a little less well known that his later work, the *Planetary Hypotheses*, proposed systems of three-dimensional bodies, composed of ether (αἰθήρ, the Aristotelian fifth, celestial element) and having spherical and planar surfaces, that were supposed to be the physical realities performing the various revolutions that the *Almagest*'s circles describe geometrically.

The three-dimensional, physical cosmology of the *Planetary Hypotheses* was not an abrupt innovation that Ptolemy imposed after the fact on the geometrical theories he worked out the *Almagest*. Its foundations include assumptions expressed, albeit sometimes briefly and in passing, in the *Almagest* and in works that Ptolemy composed before the *Almagest*, as well as ideas that were already current in Platonist and Peripatetic philosophy before Ptolemy's time, for which we have a valuable witness in the one surviving work by the Platonist philosopher Theon of Smyrna. This paper attempts to delineate some of this background and to show what elements were indeed new in the *Planetary Hypotheses*.

A. Jones (✉)
Institute for the Study of the Ancient World, New York University,
15 East 84th Street, New York 10028, USA
e-mail: alexander.jones@nyu.edu

V. De Risi (ed.), *Mathematizing Space*, Trends in the History of Science,
DOI 10.1007/978-3-319-12102-4_4

1 Theon of Smyrna

Theon of Smyrna "the Platonist," the author of an extant (though incomplete) book *On the Mathematics Useful for Reading Plato* in addition to several other works of Platonic scholarship that have not come down to us, is almost certainly identical to a "Theon the Platonist Philosopher" whose portrait bust, purchased in Smyrna in the 17th century, is now in the Capitoline Museum.[1] The inscription on the bust's base identifies not only the subject but also the person who commissioned it, Theon "the Priest," who was the philosopher's son. Technical as well as stylistic grounds securely date the portrait to the later part of Hadrian's reign, in the 120s or 130s of our era, making Theon an older contemporary of Ptolemy, though we have no reason to believe that Ptolemy was acquainted with Theon or his writings.[2]

Theon's book consists of several sections on various branches of mathematics (broadly conceived), of which the one on astronomy is the longest.[3] The title accurately indicates Theon's purpose and intended readership: philosophical students who had a limited mathematical education—chiefly the earlier parts of the *Elements* (H16)—and who could benefit from some background on topics such as properties of whole numbers, harmonics, and astronomy that are alluded to in the dialogues of Plato that were widely read at this period, among which the *Timaeus* and *Republic* were prominent. Theon acknowledges a heavy dependence on earlier works, mostly by philosophical writers in the Platonist or Peripatetic traditions rather than mathematicians or scientists.[4] Most of the astronomical section, according to his own repeated statements, came from a certain Adrastos, whom we know from other authors who refer to him as a Peripatetic from Aphrodisias. In his commentary on Ptolemy's *Harmonics* (ed. Düring 1932, 96) Porphyrios (3rd century AD) quotes as from Adrastos's *Commentary on the Timaeus* a passage that we find *verbatim* in Theon (H50–51), and this was presumably the source work for the rest of Theon's Adrastian material.[5] There are also extensive and very close parallels in Calcidius's Latin *Timaeus* commentary (4th century AD) to passages in the part of Theon dependent on Adrastos. Calcidius does not acknowledge his

[1] Musei Capitolini inv. 529. Richter (1965, 3.285).

[2] For more detailed discussion of the evidence for Theon's biography see Jones (forthcoming). The "Theon the mathematician" from whom Ptolemy obtained reports of observations of Mercury and Venus at greatest elongation (*Almagest* 9.9 and 10.1–2) can hardly have been Theon of Smyrna, as has often been suggested. No astronomer of the second century AD could have been unaware that the planets exhibit two anomalies or could have believed that their stationary points occur when they are at greatest elongation from their mean longitudes.

[3] References to passages in Theon's book will be by the pages of Hiller's edition, Hiller (1878).

[4] Hence the absence of references to Ptolemy or of knowledge of the *Almagest*'s contents can hardly stand an argument that Theon's book was written before the *Almagest* (completed soon after AD 146/147, the date of Ptolemy's *Canobic Inscription*). The *only* good evidence for Theon's date is the bust, which portrays him as a man in his prime, so if he was still living when it was made, he might still have been living and writing in the 150s.

[5] Earlier Porphyrios (ed. Düring 1932, 7) quoted another passage from Adrastos, without specifying the work's title, that is matched *verbatim* in Theon (H50).

source; but if he was using Theon, it is hard to explain why he avoided drawing on anything that Theon identifies as his own contribution or as coming from source texts other than Adrastos, so it seems most probable that Calcidius had access either directly to Adrastos's book or to a digest independent of Theon's. If this is correct, Theon hewed closely to Adrastos in wording as well as content.

Adrastos is usually assigned to the period immediately preceding or even contemporary with Theon, but while this is certainly possible, nothing in his reported writing is incompatible with a date anywhere within the first century AD or even the first century BC.[6] The latest earlier author to whose work he is known to have referred is Hipparchos (fl. 160s–120s BC). The hypotheses for the Sun's motion that he discussed are versions of the one that Ptolemy attributes to Hipparchos (though Adrastos, at least as we read him through the filters of Theon and Calcidius, does not expressly associate them with Hipparchos), and he apparently believed that the planets, like the Sun, exhibited only the single periodic anomaly manifested by their synodic cycles. Ptolemy (*Almagest* 9.2) tells us that Hipparchos had written a book showing that hypotheses for the planets that assumed only the synodic anomaly were inconsistent with observations. It does not follow, of course, that the publication of such a book would have led all astronomers, let alone philosophers, to instantly abandon single-anomaly theories; nevertheless, the closer one gets to Ptolemy's time, the less likely it would be that an author with up-to-date knowledge of the astronomical literature would have been unaware of the planets' zodiacal anomaly and the consequent need for hypotheses more complicated than Adrastos's and Theon's simple epicycles or eccenters.[7] We know that such hypotheses were circulating at least as early as the mid first century AD, because Pliny (d. AD 79), in his confused discussion of the planets (2.13/63–64), reveals an awareness, if not an understanding, of hypotheses that involved an eccentricity to account for zodiacal anomaly. If Adrastos lived as late as is generally supposed, therefore, he was quite out of touch with contemporary astronomy, something that can be said about Theon without reservation.

A point of interest concerning Adrastos is that a philosopher who identified himself as a Peripatetic in this period (i.e. anywhere between the 1st century BC and the early 2nd century AD, given the uncertainty about his date), and whose other known writings were devoted to Aristotelian topics,[8] would have composed a didactic commentary on the *Timaeus*, and moreover one on which a subsequent

---

[6] See for example Moraux (1984, pp. 294–295), and (Sharples 2010, p. 22), neither of whom give any reason for their confidence that Adrastos wrote in the first half of the 2nd century AD except for the *terminus ante quem* provided by Theon.

[7] The resistance of astronomers and astrologers, even after Ptolemy's time, to accept Hipparchos's discovery of precession is a striking illustration of how scientific arguments that may appear cogent to us (reading them as they are presented in Ptolemy's *Almagest*, since the relevant works by Hipparchos have not survived) could be disregarded in the contemporary scientific community; see Jones (2009). The inadequacy of single-anomaly planetary models would have been harder to ignore, however, especially after the Babylonian predictive models for planetary phenomena had become prevalent in the Greek world.

[8] Moraux (1984, pp. 314–330).

Platonist philosopher would rely for so much of his material. This was symptomatic of the closeness that existed between Middle Platonism and contemporary Peripateticism; inasmuch as Theon is our most abundant informant for the approach to astronomical cosmology that supplied the building blocks for Ptolemy's, we cannot label this approach exclusively as belonging to one of these schools or the other.

Because Theon's *The Mathematics Useful for Reading Plato* is a comparatively neglected text, I will summarize his treatment of celestial modelling in some detail, in addition to providing in an appendix an English translation of the passage that is most central to this paper's concerns.[9]

The section on astronomy in Theon's book (H120–205) begins with material, drawn from Adrastos, that is similar in scope to *Almagest* 1.2–8 (H120–133): the spherical shape and revolution of the heavens and the spherical shape, central location, and minuscule relative size of the Earth, supported by mostly empirical arguments; definitions of the celestial equator, tropic circles, and arctic circles; and explanation of the inclined zodiacal belt as the apparent pathway of the Sun, Moon, and planets. After a general discussion of the planets and their motion and phenomena (H134–147), Theon takes up the topic of their apparent anomalies.

The first portion of this passage (H147–152) is ascribed through repeated citations to Adrastos. It begins with the phenomena that, while the Sun and Moon are always seen to move eastward relative to the fixed stars, the planets exhibit apparent reversals of direction. The reality, however, is that the cosmos consists effectively of three tiers of entities in motion: the outermost, which moves with uniform revolution concentric with the cosmos; below this, the circular but *apparently* nonuniform motion of the Sun, Moon, and planets, and innermost, the truly irregular motion of bodies that experience generation and corruption. The entities of the outer tiers are described as divine and eternal. Physical necessity dictates that the other heavenly bodies must, like the fixed stars, move in a simple, uniform, and orderly manner, where Adrastos defines uniform motion as travelling equal intervals in equal times and orderly motion as motion without stoppings and reversals of direction. Nevertheless all the heavenly bodies (other than the fixed stars) exhibit apparent nonuniformity while some (the five planets) also exhibit disorderly motion, i.e. changes of apparent direction. Two causes are given for the appearance of nonuniformity, firstly that the bodies travel on circles that are distinct from (i.e. not concentric with) the zodiac that they are seen as traversing, and secondly that they travel on combinations of more than one circular path.

To illustrate this general explanation (H152–166), Theon cites the phenomenon that the intervals between the equinoxes and solstices are observed as being unequal in duration, which means that the Sun appears to traverse the four equal quadrants of the zodiacal circle in unequal times. Two hypotheses are shown to be able to account for this: the eccentric hypothesis, according to which the Sun travels uniformly on a circle that is not concentric with the cosmos, and the epicyclic

[9] Theon's book has twice been translated into French (Dupuis 1892 and Delattre Biencourt 2010). I am not aware of any complete translations into other modern languages except for the English version of Lawlor and Lawlor (1979), which is derived from Dupuis's French, not Theon's Greek.

hypothesis, according to which the Sun travels uniformly on an epicyclic circle that itself travels uniformly on a circle concentric with the cosmos.[10] Theon castigates the "mathematicians" (μαθηματικοί) for disagreeing among themselves whether the motions of the heavenly bodies are produced only by epicycles on deferents or only by eccenters, since, he says, it will be shown subsequently that the heavenly bodies trace (γράφοντες) all three kinds of circle as nonessential consequences (κατὰ συμβεβηκός). (It will only later become clear that Theon means this in two ways, firstly that each hypothesis, taken as the valid one, generates the circular paths of the other hypothesis incidentally, and secondly that there exists a third hypothesis, the valid one, that incidentally generates the paths of both the epicyclic and eccentric hypotheses.) Using diagrams similar in kind to those of the *Almagest*, Theon demonstrates how each of these hypotheses "save the phenomena," i.e. lead, under appropriate conditions, to the observed inequality of the seasons.

Theon now (H166–172) raises the question (which he tells us Hipparchos considered worthy of mathematical understanding) of how it is that two seemingly very different hypotheses can lead to the same phenomena.[11] In response, he tells the reader that Adrastos demonstrated that the eccentric hypothesis arises as a nonessential consequence of the epicyclic, and vice versa.[12] In fact Theon gives two demonstrations of the former claim, one applying just to the situations where the center of the epicycle is situated on the apsidal line or on the line through the center of the cosmos and perpendicular to the apsidal line, while the second demonstration is for the general situation. The demonstration of the converse is also general.

These demonstrations, Theon continues (H172–174), can be extended to the Moon and planets, with the qualification that whereas the Sun has *almost* identical periods of revolution in longitude, latitude, and anomaly so that a model with a fixed eccenter or one with the same periods of revolution for the epicycle eastward around the center of the cosmos and of the Sun around its epicycle (in the opposite sense of revolution) suffice, whereas these periods are significantly different from each other for each of the remaining bodies. This diversity of periods, which in the case of the five planets is manifested in their retrogradations, has to be reflected in

---

[10] Theon cites the specific intervals 94 1/2 days from vernal equinox to summer solstice, 92 1/2 days from summer solstice to autumnal equinox, 88 1/8 days from autumnal equinox to winter solstice, and 90 1/8 days from winter solstice to vernal equinox; the first two of these are the ostensibly observed intervals from which Ptolemy derives his eccenter model, which he attributes to Hipparchos (*Almagest* 3.4). Theon's eccenter model also has identical parameters to Ptolemy's, with eccentricity 1/24 of the eccenter's radius in the direction of Gemini 5 1/2°, and his epicyclic model's parameters are equivalent. Subsequently (H188), Theon will claim that Hipparchos expressed a preference for the epicyclic model, which is not something that one can infer from anything Ptolemy says.

[11] Tannery (1893, pp. 60–61), followed later by van der Waerden (1988, p. 180), tendentiously interpreted this passage as stating that Hipparchos admitted to not understanding the kinematic equivalence of the epicyclic and eccentric hypotheses, and on this basis both maintained, perversely, that Hipparchos could not have been a competent mathematician.

[12] The construction of Theon's sentence implies that Adrastos gave both the derivation of the eccentric hypothesis from the epicyclic and that of the epicyclic from the eccentric; the parenthetic phrase "as I say" (ὡς δὲ ἐγώ φημι) is probably not an indication that Theon has added the second demonstration in his own right but just a reminder that the subject was mentioned earlier.

the model; thus in an epicyclic model the period of the planet's revolution around the epicycle is in some cases faster and in other cases slower than that of the epicycle around the center of the cosmos. Using simple *Almagest*-style diagrams, Theon stipulates (H175–177) that for an epicyclic model to "save the phenomena," the direction of revolution of the body around the epicycle must be opposite to that of the epicycle's own eastward revolution, but for the planets both revolutions are in the same (eastward) direction; and again, for an eccentric model, the center of the eccenter must revolve in the eastward direction on a circle that is concentric with the cosmos and equal in size to the epicycle of the alternative model (presuming that the eccenter is the same size as the deferent of the other model), while the eccenter revolves in an unspecified direction around its center, carrying the body with it.

In bringing this passage to a close, Theon tells us that he has been following Adrastos's account of how the hypotheses of the "mathematicians" can be reconciled (προσοικειῶσαι) with each other. This repeated mention of the "mathematicians" marks the transition to the crucial passage (H177–189) translated in the appendix of this paper. There are no further nods to Adrastos in the passage. However, parts of Theon's discussion of Aristotle's cosmology (H178–179) are paralleled in Calcidius (ed. Waszink 1962, 135–136) so that this much, at least, probably comes from Adrastos, and since this includes a compact outline of the principles of Theon's "physical" hypotheses for the heavenly bodies later in the passage, it appears that Theon was at most working out a specific implementation of an approach that was already known in Peripatetic (and presumably Platonist) cosmology. It may in fact, as Evans has suggested, be the kind of planetary modelling that Geminos (mid 1st century BC) refers to as σφαιροποιία, "sphere-assemblage," though Geminos provides no details.[13]

Theon likens the "mathematicians" to the Babylonian, Chaldean, and Egyptian astronomers of old, who, he alleges, fitted mathematical methods to the phenomena with a view only to prediction. An allusion to the *Epinomis* (which Theon regards as an authentic work of Plato's) highlights the pejorative implication of this association of the "mathematicians" with non-Greeks, since it is meant to call to mind "Plato's" remark (987d) that "whatever Greeks take over from barbarians is in the end turned by them into something finer." In this instance what the Greeks—but not the "mathematicians" among them—have contributed is reasoning according to nature (φυσιολογία), that is, explanation in terms of physical causes.[14]

---

[13] Evans (2003). Evans further speculates that the three-dimensional physical approach to modelling was associated with the epicyclic and eccentric hypotheses already at the time of their introduction (which he attributes to Apollonios of Perge); but if that was the case, it is hard to understand why both Theon and Geminos (see next note) assert that mathematical astronomers took no account of physical causation.

[14] Similarly, Geminos, in a passage from his lost digest of Posidonios's *Meteorology* quoted by Simplikios (*Commentary on Aristotle's Physics* 291–292) by way of Alexander of Aphrodisias, contrasts the methods used to address celestial phenomena in astronomy (ἀστρολογία), which are mathematical, and those used in φυσιολογία, with a certain bias in favor of the latter.

Aristotle is made to supply these physical causes by way of his adaptation of the systems of “homocentric” spheres of Eudoxos and Kallippos as outlined in *Metaphysics* Λ and *De Caelo*. These spheres, being three-dimensional bodies composed of the Aristotelian “fifth body” or ether, are an appropriate cause for the motions of the heavenly bodies, as the circles and more complicated “spiral” curves of the “mathematicians” are not. According to Theon (in one of the presumed Adrastian sentences that also appear in Calcidius), not all of Aristotle’s etherial spheres were shells concentric with the cosmos, but some were solid spheres embedded in the interval between the interior and exterior surfaces of spherical shells (τὰς ἐν τ βάθει τούτων [scil. τῶν κοίλων σφαιρῶν]… στερεάς), which is, to put it mildly, a bold reading of Aristotle’s cosmology. It turns out presently that this is how Theon, or Adrastos, interprets Aristotle’s “unwinding spheres”: they are lodged like ball bearings or (in Theon’s remarkable simile) gears in the intervals between revolving shells to effect the transfers and reversals of their spinning motions.

Having thus procured Aristotle’s blessing for the principle that the heavens can be composed of an assemblage of spherical shells concentric with the cosmos and of solid spheres embedded in the thickness of the shells, Theon shows how a hypothesis can be formed out of such bodies that results in the path for a heavenly body identical to the path generated by an epicyclic or eccentric hypothesis. It is essentially a three-dimensional filling out of an epicyclic model (Fig. 1, showing the epicycle in four positions), though the diagram illustrating it in Theon’s text (Fig. 5) is necessarily just a cross-section of the hypothesis in the plane of the orbit. The planar circle of the epicyclic model becomes a great circle of a solid sphere that revolves on an axis perpendicular to the plane of the epicycle. The visible heavenly body is fixed on the surface of this solid sphere. The solid sphere in turn is assumed

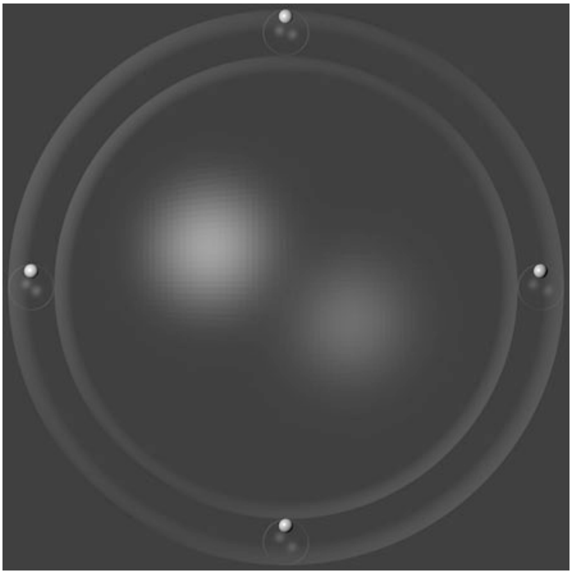


Fig. 1 Theon’s three-dimensional epicyclic hypothesis for the Sun

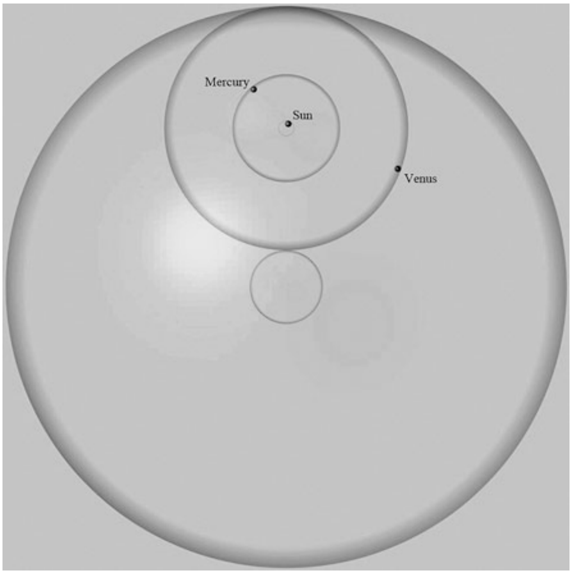


Fig. 2 Theon's nested epicyclic hypothesis for the Sun, Mercury, and Venus

to be embedded in a spherical shell concentric with the cosmos, such that the exterior and interior surfaces of the shell are tangent to the solid sphere. The shell revolves on an axis perpendicular to the plane of the deferent, so that the center of the solid sphere traces out the deferent, while the heavenly body traces out the epicycle; or, taking into consideration the combined motions of the solid sphere and the shell, the heavenly body traces out the eccenter.

Theon further points out that a common hypothesis of this kind could serve for the Sun, Mercury, and Venus; in this case, the single solid sphere tangent to the exterior and interior surfaces of the revolving shell is replaced by a spherical shell bearing Venus, enclosing a second shell bearing Mercury, which in turn encloses a solid sphere bearing the Sun (Fig. 2). It is noteworthy that Theon, while giving this model only as a possible alternative to having separate hypotheses for each body, expresses a "suspicion" that the combined hypothesis is correct, on the grounds that it would provide a suitable seat for the cosmic soul; notwithstanding the apparently mechanical conception of how revolving motions are imparted to the spheres that was ascribed to Aristotle, Theon believes that the cosmos is a single living being.

Theon's cosmological system is, in comparison to the one Ptolemy was to devise, a rather modest affair. The Sun, Moon, and planets are each assumed to have just a single periodic anomaly, and the building blocks of Theon's hypotheses, spherical "deferent" shells concentric with the cosmos in combination with solid "epicyclic" spheres or concentric shells lodged in the thickness of the shells, only suffice to model a body with a single anomaly. Although his reading of such embedded spheres into Aristotle seems anachronistic and perverse, the approach does make sense as a way of preserving a significant part of the Eudoxian cosmology while making it possible to explain anomaly as a function of motion in "depth," i.e. by means of a kind of epicyclic motion. Theon has a fairly precise conception of how to build a viable quantified model for the Sun, but his accounts of how to extend such models to the Moon and planets suffer from vagueness with respect to the dimensions and rates of revolution of the components as well as how latitudinal motion is to be effected. A technically oriented "mathematician" astronomer of his time would likely have dismissed the whole enterprise as simplistic.

2 Ptolemy's Planetary Hypotheses

The *Planetary Hypotheses* consists of several sections.[15] First (1A.1–2/BM81b–82a), Ptolemy gives a brief introduction explaining the work's purpose and relationship to the *Almagest*. The next section (1A.3–14/BM82a–88a), accounting for a little over half of Book 1, is a technical description of Ptolemy's "hypotheses" or models for the Sun, Moon, and planets as consisting of combinations of circles. This is followed by a brief section (1B.1/BM88a–88b) reviewing in a qualitative manner the various kinds of north-south motion exhibited by each of the heavenly bodies according to their models. The fourth, longer, section (1B.2–5/BM88a–92a) takes up the topic of the sizes and distances of the heavenly bodies. A discussion of the visibility conditions of the stars and planets and visual misjudgment of their apparent sizes (1B.6–7/BM92a–92b) brings Book 1 to a close. The first section of Book 2 (2.1–10/BM93a–96b) is devoted to general considerations with respect to the nature, shape, and motion of the etherial bodies that make up the systems for the heavenly bodies, followed by a detailed description (2.11–16/BM96b–101b) of the configuration of each system.[16] Ptolemy gives alternative versions for the systems of the Sun, Moon, and planets, in one of which the bodies are all bounded by complete spherical surfaces, while in the other some of the bodies are truncated by planes perpendicular to the bodies' axes of revolution, so that they take the form of

[15] Only a part of Book 1, from the beginning to close to the end of the description of the models (ending at ed. Heiberg 104 line 23 "ἰσοταχῶς"), survives in Greek; the completion of this section that appears in some manuscripts is a later restoration. Heiberg (1907, pp. 70–107) presents the Greek text facing L. Nix's German translation of the corresponding part of the Arabic version. We will use "1A" to designate this first part of Book 1. Nix's translation of the Arabic version of Book 2 follows on pp. 111–145. Morelon (1993) has edited and translated into French the Arabic version of Book 1, and Goldstein (1967, pp. 13–55) provides a facsimile of the manuscript BM Arab. 426 with an apparatus including collation of Leid. Arab. 1155; for his English translation of the part of Book 1 not extant in Greek ("1B") see pp. 5–9. Unfortunately the patchwork state of publication of the *Planetary Hypotheses* makes a consistent and convenient system of reference difficult. We will refer to passages in two ways: first by the numbers of the sections into which the text has been divided by Heiberg and Goldstein (unfortunately in his translation only, and numbered starting with 1 instead of continuing Heiberg's sequence), and secondly by the folios of BM Arab. 426, which are indicated in Goldstein's facsimile and the margins of Morelon's text. Thus "1A.2/BM81b" means Sect. 2 of the first part of Book 1 (appearing on pp. 72–73 in Heiberg 1907) and f. 81^b of BM Arab. 426 (appearing on p. 13 of Goldstein 1974 and pp. 16–17 of Morelon 1993).

[16] In the absence of a critical edition of Book 2, the analysis of this part by Murschel (1995) is indispensable.

disks or rings. Then a short section (2.17/BM101b–102a) concerns the number of distinct etherial bodies in Ptolemy's entire cosmic system, and in conclusion (2.18/BM102a–102b) Ptolemy explains how a "table-top" model of his system could be made to show the agreement of the theories with observation, with the components set in appropriate positions according to a set of mean motion tables appended to the end of the *Planetary Hypotheses* (the tables have not survived).

The somewhat disjointed structure of the *Planetary Hypotheses* is at least in part a reflection of the fact that Ptolemy did not compose the work with a single category of reader in mind. In his preface (1A.1/BM81b) he identifies his intended readership as "ourselves and those who choose to set these things [*scil.* the celestial models] in an instrumental construction" (ὑπό τε ἡμῶν αὐτῶν καὶ τῶν εἰς ὀργανοποιίαν ἐκτάσσειν αὐτὰ προαιρουμένων). Whom does he mean by "ourselves"? Ptolemy's addressee is the same Syros, otherwise unknown to us, to whom he dedicated most of his astronomical writings, including the *Almagest*, as well as the astrological *Tetrabiblos* (but not the *Criterion*, *Harmonics*, or *Geography*). Syros must therefore have been adept in mathematics and mathematical astronomy; but it is not primarily in this capacity that Ptolemy seems to be speaking of "ourselves" in the *Planetary Hypotheses* but rather as philosophers interested in the cosmological issues that flowed chiefly from Plato and—especially—Aristotle.[17]

Thus the section in Book 1 in which Ptolemy singles out for special attention the north-south aspect of the apparent motions of the heavenly bodies, and differentiates between the kinds of north-south motion that are due respectively to the obliquity of the ecliptic, the inclination of the eccenter, and the inclination of the epicycle, seems pointless from either an astronomical or a mechanical point of view. Its ultimate motivation is in Aristotle, *On Generation and Corruption* 2.10, where the annual north-south motion of the Sun caused by the obliquity of the ecliptic is characterized as alternate approach and recession, and assigned the key role of instigating the cycles of generation and corruption in the sublunary part of the cosmos. Theon (H148–149) shows how this had come to be generalized to apply to the other heavenly bodies: "For the sake of the numbering of time and the transformation of the things near the Earth and far from the Earth, the travel of the wandering (stars) came to be; for the things here (below the heavens) also transform in all ways together with their (scil. the heavenly bodies') turnings (τροπαί) as they approach and recede." Extension of the term τροπαί, which normally refers to solstices, to the reversals of north-south motion of the Moon and planets meant that their latitudinal cycles were also regarded as significant in maintaining the

[17] Ptolemy's philosophical positions, though influenced in some respects by the other Hellenistic sects, are closest to Middle Platonism and contemporary Peripateticism (Feke and Jones 2010). In some outstanding recent discussions of the *Planetary Hypotheses* (Murschel 1995, p. 36; Swerdlow 2005, p. 66) the significance of the "ourselves" in the preface to the *Planetary Hypotheses* is passed over, resulting (in my opinion) in an overemphasis on the construction of instruments to simulate and represent the celestial bodies as if this was the primary or sole purpose of the work.

processes of change in the sublunary world; hence Ptolemy's concern with analysing the more complex patterns of north-south motion into their constituents.

In Book 2, Ptolemy's presentation of the three-dimensional physical systems for the heavenly bodies is intermixed with extended criticisms of the cosmology of Aristotle's *Metaphysics* Λ.8. Again from the point of view of the mathematical astronomer this appears to be a strange preoccupation with a long abandoned theory, one to which Ptolemy had not made any allusion in the *Almagest*; but we have seen from the example of Theon that for a Platonist or Peripatetic of Ptolemy's time the introduction of eccentric and epicyclic motion into the apparatus of astronomical modelling had not made Aristotle's, or indeed Plato's, cosmological pronouncements irrelevant. Theon diminishes the conceptual interval between the fourth century BC cosmologies and the not-quite-as-out-of-date simple eccentric and epicyclic models that he describes by attributing opinions to Plato and Aristotle that we would consider anachronistic: the interpretation of Aristotle's "unwinding spheres" as epicyclic spheres rolling between Eudoxos's homocentric spheres, and the suggestion that the Myth of Er reveals Plato as an advocate of epicyclic modelling. For Ptolemy, how corporeal spheres transmit or do not transmit their motions to the spheres that they enclose is a problem that remains crucial for contemporary physical cosmology, and that Aristotle mishandled.

Aristotle's purpose in *Metaphysics* Λ.8 was not to describe for its own sake a physical system of spherical shells conforming to the astronomical hypotheses of Eudoxos and Kallippos, but to obtain a tally of the number of divine unmoved movers that must exist on the hypothesis that one mover is required for each sphere; he concludes that it is "plausible" (εὔλογον) that there forty-seven movers for forty-seven spheres. It is without doubt in response to Aristotle that Ptolemy takes the trouble of counting the number of moving parts in his own cosmic system, finding the total to be as few as thirty-four or indeed as few as twenty-two, depending on whether one requires all the spheres and spherical shells to be complete or allows some to be truncated.[18] Ptolemy does not ascribe the cause of their motion to unmoved movers, but rather to planetary souls; each of the visible heavenly bodies is the seat of a single soul that causes and governs the motions of all the etherial bodies, visible and invisible, that constitute that body's system. Hence each heavenly body's system is a kind of divine animal whose mobile components can be likened to the limbs of a terrestrial animal. One passage (2.3/BM93a) appears to speak also of the entirety of the cosmos as a single animal, though such a unitary conception does not arise in any obvious way in his detailed account of the systems and their operation.[19] Like the question of how many moving bodies exist in the heavens, the question of the causes of their motions was extraneous to the

[18] In the "complete spheres" systems some of the bodies both are enclosed within and enclose other bodies that have identical revolutionary motions. Truncation allows these neighboring bodies to become a single body wrapping around the top and bottom.

[19] Murschel (1995, p. 38). The expression in question (*al-ḥayawān al-kullī*) may, however, render a Greek expression such as (μόρια) τοῦ παντὸς ζ ου, meaning just "the animal taken as a whole," as Nix interprets it.

mathematical astronomy of the *Almagest*, but in writing the *Planetary Hypotheses* Ptolemy evidently anticipated readers for whom these were more important than the technical details of planetary motion.

Those technical details, as laid out in terms of *Almagest*-style circles in Book 1 ("for the time being fitting the motions to the circles themselves, as if they were liberated from the spheres that contain them", 1A.2/BM82a) and in terms of solids in Book 2, would obviously be directed at the craftsmen whom Ptolemy exhorts in his preface to construct tangible models of the heavens that make visible the realities rather than the mere phenomena, that is, the eccenters and epicycles rather than the apparent changing speeds and the alternation of direct and retrograde motion. Beyond this insistence that the models should show what we *cannot* directly see in the sky, not just mimic what we *can* see,[20] Ptolemy is not very specific about the form the models should take, and he allows for both models composed of unconnected pieces that can be slid manually into appropriate orientations and models in which the motions are coordinated in appropriate ratios by "mechanical methods" (διὰ τῶν μηχανικῶν ἐφόδων), probably meaning gear trains. Book 1's descriptions of the celestial models are expressed in the language of geometrical construction, which is conceptual (e.g. employing such passive imperatives as ὑποκείσθω, "let there be hypothesized," and νοείσθω, "let there be imagined"), not the active language of mechanical construction; a glance at the chapters of the *Almagest* (e.g. 1.12 and 5.1) where Ptolemy describes how to construct instruments of observations shows the difference. In the *Planetary Hypotheses* Ptolemy is not giving instructions for making instruments but rather laying out the theory that someone else must design instruments to reproduce as best they can.

One kind of hand-set instrument that Ptolemy probably has in mind is an equatorium, in which the deferents and epicycles are simulated by revolvable graduated disks that can each be put in its appropriate position for a given date according to the tables of mean motions that originally stood at the end of Book 2.[21] This could serve as an analogue computing device, but Ptolemy apparently sees its value not so much in practical application (say for calculating planetary data for astrology) as in allowing one to confirm graphically that the theories are in precise agreement with observed phenomena (2.18/BM102a). It is worth keeping in mind that astronomical demonstration-models were valued in antiquity as instruments of philosophical instruction: Posidonios possessed one (Cicero, *De Natura Deorum* 2.88), Cicero repeatedly invoked their properties in his philosophical dialogues (*De Re Publica* 1.21–22, *Tusc. Disp.* 1.36, in addition to the passage just cited), and Theon of Smyrna, whose likening of the transfer of motion in celestial spheres to

[20] Ptolemy invokes the same rationale in *Geography* 1.1 for maps of the world, where the map is said to allow us to grasp the reality of the spherical Earth and our place on it, which we cannot see directly because we are too close to it and it is too big.

[21] Swerdlow (2005, p. 66) asserts that the subject of the first part of Book 1 *is* the construction of equatoria, though he concedes that the text does not explain how one would physically implement the three-dimensionality of the models with their variously inclined circles.

astronomical gearwork has already been mentioned, had a model of Plato's spindle and whorls (H146).[22] So the dichotomy among Ptolemy's intended readers is not as profound as it might initially appear.

When we say that Ptolemy in the *Planetary Hypotheses* is viewing the modelling of the heavens more from a philosopher's point of view than from a mathematical astronomer's, we must not forget that Ptolemy himself rejected such a dichotomy. His position, as set out in *Almagest* 1.1, was that mathematical astronomy is a part of mathematics proper (and not merely a science *dependent* on mathematics), that mathematics is one of the divisions of theoretical philosophy together with theology and physics, and that for human beings mathematics is the most valuable of the three because its subject matter, encompassing spatial properties and motions of unchanging bodies, is stable and knowable. Mathematical knowledge, he asserts, even confers some limited degree of illumination on theology (whose subject, the ultimate causes of motion and change, is inaccessible to the senses) and physics (whose subject, material properties of bodies, is unstable and irregular). In some sense the *Planetary Hypotheses* could be seen as an effort in this direction, using the "firm" results of mathematical astronomy to obtain an understanding of the physics and theology of the divine heavens. That he still regarded this understanding as less secure than the mathematical models is apparent in his frequent appeals in the *Planetary Hypotheses* to plausibility and likelihood, his use of analogies, and his readiness to offer more than one answer to a problem while making it clear which answer he favors.

Theon of Smyrna's physical modelling, as we saw, was based on three varieties of etherial body: spherical shells concentric with the cosmos to serve the function of the deferents in epicyclic models, solid spheres embedded within the "deferent" shells to serve the function of the epicycles, and (for the special case of the Sun-Mercury-Venus system) spherical shells again embedded within the "deferent" shells to serve the function of epicycles that are concentric with other epicycles. Ptolemy's theories of planetary motion were significantly more complex than Theon's, yet surprisingly in his first approach to giving them a physical interpretation he added only one new basic body type to Theon's repertoire, namely eccentric spherical shells embedded within "deferent" shells to serve the function of eccenters. In the case of the Sun, the visible body is directly embedded in such an eccenter shell (Fig. 3, showing the planet in four positions), while for the Moon and planets the solid or solids acting as the epicycle are embedded in the shell (Fig. 4). The idea of eccentric shells could well have been older than Ptolemy; the fact that Theon does not speak of it may indeed merely mean that he considered it to be physically less satisfactory than his epicyclic spheres.

We saw that Theon took over from Adrastos a definition of uniform motion as "traversing equal intervals in equal times" (τὸ τὰ ἴσα διαστήματα ἐν ἴσοις χρόνοις διανύειν, H151); this is expressed rather naïvely since the meaning of "intervals" when the motion is not rectilinear is not explained, but in the demonstrations Theon

[22] I have argued in Jones (2012) that the gearwork Antikythera Mechanism was likewise intended as a didactic instrument, not as a computer.

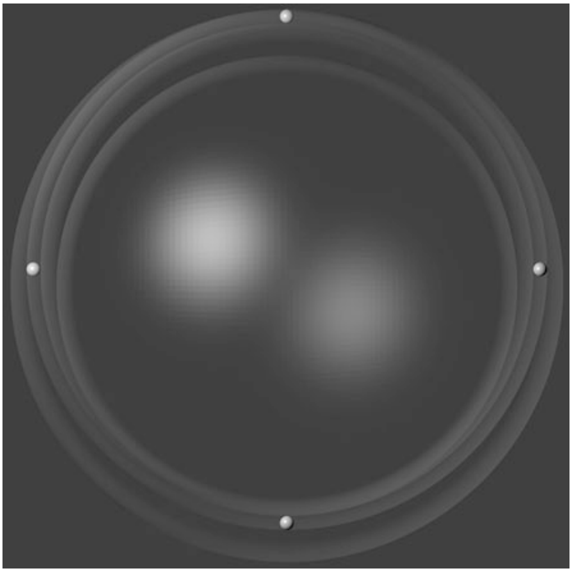


Fig. 3 Ptolemy's three-dimensional eccentric hypothesis for the Sun

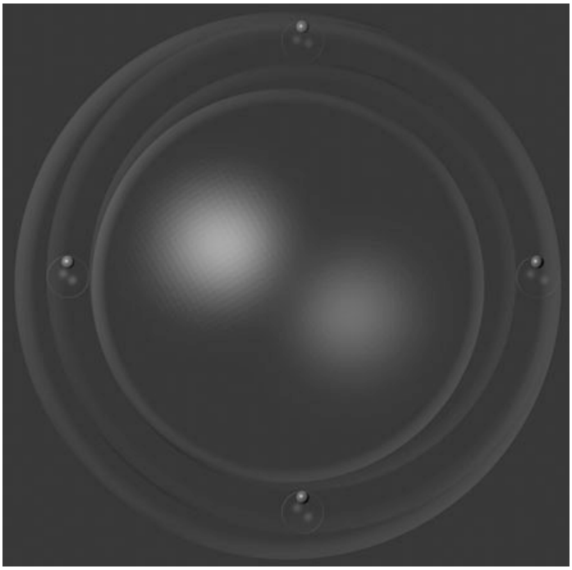


Fig. 4 Ptolemy's three-dimensional epicycle-and-eccenter hypothesis

clearly assumes that a uniformly moving body must be traversing equal arcs on its actual circular path. Again rather naïvely, he sees no need to define the frames of reference for the motions; in his practice, the frame of reference turns out to be zodiacally fixed for motion of a body on an eccenter or of an epicyle on a concentric deferent, but geocentrically fixed for motion of a body on an epicycle.

By contrast, Ptolemy takes it as a given, though in neither the *Almagest* nor the *Planetary Hypotheses* does he offer an explicit definition of the principle, that a motion along a circular path is uniform if there is *some* fixed point inside the circular path of the moving body or point such that the radius from the fixed point to the body or point sweeps out equal angles in equal times. In his eccentric solar model this point is indeed the center of the eccenter, but for the Moon and planets the center of uniform revolution of the epicycle's center is always a defined point (i.e. an "equant") different from the eccentric deferent's center.[23] Moreover, Ptolemy *does* state as a general principle that the frame of reference for the revolution of a body around an epicycle is the radius from the equant to the epicycle's center.[24] It may seem surprising that the descriptions of the physical models in *Planetary Hypotheses* 2 contain no allusion to the equants. A spherical "deferent" shell with an embedded epicycle revolving according to an equant model would have to exhibit a periodic oscillation in its rate of rotation, which seems intuitively harder to accept as "uniform motion" than the motion of the epicycle thought of in isolation from its corporeal context. Ptolemy probably believed that a revolution sweeping out equal angles with respect to a noncentral point was sufficiently explained by ascribing the cause of motion to a rational soul, which could through sheer volition maintain the required oscillation of speed just as it maintains the exact rates of revolution of each of its components through eternity.

It is in its treatment of planetary motion in latitude that the *Planetary Hypotheses* truly embraces the three-dimensionality of Ptolemy's cosmology. Theon has vague references to latitudinal motion, but his geometrical discussions of the epicyclic and eccentric hypotheses ignore it, restricting themselves to the relation between motion in "depth" (or anomaly) and apparent longitudinal motion. His shells and spheres are really little more than a two-dimensional diagram of an epicyclic model that has sprouted a third dimension after the fact. In the *Almagest* Ptolemy always starts his empirical working out of a model as if all the circles were in a single plane, treating anomalistic motion as something that can be analysed independently of motion in latitude; inclinations of the circles with respect to the plane of the ecliptic are added on to the models as a last stage, with Ptolemy relying on the comparatively small angles of inclination to ensure that the validity of the previous planar analyses is not undermined by the neglect of them. Ptolemy's planetary tables in the *Almagest* and *Handy Tables* also treat the calculation of planetary latitude as a kind of add-on.

---

[23] The equants in the *Almagest* consist of both an equant point and an "equant circle" centered on this point, such that radii from the equant point through the epicycle's center sweep out equal *arcs* on the equant circle. In the *Planetary Hypotheses* the equants are introduced simply as points "around which the center of the epicycle is moved with constant speed (ἰσοταχῶς)".

[24] It is noteworthy that Ptolemy states this principle in *Almagest* 5.5 when he is on the point of making an exception to the rule for his lunar model, and also noteworthy that the description of the lunar model in the *Planetary Hypotheses*, though otherwise unaltered from the *Almagest*, eliminates this exception without drawing any attention to the fact (Murschel 1995, p. 50).

Released in the *Planetary Hypotheses* from the obligation to show how his models and their parameters can be deduced from observations, Ptolemy no longer treats the models like paper dolls that lie flat until one punches their limbs out from the page. In Book 1, as each circle of a model is introduced, any inclination that it has with respect to the ecliptic or to the circle that bears it is specified immediately, so that the model grows in three dimensions all at once. The circles of Ptolemy's solar model, of course, lie entirely in the plane of the ecliptic, while the Moon's model lies entirely in a single inclined plane. Things become more interesting when we turn to the models for the five planets.

In the Book 1 "bare circles" descriptions of the models, Ptolemy states that the plane of each planet's eccentric deferent is tilted at a constant inclination relative to the ecliptic, though the inclinations for Mercury and Venus are very slight, just 1/6°. Revolving around this deferent circle with equant-uniform motion is a *sphere* that Ptolemy calls the "epicyclic sphere" (ἡ ἐπίκυκλος σφαῖρα), and on this sphere we are to imagine two "little circles" (κυκλίσκοι), both being great circles of the epicyclic sphere in the inclined plane of the deferent and thus apparently identical in location though not in motion:[25] the first circle is fixed relative to the radius from the equant point to the epicycle's center, while the second revolves relative to the first "with an equal course to the aforesaid (course) of the epicycle's center" but in the opposite sense. To this second circle is attached a third circle at a fixed inclination to the second circle, and the planet revolves around this third circle with uniform angular speed (ἰσοταχῶς) with respect to its center.

Ptolemy does not explain the kinematic functions of these "little circles." The first is really there just to establish a frame of reference for motions around the epicycle; in the systems of etherial solids of Book 2 it has no counterpart. The second circle, since it is revolving uniformly around the equant together with the epicyclic sphere on which it lies while simultaneously it is revolving *on* the epicyclic sphere with an equal but opposite motion in the same plane, is actually maintaining a stationary orientation according to the frame of reference of the whole model, as defined by its apsidal line. In Book 2 the solid corresponding to this is a spherical epicyclic shell, a device that we have seen applied by Theon to a different and more elementary purpose. The third circle (in Book 2, a solid epicyclic sphere enclosed by the shell), being fixed at an inclination to the second, also maintains a stationary orientation in the model's frame of reference while always being tilted relative to the plane of the deferent. Ptolemy tells us the angle of this epicyclic tilt but, oddly, omits to specify its direction relative to the tilt of the deferent. For Mars, Jupiter, and Saturn, however, the angle of epicyclic tilt is exactly equal to the tilt of the deferent relative to the ecliptic, and there can scarcely be doubt that he intends the two tilts to cancel each other out for these planets, so that the third circle always lies in a plane parallel to the ecliptic. From a modern perspective these *Planetary Hypotheses* models come close to an "ideal" approximation of the geocentric latitudinal motions of the planets by means of an epicycle-

[25] The radii of these circles are actually arbitrary since their function is entirely to do with their inclinations and revolutions, but the natural interpretation of the phrase ἐν τ ἐπικύκλῳ σφαίρᾳ, which applies to both, is that they are on the surface of the sphere.

and-eccenter model, with the circle in the model corresponding to the Earth's orbit around the Sun being either in the plane of the ecliptic or parallel to it, and the circle corresponding to the planet's orbit around the Sun maintaining a fixed inclination to the plane of the ecliptic; the small inclinations Ptolemy gives to the deferents of Venus and Mercury are the only defect. Yet it is striking that Ptolemy does not point out that it is a consequence of his models that the epicycle's orientation is constant and, in the case of the superior planets, parallel to the ecliptic, which one might have expected to be a fact about the planetary systems deserving some attention.

The planetary latitude models of the *Planetary Hypotheses* present further enigmas when we try to relate them to his earlier efforts to model planetary latitude; we can only touch on the complex issues here.[26] The very complicated latitude models of *Almagest* 13, as well as the modified ones underlying the planetary latitude tables in the *Handy Tables* which he published after the *Almagest* and before the *Planetary Hypotheses*, are fundamentally geocentric (or perhaps we should say "equant-centric"), in that they define the tilts of the epicycles with reference to the radius from the equant to the epicycle's center, and the tilts have oscillations. Unfortunately we know nothing directly about planetary latitude theories previous to Ptolemy, but an indirect argument would suggest that models similar to those of the *Planetary Hypotheses*, with the epicycle's inclination fixed in the frame of reference of the deferent (and in the case of the superior planets, parallel to the ecliptic), may have been known in his time.[27] If this is correct, then Ptolemy must have had reasons, either empirical or coming out of an a priori belief that epicycles *ought* to operate according to geocentrically-based principles of motion, or perhaps a combination of the two, for rejecting the fixed-orientation epicycles. Though the models of the *Planetary Hypotheses* are undoubtedly superior to Ptolemy's earlier models, one cannot be certain why he came to adopt them. He does say in very general terms at the beginning of the work (1A.2/BM81b) that, as a result of continued observation, he has made modifications of the models in the *Almagest* that include not only periodicities and dimensions but also structural changes in the models themselves; and aside from the abandonment of the nonstandard definition of the epicycle's apogee in the lunar model, the *only* structural changes are in the theories of planetary latitude. However, he also writes a little further on in his preface (1A2/BM82a) as follows:

> In the case of the placements and arrangement of the circles that produce the anomalies we will employ the simpler ones among the approaches for the sake of facilitating the construction of instruments, even if some little difference (παραλλαγή) results …

[26] See Swerdlow (2005) for an excellent treatment of the developments in Ptolemy's latitude theory, arguing that they were driven by empirical considerations.

[27] The basis of the argument, which I owe to Dennis Duke (private communication), is that algorithms for computing planetary latitudes existed in Indian astronomical texts that approximated the behavior of models like those in the *Planetary Hypotheses* (Kennedy and Ukashah 1969). Greek astronomical texts not dependent on Ptolemy are a plausible source for many features of the Indian models (Duke 2005).

which seems to mean that, in contrast to the *Almagest*, some choices in the model structures have been made not for the sake of theoretical correctness but to make mechanical modelling easier, even though these changes would result in small discrepancies in the predicted appearances. Again it is difficult to see anything else that this could refer to besides the planetary anomalies. Maybe both his accounts tell part of a story that involved both continued empirical refinement of the parameters of planetary latitude and a genuine uncertainty about whether the geocentric framework of the *Almagest* models might after all be incorrect, so that he allowed considerations of mechanical simplicity to tip the scale.[28] In any case the *Planetary Hypotheses* latitude models evince a remarkable degree of three-dimensional imagination on Ptolemy's part, even if he might not have entirely believed in them.

3 Ptolemy Before the Planetary Hypotheses

The short philosophical essay *On the Criterion and the Governing Faculty* is likely to be Ptolemy's earliest surviving work.[29] Although it is not about astronomy or cosmology, its second part, which is devoted to an attempt to localize the "governing faculty" (ἡγεμονικόν) of the human body and soul, contains some glancing remarks that show points of continuity in certain presuppositions between Theon (a representative of what was probably typical Middle Platonist cosmology), the young Ptolemy, and the Ptolemy of the *Planetary Hypotheses*. In particular, *On the Criterion* (ed. Lammert in Lammert and Boer 1952, 19) presents an Aristotelian five-element theory, with earth and water as the "more material" and passive elements, air and fire as "more motive" (κινητικώτερα) and both passive and active, and ether as "always in the same condition" (ἀεὶ ὡσαύτως ἔχοντα) and solely active. Ptolemy's concern in this work is not with the heavens but with human beings, but it is obvious that if he believed ether to be present in the cosmos at all, he would have believed it to be the sole or at least primary constituent of the heavens.[30] But ether is also the most characteristic and governing element in our souls, so there is not an absolute divide between the sublunary cosmos as composed only of the four transitory elements and the heavens as composed of ether. Through the presence of ether in the human soul, we possess an intellective capacity that Ptolemy speaks of as the "more divine" (θειότερον) aspect of the soul in *both* humans and the cosmos (ed. Lammert in Lammert and Boer 1952, 22). Hence the

[28] It is surely relevant that in *Almagest* 13.2 Ptolemy conceded that his latitude models would be difficult to reproduce in a mechanical model, but there he insisted that empirical evidence has to override considerations of simplicity, especially when the criterion of simplicity is reproducibility in non-etherial materials.

[29] Its authenticity, concerning which doubts have been raised, is assured by the presence of several words and word forms that are idiosyncratic favorites of Ptolemy but rare or unattested in other authors.

[30] Aristotle, *De Generatione Animalium* 737a comes close to positing an etherial component in the soul when he speaks of the generative warmth (θερμόν) in the seed as being not fire but a *pneuma* whose nature is "analogous to the element of the stars".

Ptolemy of the *Criterion* already conceived of the cosmos as an animated being, though apparently a single "cosmic animal" with its governing faculty high up in the heavens, as that of the human being is in the head.

In the *Almagest*, although the mathematical modelling is carried out in terms of two-dimensional circles (albeit with inclinations that bring them into a three-dimensional space), Ptolemy makes it clear from the outset that both the visible and the invisible bodies in the heavens are spherical and composed of ether (*Almagest* 1.3), and there are intermittent mentions of the spheres elsewhere in the work. What is not so clear is whether at this stage Ptolemy was committed to the kind of cosmology that the *Planetary Hypotheses* sets out, with discrete, non-overlapping, but tightly packed systems of bodies.

A crucial document in addressing this question is the *Canobic Inscription*, a summary of the parameters of his astronomical models that Ptolemy erected in AD 146/147 shortly before completing the *Almagest* as we have it, so that while most of the data in the inscription agree with the *Almagest*, a few were subjected to rethinking at a late stage in the composition of the treatise.[31] Among the data that Ptolemy altered after the inscription were the mean distances from the Earth of the Sun and of the Moon at syzygies, which in the inscription are respectively 64 and 729 Earth-radii, whereas according to the *Almagest* they are 59 and 1,210 Earth-radii. The inscription does not say anything directly about the absolute or relative distances of the other planets, but it concludes with a scheme matching numbers representing musical pitches with each of the heavenly bodies and the four elements that clearly reflects at least their relative distances from the center of the cosmos:

Sphere of the fixed stars	36
Saturn	32
Jupiter	24
Mars	21 1/3
Sun	18
Venus and Mercury	16
Moon	12
Fire and air	9
Water and Earth	8

Thus Venus and Mercury appear to be situated above the Moon but below the Sun, as they are in the *Planetary Hypotheses*, but they are assigned a single pitch-number, as if they could not be ranked in terms of relative distance from the center.

In *Almagest* 9.1 Ptolemy presents it as uncontroversial that the "order of the spheres" in the cosmos from outermost inwards is that of the fixed stars, Saturn, Jupiter, and Mars, and that the Moon's is innermost, but he presents it as an open question whether Venus and Mercury are further from the center or closer to the

[31] Hamilton et al. (1987), Jones (2005).

center than the Sun, though he favors the latter arrangement, giving as his reason that it would be plausible for the Sun to separate the planets that can reach all elongations from the Sun from those that cannot. Here he says nothing about the distances of Venus and Mercury relative to each other, though whenever in the *Almagest* he presents material relating to all five planets, the order is Mercury, Venus, Mars, Jupiter, Saturn.

Now as *Planetary Hypotheses* 1B shows, it would be possible for an implementation in terms of etherial solids of the *Almagest* models for Venus and Mercury —in either order—to fit snugly in the space between the outer boundary of the solid version of the *Almagest*'s lunar model and the inner boundary of the *Almagest*'s solar model. But obviously this was not possible at the *Canobic Inscription* stage of Ptolemy's thought, since the gap was much smaller then. One can only wonder whether Ptolemy had a resolution of this difficulty that might be hinted at by the single harmonic number shared by the two planets in the inscription. Perhaps he thought it might be possible for Mercury's epicycle to be within that of Venus, though the mechanics of such an arrangement would not be as trivial as Theon's onion-like triple epicycle since Ptolemy's models have different eccentric deferents. A system in which Mercury's model is subsumed within Venus's would not have entirely filled the interval between the inscription's lunar and solar models, so close packing could not have been envisioned.

Since the rediscovery of *Planetary Hypotheses* 1B in the 1960s, it has usually been assumed that it was only after completing the *Almagest* that Ptolemy noticed the coincidence that the Venus and Mercury models would fit neatly in the Moon-Sun gap, with its implications for the possibility of a close-packed cosmology. Christián Carman has recently shown that a plausible argument can be made that Ptolemy deliberately chose his revised solar distance in the *Almagest* with such a nesting of the models in mind, and contrived through numerical fudging so that this distance would appear to have come from analysis of observations of the Sun and Moon.[32] If this reconstruction is correct, however, it then becomes a puzzle why Ptolemy did not allude to the nesting until he came to write the *Planetary Hypotheses*.

There remains one further passage in the *Almagest* that attempts to describe something about the physics of the etherial bodies. It forms part of Ptolemy's defense of his models for planetary latitude in 13.2, the very models that were radically supplanted in the *Planetary Hypotheses* by the ones described above. The characteristic of these models that is at issue is that the planetary epicycles, and in the case of Venus and Mercury also the eccentric deferents, are supposed to have periodically oscillating inclinations. Ptolemy models these oscillations by hypothesizing that the points of the wobbling circles that mark their northern or southern limits of latitude ride on small wheel-like revolving circles in planes perpendicular to the ecliptic. Significantly, these perpendicular circles would not lend themselves easily to a three-dimensional interpretation as spherical bodies. Ptolemy's concern in 13.2 is not this, however, but that a critic might object that the models are too

[32] Carman (2008).

complicated to be credible, and moreover that it would be a practical impossibility to make a working "table-top" model of these devices, so how could they function in the heavens?

Ptolemy's rejoinder has two parts: firstly, that the materials that we use to make mechanical imitations of celestial models have different properties from the etherial material of the heavens, so that it is not legitimate to infer from a failed mechanical simulation that the proposed model is unviable, and secondly that, presuming a model is viable, we have no right to object to it on the grounds that it violates criteria of simplicity that we base on our mundane experience. This second part of the defense is a dangerous (and perhaps desperate) move on Ptolemy's part since one might wonder whether it leaves any place for simplicity arguments in astronomy. But the part that concerns us here is Ptolemy's characterization of the properties of etherial bodies:

> There is no hindering nature in them, but one that is commensurate [συμμέτρου] with respect to yielding and going along with each one's natural motions, even if they prove to be opposed, such that all can pass through and be seen through absolutely all the fluids [χυμάτων], and this kind of thing can flow freely [εὐοδεῖν] not only around the particular circles but also around the spheres themselves and the axes of revolutions.

The word I have translated as "fluids," χύμα, means something that is liquid or molten and thus flows (χέω) or that has solidified from a molten state (e.g. an ingot). Ptolemy does not apply it elsewhere to the invisible etherial stuff filling the heavens, but it occurs in this sense in much later authors, especially John Philoponos, and it must have had an existence that we cannot now trace in earlier cosmological texts of the Platonist-Peripatetic tradition since this obscure passage of Ptolemy is surely not where the later writers discovered it. In any event, ether was always in principle the lightest, most evanescent of the elements, but the modelling of Theon and the *Planetary Hypotheses* nevertheless treats the bodies composed of it as having rigid forms, so in effect "solid" in consistency as well as in geometry. In the *Almagest* passage, on the other hand, ether is not merely fluid in principle but actually flowing around other bodies of etherial matter. We obtain here a fleeting glimpse of a different approach to physical cosmology that Ptolemy seems to have entertained but not fully articulated in his *Almagest* days, and that he abandoned in favor of the Theonic approach.

4 Appendix: Translation of Theon of Smyrna on the Solid Models (H177–189)[33]

And (Adrastos) expounds these things at greater length with a view to accommodating to each other the hypotheses and approaches of the mathematicians, who, while giving regard only to the appearances and the motions of the wandering

[33] This translation (based on the text and diagrams of Hiller's edition) attempts, by avoiding conventional "technical" renderings of the Greek expressions, to approximate the way

(stars) that occur by happenstance, having observed them for long time-intervals because of the natural suitability of their country—(I mean) the Babylonians and Chaldeans and Egyptians—eagerly sought certain first principles and hypotheses, to which the appearances fit; by means of this (they would be able) to make judgment with respect to the things found before and to forecast with respect to things going to happen, some of them adducing certain numerical methods, like the Chaldeans, and others (adducing) graphical (methods), like the Egyptians, but all (of them) making their methods incomplete without reasoning according to nature, whereas it is needful at the same time to make examination concerning these things in a nature-related manner; and this is the very thing that those among the Greeks who engaged in astronomy tried to do, taking the first principles and the observations of the appearances from these (people), just as Plato discloses in the *Epinomion*, as will be clear a little later when his statements have been laid out.

And Aristotle, having demonstrated at length concerning the stars collectively in *On the Heavens* that they neither are borne through the etherial body which stands still, nor rush along with it as it is borne, just as if they are liberated and by themselves, nor indeed (are they) whirled about or rolled about, but rather the nonwandering (stars) are borne by that (etherial body), being multiple, by one common (sphere) that is outside, while each one of the wandering (stars) (is borne) by multiple spheres, again in *Metaphysics* Λ says that Eudoxos and Kallippos move the wanderers by means of certain spheres. For the (principle) according to nature is that neither do the stars move, with respect to themselves, along certain circular or spiral-shaped lines and oppositely to the whole nor do certain circles themselves whirl around their own centers while bearing the stars fixed upon them, some of them in the same direction as the whole and others oppositely. For how is it even possible for such great bodies to be bound on bodiless circles?

He says that it is fitting that there are certain spheres of the fifth body situated and borne in the depth of the whole, some of them higher up, others placed below them, and some of them greater, some of them smaller, and moreover some of them hollow, and again some of them solid, (being situated) in the depth of these (hollow spheres), on which the wanderers, being fixed on in the manner of nonwandering stars, appear as being moved, by a travel that is the simple one of those (nonwandering stars), but at unequal speed because of the places (where they are situated), by happenstance in a complicated manner, and (they appear as) describing certain eccentric circles or again (describing circles that are) situated on certain other circles or (as describing) certain spirals, on which the mathematicians think that they are moved, being deceived by the turning about. So since they appear as being

(Footnote 33 continued)
that Theon's intended readers—philosophical students without training in astronomy—would have encountered his book. For details see Jones (forthcoming), which includes a translation on the same principles of the immediately preceding passage, overlapping the present excerpt in its final paragraph. The most crucial way that conventional modern translation practices distort Theon's meaning is in the representation of "forward" and "backward" directions of celestial motion.

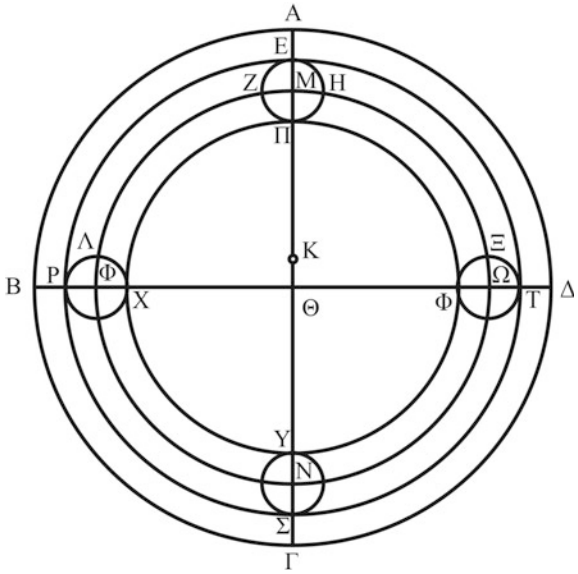
borne about together by the whole with the daily shift from risings to settings, but (also as) being borne in the opposite direction with the shift towards the trailing (stars) along the zodiacal (circle) which is inclined, and (as) also being moved a bit in breadth,[34] being seen more to the north and more to the south, and besides these things (in) height and depth, sometimes being beheld as further from the Earth and at other times as nearer to the Earth, Aristotle says that each (wanderer) is borne by means of multiple spheres that were previously hypothesized. Eudoxos says that Sun and Moon are fixed by three spheres, one (being) the (sphere) of the nonwandering (stars) that is whirled about the poles of the whole and that drags along by strength all the other (spheres) collectively from risings to settings, and a second (sphere) being borne around an axis at right angles to the (circle) through the middle of the zodiacal signs, by means of which again, each (of the Sun and Moon) appears as making the shift in length towards the trailing zodiacal signs collectively, and a third (sphere being borne) about an axis at right angles to the circle that is inclined with respect to the (circle) through the middle (of the zodiacal signs) in the breadth of the zodiacal signs, by means of which each is borne in its distinct motion in breadth, one in a greater and the other in a smaller interval, coming to be further north and further south of the (circle) through the middles of the zodiacal signs; and each of the other wandering (stars) by means of four (spheres), with another (sphere) added in the case of each, by means of which each one's[35] depth will be made. He says that Kallippos, with Kronos (Saturn) and Zeus (Jupiter) excepted, added certain other spheres to the others, by two for Sun and Moon, and by one for the rest. Next he additionally reasons that if (the spheres) when put together are going to save the phenomena, in the case of each of the wandering (stars) there are still other spheres, fewer by one than the bearing (spheres), (namely) the unwinding (spheres), either uttering his own opinion or that of those (*scil.* Eudoxos and Kallippos).

For since they thought that it was according to nature that all things be borne in the same direction, seeing the wandering (stars) also shifting in the opposite direction, they assumed that there had to be certain other spheres, obviously solid ones, between bearing (spheres), which unwind by their own motion the bearing (spheres) in the opposite direction, being in contact with them, just like the so-called "disks" in mechanical sphere-constructions, being moved about the center with a certain distinct motion, by the engagement of the teeth move and unwind in the opposite direction the underlying and juxtaposed (disks). The (principle) according to nature in fact is that all the spheres are borne in the same direction, being drawn about by the outermost (sphere), but according to their distinct motion, because of the arrangement of the placement and the places and the sizes, some of them travel faster, some of them slower in the opposite direction about distinct axes that are also inclined with respect to the sphere of the nonwandering (stars); so that the stars on them, being borne by the simple and uniform motion of these (spheres),

[34] inserting ⟨κατὰ⟩ πλάτος.

[35] reading ἑκάστου for the manuscript's ἕκαστον.

Fig. 5 Theon's diagram



seem by happenstance to make certain compound and nonuniform and complicated travels. And they describe different circles, some of them concentric, some eccentric, and some epicycles. For the sake of comprehension of the things that have been said (above), it is fitting to set out briefly also concerning these things, in accordance with the diagram that seems to us to be necessary for the sphere-constructions (Fig. 5).

Let there be a hollow sphere of the nonwandering (stars) ΑΒΓΔ about center Θ of the whole in depth ΑΕ; (and let there be) diameters of it ΑΓ, ΒΔ; and let circle ΑΒΓΔ be imagined, a greatest (circle) and through the middles of the zodiacal signs; and (let there be imagined) a certain other hollow sphere of a wanderer below it around the same center, ΕΡΣΤ and ΠΧΥΨ, in depth ΕΠ; and in this depth (let there be imagined) a solid sphere ΕΖΠΗ, bearing the wandering (star) fixed upon itself at Ε. And let them all be borne in the same direction uniformly in simple motions from risings to settings, but let the (sphere) that delimits the breadth of the wanderer alone be borne in the opposite direction, or (let it be borne) in the same direction but let it be trailing because of its slowness; for either way the phenomena will be saved. But *let*[36] the (sphere) of the nonwandering (stars be borne) about an axis at right angles to *the plane of the equinoctial (circle); and (let) the hollow (sphere) of the wanderer (be borne) about an axis at right angles to the* same plane in which also is the circle delimiting the breadth, which is inclined with respect to the (circle) through the middles of the zodiacal signs. And let the sphere of the unwandering (stars) be borne fastest; and the hollow (sphere) of the wanderer more slowly than this one in the opposite direction, so that in a certain defined time-period it goes

[36] Italics represent text restored in a lacuna by Martin.

about the whole (sphere) of the unwandering (stars), or, as some think, it trails; which is the truer opinion will be stated elsewhere; and let it bear the solid sphere that has the wandering (star); the solid sphere, being borne about its own axis uniformly, will be restituted to the same (situation), being borne in the same direction as the unwandering (sphere); either it will be restituted to the same (situation) in an equal time-period to that in which also the hollow (sphere) of the wandering (star), being borne in the opposite direction, goes about the (sphere) of the unwandering (stars) or it trails, either faster or slower.

Let it first be restituted in the same (time-period); and let M be center of the sphere; and with center Θ and radius ΘM let circle MΛNΞ be described; and with straight (line) EY having been divided in two at K, with center K and radius KE let circle EΛYΞ be described, being eccentric with respect to the whole. It is obvious that in the time-period in which the hollow sphere of the wandering (star) trails by the (sphere) of the unwandering (stars) while bearing the solid (sphere), center K of the solid sphere will go through concentric circle MΛNΞ, while seeming to be borne in the opposite direction while also bringing along the solid sphere, and the wandering (star) at E will describe the circle EHΠZ on the solid sphere, which (circle) becomes an epicycle on concentric MΛNΞ, while itself being borne in the same direction as the whole; but by happenstance it will also describe the eccentric (circle) EΛYΞ which is equal to the concentric, circumscribing it in the opposite direction to the whole; it will also seem to those who see from Θ to traverse the zodiacal (circle) ABΓΔ, advancing in the direction of the trailing (zodiacal signs) oppositely to the travel of the whole; it will also appear as being moved over a breadth in proportion of the inclination of the plane with respect to the (circle) through the middles of the zodiacal signs, to which plane the axes of its spheres are at right angles; it will always make its greatest distance at the same place and it will seem to move the least, say at point A of the zodiacal (circle), whenever the center of the solid sphere is on straight (line) AΘ at M, and the wandering (star) itself at E; contrarily the least distance will always be restituted and it will seem to move the most, say at point Γ of the zodiacal (circle), whenever, with the hollow sphere having changed its position in the opposite direction, the center of the solid (sphere) comes to be at N on straight (line) ΘΓ, and the wandering (star) itself at Γ, that is, at Y. However, it will make its mean distances and mean motions in two places, when it comes to be at the bisections of the epicycle EZΠH and of the concentric (circle) MΛNΞ, say the (bisections) Z, H, which, because of the change of position of the spheres in the opposite direction or the trailing, come to be the same as the bisections Λ, Ξ of the eccentric circle EΛYΞ and of the concentric (circle) MΛNΞ, appearing at the points in between A, Γ on either side, (namely) B, Δ on the zodiacal (circle), say Φ, Ω; all these things appear concerning the Sun, because of the fact that its time-periods of restitution all are equal so far as sense perception is concerned or are found very close to one another—I mean the (time-period) of the length and of the breadth and of the depth—and the corresponding points of both spheres always are seen to coincide at their corresponding motions at the same places and in the same zodiacal signs.

Since as a consequence of a travel of the spheres of such a kind and according to nature, uniform and simple and ordered, but inclined and only because of slowness trailing the unwandering (stars) or only the (sphere) bearing the solid (sphere), that is the epicycle, being borne in the opposite direction, by happenstance there arises a complicated and compound and nonuniform travel of the wandering (star), because of *the travel of the hollow (sphere)*[37] towards the trailing zodiacal signs which arises[38] either in reality or by trailing, and because of the inclination being beheld in a certain breadth of the zodiacal signs, and because of the whirling of the solid (sphere) about its own axis, sometimes seeming to be in height and because of this slow, sometimes in depth and because of this faster, and simply nonuniform, and because of this also seeming to be on the epicycle and on the eccentric (circle), it is also obvious that reasonably the mathematicians' hypotheses of their travel, (namely) the (hypothesis) according to epicycle and (that) according to eccentric (circle), follow each other and are in concord with each other, since both are consequences of the (hypothesis) according to nature, but by happenstance, at which Hipparchos marvels, especially in the case of the Sun because of the equality of time-periods of the travels of its spheres which is accurately completed, while in the case of the others (it is not completed) so accurately because of the fact that the solid sphere of the wanderer is not restituted in the same time-period in which the hollow (sphere) of the unwandering (stars) either trails or goes about in the opposite direction, but in the case of some of them faster, in the case of others slower, so that their corresponding motions do not coincide at the same points of the spheres at the same places, but always diverge, and also the inclinations of the spheres is in multiple breadths, and because of these things their time-periods of restitution of length and breadth and depth are unequal and different, and they make their greatest and least and mean distances and motions at different places at different times and in all the zodiacal signs, and moreover, because of the fact that, as we say, the corresponding motions diverge also (?) at the corresponding points of the spheres, the wandering (stars) do not seem to describe circles by their by-happenstance motions, but certain spirals. So in the case of each of the wandering (stars) one must think of the hollow sphere as distinct and bearing in its depth the solid (sphere), and (one must think of) the solid (sphere) as distinct, bearing on its distinct surface again the wandering (star).

In the case of the Sun and Light-bearer (Venus) and Gleamer (Mercury) it is possible for both (spheres) to be distinct for each, but (for) the hollow (spheres) of the three (spheres), being of equal course, to go about the sphere of the unwandering (stars) in the opposite direction in an equal time-period, while the solid (spheres) have their centers on one straight line, that of the Sun being smaller, that of Gleamer being larger, and that of Light-bearer being still greater than this one. But it is also possible for the hollow sphere of the three to be one, and the solid (spheres) of the three to be in the depth of this (hollow sphere) about the same center as each other, the (sphere) of the Sun being smallest and really solid, and the

[37] Hiller's restoration of a lacuna.

[38] reading γινομένης for the manuscript's γινομένη.

(sphere) of Gleamer about this one, and the (sphere) of Light-bearer thereupon enclosing both and filling the whole depth of the hollow and common (sphere); through this, these three make their trailing or travel in the opposite direction in length through the zodiacal signs of equal course, but the others not likewise, they are always seen catching up with each other and being caught up and standing in front of each other, with Hermes (Mercury) standing away at most about twenty degrees on either side of the Sun to the setting or to the rising, and the (star) of Aphrodite (Venus) (standing away) at most fifty degrees. One might suspect that the truer placement and arrangement is this one, so that this would be the place of the ensoulment of the cosmos, as cosmos and living thing, the Sun being as it were heart of the whole, greatly warm because of its motion and size and the companionship in journey of the things about it. For in ensouled beings the middle of the creature, that is of the living thing in the way that it is a living thing, is one thing, and (the middle) of the size is another thing; for example, as we said, the middle of the ensoulment of ourselves, as human beings and living things, is one thing, (namely) the (place) around the heart, which is always moving and greatly warm and because of these things is beginning of every power of the soul, for example spiritual (power) and (power) of impulse with respect to place and (power) of appetite and (power) of imagination and (power) of intellect, while there is another middle of our size, such as the (place) about the navel. Similarly too the middle of the whole cosmos, to liken the greatest and most worthy of honor and divine on the basis of slight and random and mortal things, is the (place) about the earth, which is chilled and immobile; but as cosmos and in the way that it is cosmos and living thing, the middle of its ensoulment is the (place) about the Sun, as it were being heart of the whole, starting from which they relate that its soul makes its way through the whole body, being stretched from the extremities.

It is obvious that whereas because of the aforesaid causes both hypotheses follow each other, the (hypothesis) according to the epicycle seems more collective and more general and very close to the (hypothesis) according to nature; for the greatest circle of the solid sphere, which the wandering (star) describes on it by its travel on it, is the epicycle; but the eccentric (circle) is absolutely separated from the (hypothesis) according to nature and is described more by happenstance. Comprehending this, Hipparchos too praises the hypothesis according to epicycle as being his own, saying it is more credible that all the heavenly things are situated in equilibrium and joined together likewise with respect to the middle of the cosmos; nevertheless he himself, because he did not proceed from reasoning according to nature, did not comprehend accurately what is the travel of the planets that is according to nature and accordingly true is, and what is the (travel that is) by happenstance and apparent; even he hypothesizes that each one's epicycle is moved on the concentric circle, and the wandering (star) on the epicycle.

Plato too seems to believe that the (hypothesis) according to epicycle is more powerful, but that the things that bear the wandering (stars) are not spheres but circles, just as at the end of the *Republic* he riddles by the whorls fitted in one another; but he uses the words in a more common manner, and often calls spheres "circles" and "poles," and axes as "poles."

References

Carman, C. (2008). Rounding numbers: Ptolemy's calculation of the earth-sun distance. *Archive for History of Exact Sciences, 63*, 205–242.

Delattre Biencourt, J. (2010). *Théon de Smyrne. Lire Platon. Le recours au savoir scientifique: arithmétique, musique, astronomie*. Toulouse: Anacharsis.

Duke, D. (2005). The equant in India: The mathematical basis of ancient Indian planetary models. *Archive for History of Exact Sciences, 59*, 563–576.

Dupuis, J. (1892). *ΘΕΩΝΟΣ ΣΜΥΡΝΑΙΟΥ ΠΛΑΤΩΝΙΚΟΥ ΤΩΝ ΚΑΤΑ ΤΟ ΜΑΘΗΜΑΤΙΚΟΝ ΧΡΗΣΙΜΩΝ ΕΙΣ ΤΗΝ ΠΛΑΤΩΝΟΣ ΑΝΑΓΝΩΣΙΝ. Theon de Smyrne philosophe platonicien exposition des connaissances mathématiques utiles pour la lecture de Platon*. Paris: Hachette.

Düring, I. (1932). Porphyrios Kommentar zur Harmonielehre des Ptolemaios. *Göteborgs högskolas Årsskrift, 38*, 2.

Evans, J. (2003). The origins of Ptolemy's cosmos. In S. Colafrancesco & G. Giobbi (Eds.), *Cosmology through time: Ancient and modern cosmologies in the Mediterranean Area Conference Proceedings* (pp. 123–132). Milan: Mimesis.

Feke, J., & Jones, A. (2010). Ptolemy. in L. Gerson (Ed.) *The Cambridge history of philosophy in late antiquity* (Vol. 1, pp. 197–209). Cambridge: Cambridge University Press.

Goldstein, B. R. (1967). The arabic version of Ptolemy's planetary hypotheses. *Transactions of the American Philosophical Society NS*, *57*(4) (Philadelphia: American Philosophical Society).

Hamilton, N. T., Swerdlow, N. M., & Toomer, G. J. (1987). The canobic inscription: Ptolemy's earliest work. In J. L. Berggren & B. R. Goldstein (Eds.), *From ancient Omens to statistical mechanics* (pp. 55–73). Copenhagen: University library.

Heiberg, J. L. (1907). *Claudii Ptolemaei Opera quae exstant omnia. II. Opera astronomica minora*. Leipzig: Teubner.

Hiller, E. (1878). *Theonis Smyrnaei Philosophi Platonici Expositio Rerum Mathematicarum ad Legendum Platonem Utilium*. Leipzig: Teubner.

Jones, A. (2005). Ptolemy's canobic inscription and heliodorus' observation reports. *SCIAMVS, 6*, 53–97.

Jones, A. (2009), Ancient rejection and adoption of Ptolemy's frame of reference for longitudes. in A. Jones (Ed.) *Ptolemy in perspective. Use and criticism of his work from antiquity to the nineteenth century*, (Archimedes 23, pp. 11–44). Dordrecht-New York: Springer.

Jones, A. (2012). The Antikythera Mechanism and the Public Face of Greek Science. in *Proceedings of Science* PoS (Antikythera & SKA) 038. http://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=170.

Jones, A. (forthcoming). Theon of Smyrna on the apparent motion of the planets.

Kennedy, E. S., & Ukashah, W. (1969). Al-Khwārizmī's planetary latitude tables. *Centaurus, 14*, 86–96.

Lammert, F., & Boer, E. (1952). *Claudii Ptolemaei Opera quae exstant omnia. III. 2. De judicandi facultate et animi principatu. Pseudo-Ptolemaei Fructus sive Centiloquium*. Leipzig: Teubner.

Lawlor, R., & Lawlor, D. (1979). *Mathematics useful for understanding plato*. San Diego: Wizards Bookshelf.

Moraux, P. (1984). *Der Aristotelismus bei den Griechen: von Andronikos bis Alexander von Aphrodisias. 2. Der Aristotelismus im I. und II. Jh. n. Chr*. Berlin: De Gruyter.

Morelon, R. (1993). La version arabe du Livre des hypothèses de Ptolémée. *Mélanges de l'Institut dominicain d'études orientales du Caire, 21*, 7–85.

Murschel, A. (1995). The structure and function of Ptolemy's physical hypotheses of planetary motion. *Journal for the History of Astronomy, 26*, 33–61.

Richter, G. M. A. (1965), *The portraits of the Greeks* (3 Vols.). London: Phaidon Press.

Sharples, R. W. (2010). *Peripatetic philosophy, 200 BC to AD 200: An introduction and collection of sources in translation*. Cambridge: Cambridge University Press.

Swerdlow, N. M. (2005). Ptolemy's theories of the latitude of the planets in the Almagest, handy tables, and planetary hypotheses. in J. Z. Buchwald & A. Franklin (Ed.), *Wrong for the right reasons* (Archimedes 11, pp. 41–71). Dordrecht-New York: Springer.

Tannery, P. (1893). *Recherches sur l'histoire de l'astronomie ancienne*. Paris: Gauthier-Villars.

van der Waerden, B. L. (1988). *Die Astronomie der Griechen: Eine Einführung*. Darmstadt: Wissenschaftliche Buchgesellschaf.

Waszink, J. H. (1962). *Timaeus a Calcidio translatus commentarioque instructus. Plato Latinus 4*. London: Warburg Institute.

Proclus' Conception of Geometric Space and Its Actuality

David Rabouin

The main aim of this paper is to present Proclus' philosophy of geometric extension not so much from the point of view of what he says about it as what he does with it. I will henceforth pay particular attention to the role of spatial configurations in the *practice* which he describes. My motivations are twofold. First, although Proclus' philosophy of geometry has received quite a lot of attention in the scholarship, this attention has remained mainly inspired by the philosophical doctrine expounded in the *Prologues.*[1] It did not engage much, at least in a systematic way, with the material given in the actual commentary of Euclid's propositions and the mathematical practice there described. As a consequence, the complexity and the flexibility of Proclus' views on the geometric imagination were not always well rendered. I would like to complete existing descriptions by paying more attention to these details, although, as I will indicate, they may sometimes introduce important nuances, if not tensions, in the philosophical system. Second, Proclus provides indications throughout his commentary about the geometric practice which go far beyond his own specific philosophical agenda. He deals, for example, with objections that other mathematicians and philosophers raised against Euclid's proofs. These objections, which sometimes stem from views opposite to his own (typically, Epicurean objections of an "empiricist" flavor), were taken seriously enough to ask for answers which Proclus also mentions (or sometimes even initiates). For the modern reader, these passages are precious because they provide, by

D. Rabouin (✉)
Université Paris 7—CNRS Laboratoire SPHERE UMR 7219, 5 rue Thomas Mann,
75205 Paris Cedex 13, France
e-mail: david.rabouin@wanadoo.fr

[1] See Breton (1969), Charles-Saget (1982), Mueller (1987), O'Meara (1989), Cleary (2000), Nikulin (2002) and Lernould (2010).

V. De Risi (ed.), *Mathematizing Space*, Trends in the History of Science,
DOI 10.1007/978-3-319-12102-4_5

contrast, testimonies about certain conditions of the practice which seem to have been accepted by the various interlocutors (whatever their philosophical background may have been). I would like to reconstitute some of these conditions and compare these reconstructions with recent works dealing with ancient Greek geometric practice, especially as regards the use of diagrams.[2]

My approach to Proclus will therefore be partly instrumental. Instead of focusing on the quite rigid Neoplatonic framework set up in the *Prologues*, which has already received considerable attention and which in Proclus' eyes was probably the most important aspect of his interpretative enterprise, I will try to extract a more flexible approach of geometric practice found throughout his commentary.[3] My handling of the question will be focused on the role of "impossible" diagrams, be they related to the picturing of infinite or indivisible entities, to objections raised against this or that construction or to proofs by *reductio*. In the first part of the paper, I will motivate this choice with the first series of examples. As we will see, it so happens that Proclus gives important indications on these cases about the role of the geometric imagination. He insists on this occasion, more so than in the *Prologues*, on the *autonomous* activity of the imagination and on its relative *opacity* to the understanding. These two features will constitute the *leitfaden* of my study. In the second part of the paper, I will show that this situation is not specific to the first set of examples and conforms to the frequent use of other "impossible" representations in Ancient Geometry (a fact which indicates accordingly that they are by no means "impossible"). I will document some of their appearances in Proclus, detail how the geometric imagination is involved in these situations and indicate several related philosophical issues. In the last part of the paper, I will focus on these philosophical issues and compare them with recent attempts to characterize Ancient Geometric Practice. As a conclusion, I will try to show how this description may help us tackle broader issues dealt with in contemporary philosophy of mathematics.

---

[2] In particular the studies by Manders (2008b), and Netz (1999). As regards the characterization of geometric "practice", I follow here an insightful remark made by Netz in passing (p. 2): "what unites a scientific community need not be a set of beliefs. Shared beliefs are much less common than shared practices. This will tend to be the case in general, because shared beliefs require shared practices, but not vice versa. And this must be the case in cultural settings such as the Greek, where polemic is the rule, and consensus is the exception. Whatever is an object of belief, whatever is verbalisable, will become visible to the practitioners. What you believe, you will sooner or later discuss; and what you discuss, especially in a cultural setting similar to the Greek, you will sooner or later debate. But the real undebated, and in a sense undebatable, aspect of any scientific enterprise is its non-verbal practices".

[3] Although the extant text of Proclus' commentary deals only with the first book of Euclid's *Elements*, my interest will not be about the practice attributed to a particular mathematician ("Euclid") and developed into a text entitled the *Elements*, but on the common practice spanning centuries (say, at least until Proclus) by practitioners of Euclidean Geometry and which remained quite stable through various interpretations (involving different cultural settings, conceptual debates, ideological criticisms, etc.). Moreover, what I will describe easily extends to other classical authors such as Archimedes and Apollonius.

1 Proclus on Geometric Space

Before entering into the flesh of Proclus' views, a word of justification about the choice of the topic seems to be in order. Indeed it may sound strange to aim at studying "geometric space" in Proclus, as it does more generally for ancient Greek mathematics. A usual claim would be, quite on the contrary, that ancient Geometry was a science of figures in contrast with our modern Geometry which takes space and transformations in space as its primary object of inquiry. In his seminal study on Euclid's *Elements*, Ian Mueller stresses this difference between the modern "structural" viewpoint and that of the Ancients:

> For Hilbert geometric axioms are characterized by an existent system of points, straight lines, etc. At no time in the *Grundlagen* is an object brought into existence, constructed. Rather its existence is inferred from the axioms. In general Euclid produces, or imagines produced, the objects he needs for a proof (…). It seems fair to say then that in the geometry of the *Elements* there is no underlying system of points, straight lines, etc. which Euclid attempts to characterize. Rather, geometric objects are treated as isolated entities about which one reasons by bringing other entities into existence and into relation with the original objects and one another.[4]

A source of great confusion on this issue was the description of the birth of "Modern Science" as characterized, to repeat Koyré's famous wording by "the replacement of the Aristotelian conception of space—a differentiated set of inner worldly places, by that of Euclidean geometry—an essentially infinite and homogenous extension—from now on considered as identical with the real space of the world".[5] Indeed, in this picture, it seems that Euclidean Geometry is already, by contrast to Aristotelian philosophy, a science of infinite (homogeneous) extension. But, as emphasized by E. Grant in a manner very close to that of Mueller:

> There is nothing in Euclid's geometry to suggest that he assumed an independent, infinite, three-dimensional, homogeneous space in which the figures of his geometry were located. In a purely geometric sense, such a space would have been superfluous because every geometric figure has its own internal space. Moreover, if the space of the geometric figure and the independent space it is alleged to occupy are conceived as indistinguishable, an infinity of spaces could be postulated in one and the same place. (…) Euclidean geometric space was the space of geometric figures of any size whatever and when applied to material bodies was conceived as an internal space.[6]

---

[4] Mueller (1981, p. 14).

[5] Koyré (1957, Preface, p. viii).

[6] Grant (1981, p. 17). The last sentence of the quote is intended to show the proximity between Euclid and Aristotle's concept of place. This goes along with another important historical rectification emphasized by Grant: "the adoption of an infinite space in the seventeenth century resulted primarily from the divinization of space—a process begun in the fourteenth century—and to a lesser extent, from the needs from physics and cosmology. But it did not arise from any straightforward application of an alleged Euclidean geometric space to the physical world" (note 49, p. 273).

Considering these debates, it might seem that the attribution of a doctrine of "geometric space" to Greek mathematicians amounts to some form of anachronistic projection of our modern "structural" viewpoint (i.e. what we *now* call "Euclidean Geometry"). It is therefore worth recalling that although we have no way of knowing what Euclid thought about the subject, there was at least one Ancient Greek (although much later than Euclid) who had no trouble finding infinite space in his geometry: Proclus! Moreover, he gave strong justifications for this claim.

The discussion on this topic occurs mainly in the commentary of proposition I.12, which asks: "To a *given infinite straight line*, from a given point which is not on it, to draw a perpendicular straight line" (my emphasis). Proclus remarks that the condition ("from a given point which is not on it") cannot be known to be satisfied if we take an arbitrary finite segment and a point at random in the plane. Indeed we would have no way of knowing whether or not the given point would be on the line when produced (a situation dealt with in the previous proposition and asking for another construction cf. *In Eucl.* 284).[7] This is, according to him, why Euclid was led to posit an infinite straight line given *in actuality* (κατ᾽ ἐνέργειαν). He insists then on the fact that "if there is an infinite line, there will also be an infinite plane, and infinite in actuality if the problem is to be a real one" (*In Eucl*. 284, 19–21). This last stance transfers immediately to three dimensions, since proposition XI.11 of the *Elements* asks for the construction of a perpendicular to a plane and relies directly on I.12. If the plane in I.12 is infinite in actuality, so will be the three dimensional space containing it (by repeating the construction twice, if needed).

One could wonder if too much emphasis can be put on a single mention which remains isolated in Euclid and, by consequence, in Proclus. But the commentator also emphasizes that the *Elements* express at several occasions the fact that a straight line is given as *finite* (prop. I.1, I.10), a specification which would be pointless if there were no possibility for a straight line to be given as *not* finite. He also indicates that the condition is not specified in certain propositions, because it is implicit in the position of the problem (see for example *In Eucl.* 208, 11–12; 223, 16; 224, 4; 277, 18–24). Although these terminological remarks are not enough to tell us about Euclid's own conception, they certainly indicate that the acceptance of an actual infinite in Geometry was more than a strange *hapax legomenon* in Proclus' discourse. Moreover, one has to remember that the construction involved in I.12 is ubiquitous in the *Elements* (and more generally in Ancient Greek Geometry).

The epistemological problem raised by this description is of course that an actual infinite space seems inconceivable if one looks either in the realm of sensible things, in which only finite magnitudes have existence (culminating according to Proclus with the last sphere which gives a boundary to all material entities), or in the realm of intelligible where there is a platonic 'idea' of infinity, but not endowed

[7] The Greek text is Proclus (1873). I will follow the English translation by Morrow (1992). For a contemporary discussion of the problem raised by the choice of arbitrary points in the *Elements* and their relation to the presupposition of an infinite system of objects, see Mumma (2012, p. 117; for propositon I. 11 and 12, see in particular note 14).

with spatial extension (*In Eucl.* 284–285). To this traditional dilemma, Proclus proposes a solution, which is of particular interest to us:

> It remains, then, that the infinite exists in the imagination, only without the imagination knowing the infinite. For when imagination knows, it simultaneously assigns to the object of its knowledge a form and limit, and in knowing brings to an end its movement through the imagined object; it has gone through it and comprehend it. The infinite therefore is not the object of knowing imagination, but of imagination that is uncertain about its object, suspends further thinking and calls infinite all that it abandons, as immeasurable and incomprehensible to thought. Just as sight recognizes darkness by the experience of not seeing, so imagination recognizes the infinite by not understanding it. It produces it indeed, because it has an indivisible power of proceeding without end. (*In Eucl.* 285, 5–17)

This does not contradict what Mueller told us.[8] Space does not appear here as an object of inquiry or a system of objects which would be treated mathematically as such. Indeed Proclus insists on the fact that it is given by imagination as something that the understanding *cannot grasp*. In this sense, geometric space acts only in the background of the theory. But this background, although there is *stricto sensu* no concept of it, is nonetheless presented as a necessary condition for a proper study of geometric objects. It is not something that we know, but something that we need in order to know.[9] Hence one should not confuse the acceptance of an infinite geometric space and the fact of taking it as an object of mathematical study. The purpose of this paper is to elucidate the first of these directions by embedding it, following Proclus, in a more general view on the role of the imagination in Geometry and on the way space is *used* not as an object of study, but as a tool for the study of mathematical objects.

These remarks are of great importance from a historiographical point of view. Even if modern scholars are more cautious than their predecessors with regards to diagnosing "revolutions" in science, they often leave untouched the above mentioned thesis, according to which the birth of an infinite *mathematical* space in the Renaissance was an important breakthrough (the existence of an infinite physical space being already present in certain Ancient philosophers). In this regard, it

[8] It does seem, however, to directly contradict Grant's declarations: according to Proclus, we certainly need an actual infinite extension to perform Euclidean geometry.

[9] "The understanding from which our ideas and demonstrations proceed does not use the infinite for the purpose of knowing it, for the infinite is altogether incomprehensible to knowledge; rather it takes it hypothetically and uses only the finite for demonstration; that is, it assumes the infinite not for the sake of the infinite, but for the sake of the finite" (*In Eucl.* 285–286). Compare with Mumma (2012, p. 117): "In the construction stage, most steps produce unique geometric objects from given ones and so can be represented logically by functions. Yet one kind does not: the free choice of a point satisfying non-metric positional conditions. Such points have an *indefinite* character within proofs. Their precise identity is not fixed relative to other given objects in the configuration. The natural logical representation for what licenses their introduction are thus existential statements, asserting the existence of a point satisfying certain non-metric positional conditions. *And so, though the geometric reasoning in* ***Eu*** *is always performed with a particular finite diagram, it still seems to presuppose a domain of geometric objects, i.e. the domain over which the quantifiers of these propositions range*" (my emphasis). The link between "indefinite" objects and the presupposition of an infinite domain of objects is of particular interest.

should be recalled not only that it was a possibility foreseen in Ancient times, but that, symmetrically, such an idea was controversial amongst early modern thinkers. Moreover the Proclean position of the problem, which distinguishes between what is given to understanding and what is given to imagination, has strong echoes in these discussions. In particular, one has to keep in mind that Descartes, one of the alleged "heroes" of the modern identification between mathematical space and the real world, *rejected* the idea of an infinite extension as conceptually given and introduced the idea of "indefinite" for what is given as infinite to the imagination.[10] It should also be stressed that in Proclus one can find not only the acceptance of an infinite extension given by imagination, but a dynamical conception of geometric space in which *transformations* are brought to the fore (in contrast to the study of properties of static figures, often presented as a characteristic of Ancient geometry). In fact, the history of the "Euclidean" tradition shows continuous evolution toward this view[11] and produced as early as Ibn al-Haytham and al-Sijzi presentations of Geometry in which space and transformations became of primary importance.[12]

It is no big surprise that geometric space makes its appearance in the case of entities such as the infinite straight line or points taken at random in the plane (i.e. not as intersection of given lines or circles).[13] It provides us with a typical situation in which the spatial proxy acting in the background of the theory and the conceptual apparatus operating on its surface are not in perfect accordance—whereas other regular situations (but hardly all of them, as I will argue in Sect. 2) could let us think that they evolve in a perfect parallelism. There are spatial configurations which, so to speak, do not "correspond" to conceptual configurations. This relative opacity of space, as I will call it following the Proclean metaphor of darkness and sight, will be the leading topic of my inquiry. Note however that the problem here is not presented in the form which a modern reader could expect: it does not come from the fact that we, human beings with finite resources, are unable to properly represent an actual infinite. Quite on the contrary since Proclus credits human knowledge with being able to represent infinity thanks to imagination! What is

[10] "I do distinguish here between 'indefinite' and 'infinite'; strictly speaking, I designate only that thing to be 'infinite' in which no limits of any kind are found. In this sense God alone is infinite. However, there are things in which I discern no limit, but only in a certain respect (such as the extension of imaginary space, a series of numbers, the divisibility of the parts of a quantity, and the like). These I call 'indefinite' but not 'infinite,' since such things do not lack a limit in every respect." ("Reply to First Set of Objections", translation in Ariew 2000, p. 155; AT 7: 113). The close connection between the concept of 'indefinite' and imagination is emphasized as early as *Le Monde* (AT 11: 31–33, transl. Ariew 2000, p. 36) and persists until the *Principles of Philosophy* (I § 26–27, AT 8a: 15, transl. Ariew 2000, pp. 237–238). On Newton's proximity with Proclus at the time of the *De gravitatione*, see Domski (2012).

[11] See Vitrac (2005, pp. 1–56), which gives a prominent role to Proclus.

[12] See Rashed (2001, Introduction, pp. 1–11 and Chap III, pp. 655–685: "Ibn Al-Haytham et la géométrisation du lieu"), Crozet (2010), De Vittori (2009).

[13] The word used in the *Elements* is τυχὸν σημεῖον. Before I.12, it appears in I.5, I.9 and I.11.

particularly interesting is precisely that we are able to represent something to which we have no conceptual access.

*

The commentary of Eucl. I.12 provides the occasion to recall the general epistemological framework set up by Proclus in the *Prologues*. It also allows us to identify important nuances introduced in the details of the commentary. This is particularly the case with the *autonomous* activity which the imagination acquires with increasing clarity as the commentary unfolds.

In the beginning of the book (*Prologue* 1), under the guise of an orthodox Neoplatonism, Proclus has emphasized the intermediate nature of mathematical entities which are neither simple nor dispersed in the uncontrolled diversity of Becoming, neither mere objects of intellect (*nous*) nor of opinion (*doxa*) (*In Eucl.* 3–5). In the second *Prologue*, however, paying closer attention to geometric objects, he introduces a more original view in which this intermediary position is not linked solely to the position of discursive thinking (*dianoia*), but to the intervention of a kind of "matter" identified with imagination (*hylè phantastikè* cf. *In Eucl.* 51–52). Here, Proclus is very clear about the fact that he is not following an orthodox platonic doctrine.[14] He then goes so far as to assert that *dianoia* is *unable* to access geometric ideas by itself and *needs* the help of imagination to seize them:

> When, therefore, geometry says something about the circle or its diameter, or about its accidental characteristics, such as tangents to it or segments of it and the like, let us not say that it is instructing us either about the circles in the sense world, for it attempts to abstract from them, or about the form in the understanding. For the circle [in the understanding] is one, yet geometry speaks of many circles, setting them forth individually and studying the identical features in all of them; and that circle [in the understanding] is indivisible, yet the circle in geometry is divisible. Nevertheless we must grant the geometer that he is investigating the universal, only this universal is obviously the universal present in the imagined circles. Thus while he sees one circle [the circle in imagination], he is studying another, the circle in the understanding, yet he makes his demonstrations about the former. For the understanding contains the ideas but, *being unable to see them when they are wrapped up, unfolds and exposes them and presents them to the imagination sitting in the vestibule*; and in imagination, or with its aid, it explicates its knowledge of them, happy in their separation from sensible things and finding in the matter of imagination a medium apt for receiving its forms. (*In Eucl.* 54–55, my emphasis)

We find here for the first time the crucial idea that imagination provides *dianoia* with a knowledge which it is unable to access by itself. At this stage, however, it would still be possible to understand this doctrine as a mere complement to what has been stated in the first prologue. One could claim that the discursive movement remains attached only to *dianoia*, imagination being just a proxy on which this

---

[14] "We are not unaware of what the philosopher Porphyry in his *Miscellaneous Inquiries* and most of the Platonists have set forth, but we believe that what we have said is more in agreement with the principles of geometry and with Plato's declaration that the objects of geometry are understandables" (*In Eucl.* 56).

movement is transcribed and which allows the manifestation of pure and simple ideas. The famous metaphor of the *projection* on the "receptacle" of imagination seems to go in this direction:[15]

> We invoke the imagination and the intervals that it furnishes, since the form itself is without motion or genesis, indivisible and free of all underlying matter, though the elements latent in the form are produced distinctly and individually on the screen of imagination. What projects the images is the understanding; the source of what is projected is the form in the understanding; and what they are projected in is this "passive nous"[16] that unfolds in revolution about the partlessness of genuine *Nous.* (*In Eucl.* 56)

This seems even clearer in the metaphor of the surface on which the understanding *writes* its mathematical concepts (or "ratios" since Proclus designates both as *logoi*). Imagination is then presented as a plane mirror[17] on which discursive knowledge contemplates itself: "We must think of the plane as projected and lying before our eyes and the understanding as writing everything upon it, the imagination becoming something like a plane mirror to which the ideas of the understanding send down impressions of themselves" (*In Eucl.* 121).

One important thing to notice in these various declarations is that imagination is identified with the spatial support, Proclus establishing a strong connection between what could appear as a faculty of the soul and what I have designated above as a form of "geometric space". In this sense, space is not just a screen on which the *dianoia* projects its concepts, it is also, and by the same token, a screen on which the soul recognizes its rational activity. But, in any case, the activity seems situated on the discursive side, the apparent activity on the surface being just a reflection of the dianoetic activity. This idea is well expressed in the description of the geometric "figure" where the mirror is presented as a surface on which the seer and the seen coincide (because in the mirror I see myself seeing):

> Therefore just as nature stands creatively above the visible figures, so the soul, exercising her capacity to know, projects on the imagination, as on a mirror, the ideas of the figures; and the imagination, receiving in pictorial form these impressions of the ideas within the soul, by their means affords the soul an opportunity to turn inward from the pictures and attend to herself. It is as if a man looking at himself in a mirror and marveling at the power of nature and at his own appearance should wish to look upon himself directly and possess such a power as would enable him to become at the same time the seer and the object seen. (*In Eucl.* 141)

[15] On Proclus's "projectionism", see Mueller in Morrow (1992, p. xxvi), Mueller (1987) and O'Meara (1989).

[16] This is the way Aristotle designates imagination in *De Anima* 430a24. As explained a few pages earlier by Proclus, this expression should be taken *cum grano salis* since there is no such thing as a *passive* nous in his view (*In Eucl.* 52).

[17] A metaphor coming from Plato (especially *Timaeus* 70e–f) and elaborated by Plotinus, see Claessens (2012, where the relevant literature is mentioned).

However, as we have seen in I.12, imagination does not limit itself to reflecting a dynamic coming from outside. It is also endowed with a *dynamic of its own.*[18] In the case of infinity, no reflexivity is allowed since the dynamic attached to it is presented as *irreducible* to conceptualization: with infinite entities (be they straight lines, planes, three dimensional space), imagination provides knowledge with forms of representation which must remain opaque to it. This is why I talked about important nuances introduced in the course of the commentary.

This autonomous and irreducible activity of imagination appears in other situations. In fact, it could be detected as early as the first definition of the first Book, which will serve me as a second basic example. As is well known, def. 1 of the *Elements* defines the *point* as "what has not part". A traditional difficulty here is to determine how the geometer could represent in a spatial diagram, endowed *de jure* with infinite divisibility, an indivisible entity (a problem which, like for the infinite line, immediately transfers to higher dimensions: length without breadth, surface with length and breadth only). This is a symmetrical, and more traditional, problem than the one posed by I.12: we have here a clear concept, but no possible image of it. What is interesting is Proclus' answer, which once again amounts to stress the irreducible role of imagination in its dynamical aspect (*phantastikès kinesis*):

> But someone may object: How can the geometer contemplate a partless something, a point, within the imagination if the imagination always apprehends things as shaped and divisible? For not only ideas in the understanding, but also the impressions of intellectual and divine forms, are accepted by the imagination in accordance with its peculiar nature, which furnishes forms to the formless and figures to what is without figure. To this difficulty we reply that the imagination in its activity is not divisible only, neither is it indivisible. Rather it moves from the undivided to the divided, from the unformed to what is formed. (*In Eucl.* 94–95)

Although it is not entirely clear what Proclus has in mind when claiming that imagination is "not divisible, neither indivisible" (literally: "not divided, neither undivided"), the insistence on the dynamical aspect is striking. It brings to the fore the role of the action (in this case: division) which generates objects (the indivisible), in contrast to the object given as such to *dianoia* and necessarily endowed with non-incompatible properties (either divided, or undivided). As images (unlike concepts-*logoi*, given solely by definitions), "points" carry information about what we can or cannot do (in this case: divide any further), although they are endowed with properties allowing us to do what they forbid.[19] What I would like to stress is,

[18] Reflecting in the first *Prologue* on the famous Aristotelian metaphor of the soul as a wax tablet, Proclus objects that it is rather a tablet "writing itself", a first occurrence of the idea of an active surface (*In Eucl*. 16, 10: γράφον ἑαυτὸ). On the autonomy of imagination in Proclus, see Claessens (2012).

[19] For a modern reading of "points" in Euclidean Plane Geometry insisting on the role of division, see Panza (2012, p. 73). According to Panza: "geometrical points are not represented by elementary diagrams. They are rather represented by extremities or intersections of lines, some of which are possibly elementary diagrams, whereas others are parts of such elementary diagrams resulting from dividing them through intersection". For the problems arising from such a view see Mumma (2012) quoted note 10.

once again, the relative opacity attributed to geometric imagination. It manifests itself in the fact that the dynamic of imagination is not presented as a projection of some discursive reasoning here. Quite the contrary since imagination allows us to circulate between properties ("not divided, neither undivided") which are conceptually incompatible.

In this regard, the discussion on the bisection of a line segment (prop. I.10) is also interesting. Proclus begins by recalling that the very act of bisecting seems to contradict a view on geometric entities as constituted by multiplicities of points.[20] A multiplicity being something which can be numbered, it would not be possible to cut in two parts a segment composed of an odd number of points (*In Eucl.* 278). But Proclus also refuses to consider infinite divisibility as an assumption entailed in the position of continuous magnitudes. All we need, according to him, is mere continuity described in terms of contact and possibility of action (division).[21] Infinite divisibility is then inferred as a consequence of this possibility by the (provable) fact that there exist incommensurable magnitudes (i.e. for which the alternate subtraction of one from the other cannot be a finite process cf. *Elements* X.2). Once again, we see imagination circulating between indivisibility and divisibility, but refraining from fixing these features into objective forms (a line constituted of points or an actual infinite division attached to continuous magnitude in and of itself). This makes it possible to relieve the traditional paradox of positing at the same time divisibility with no end and indivisible entities—or, better, in parallel to I.12, it transfers the paradoxical aspect to a dynamic of imagination which does not have to be fully transparent to conceptual determinations.

The infinite straight line and the point are however a very particular type of representations. If along with Peirce we define a diagram as a type of "icon" characterized by the fact that certain relations in the situation represented are carried by the *representamen*, they lead to the following paradoxical situation: the very nature of the *representamen forbids* us from mapping what are supposed to be the *characterizing* properties of the objects under study (non-divisibility for the point; actual infinity for the infinite straight line). For this very reason, it may look as if they constitute exceptions, related to the mysterious treatment of the infinite and indivisibility. As such, they could be treated with particular conventions of representation. What I would like to argue in the next section is that they are far from being isolated cases, although they may have pushed Proclus to explicate things which would have otherwise remained implicit (the autonomous position of geometric space and its relative opacity). Proclus' commentary is full of other diagrams which raise the same type of issues, and for a very good reason: so is ancient Greek geometry.

---

[20] An argument already put forward by Sextus Empiricus (*Adv. Math.* IX 282–283) and coming from the Epicureans, although they used it in an opposite strategy (in order to show that one cannot follow geometers in their claims) cf. Benatouïl (2010, pp. 156–157).

[21] This would go in the same direction as Panza's general discussion on continuity (see note 19 above and Panza 2012, Sect. 1.3.1, p. 72 sq.).

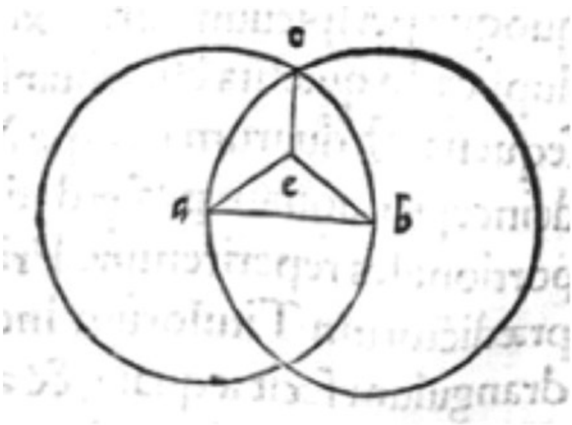


Fig. 1 Zeno's objection to I.1 (diagram from Barozzi 1560, p. 122)

2 Geometric Imagination in Practice

Let us look at the very first proposition of the first book of Euclid's *Elements* and what Proclus has to tell us about it. As is well known, this proposition asks to construct an equilateral triangle on a given finite straight line. The resolution goes on to construct the two circles whose centres are the extremities of the given segment and whose radius is the given segment. As was made famous by objections from modern geometers such as Pasch or Hilbert, a hidden assumption in the construction is that the two circles meet in a point which will serve as apex for the sought triangle (in Euclid's construction, only one of the two points of intersection is considered). What is less well known, however, is that a similar objection was already raised in ancient times. It is attributed by Proclus to Zeno of Sidon (*In Eucl.* 214), an Epicurean philosopher from the first century BC, and it deals not so much with the existence of the point of intersection as with its being well determined. How are we to be sure that we are not in a situation in which AC and AB have a segment in common and the apex of the sought triangle is not well determined? (Fig. 1).

The objection may strike us as odd, since it relies on a diagram in which straight lines are represented as not straight. Of course we know that geometric drawings are not to be taken as exact representations. But even if we don't pay attention to the drawing itself (as was certainly the case for Proclus who repeatedly dismisses any direct relation between mathematics and perception), the situation would remain problematic. Indeed one main difference between ancient and modern geometry is that some information has nonetheless to be retrieved from the diagram, be it interpreted as the concrete drawing or as some form of idealized counterpart.[22] These pieces of information typically feature inclusions of one region into

[22] For a survey on this question, see Manders (2008a).

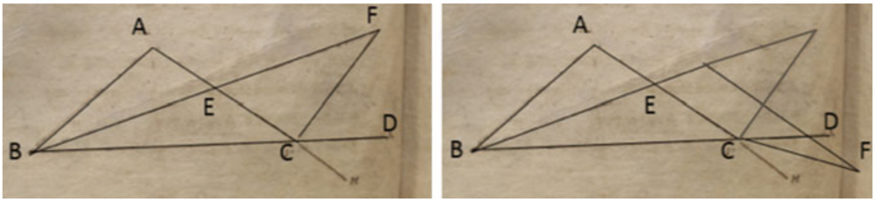


Fig. 2 *Elements* I.16 (*left*) and an alternative diagram with broken lines (*right*). As is well known, we do not have access to ancient Geometric diagrams and have to rely on late copies. For a survey on the question in the first books of Euclid's *Elements*, see Saito (2006) and more recently Saito and Sidoli (2012).

another.[23] Yet by allowing straight lines to be imagined as broken, we seem to run into serious difficulties when conducting this type of argument.

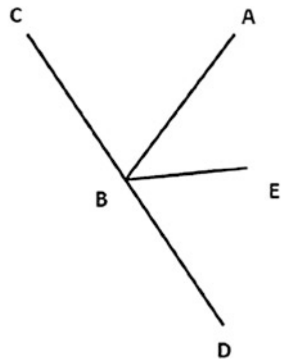
To make this point clear, let us read a typical Euclidean proposition. Consider, for example, prop. I.16 which states that "in any triangle, if one side is produced, then the exterior angle is greater than either of the interior and opposite angles". The demonstration asks us to bisect AC at E, join BE and produce it "in a straight line" to F. We then make EF equal to BE to obtain a triangle EFC similar to ABE, so that the angle BAE is equal to ECF. At that stage, we rely on the fact that ECD is greater than the angle ECF and we can therefore conclude that the angle ACD is greater than ECF (equal to BAE). The last information (ECD is greater than ECF) is a typical example of diagrammatic attribution (inclusion from one region into another leading to the conclusion that one is greater than the other). Let now suppose that we imagine another diagram allowing the "straight line" EF to be broken (as in the diagram on the right). In this case, angle ECF will contain angle ECD and a crucial step in the demonstration would not hold (Fig. 2).

Considering this situation, it may sound safe to reject the alternative diagram, and by the same token Zeno's objection, on the ground that the diagrams are not admissible. This is not, however, what Proclus does (nor, apparently, any of his predecessors).[24] He takes Zeno's objection seriously and answers by a proof. In order to do so, he first relies on the Euclidean definition of the straight line, which he takes to involve that the line is the shortest path and hence unique

[23] This belongs to the class of what Ken Manders has described as "co-exact" attributes: "The only claims based on diagram appearance in a demonstration recognize conditions that are insensitive to the effects of a range of variation in diagram entries: lines and circles that are not perfectly straight or circular, and cannot be taken to be without thickness. As we distort the 'circles' in I.1, their intersection point C may shift but it does not disappear. Such conditions I call co-exact. They include: part-whole relations of regions, segments bounding regions, and lower-dimensional counterparts" (Manders 2008a, p. 6).

[24] Other authors before him considered Zeno's objections and Proclus mentions a complete book written by Posidonius on this issue (see *In Eucl.* 216, 20).

Fig. 3 *Elements* I.14



(*In Eucl.* 215).[25] The reason for this strategy, in which the diagram is not rejected from the outset, seems to be not so much that Proclus realizes the difficulty hidden in the question of intersection and the interest of Zeno's objection, but that such a diagram *is* admissible as such in Euclid (and more generally in Ancient Geometry). Indeed in many demonstrations in which the *Elements* proceed *ad absurdum*, they involve situations in which straight lines have to be imagined as not straight.[26]

Take for example *El.* I.14 which states that: "If with any straight line, and at a point on it, two straight lines not lying on the same side make the sum of the adjacent angles equal to two right angles, then the two straight lines are in a straight line with one another". The demonstration proceeds *ad absurdum* by supposing that the two lines CB and BD, touching AB in B and satisfying the conditions on angles, are not on a straight line (although they are drawn in a straight line in the extant diagrams!). Then a line BE is introduced, which is supposed to be "in a straight line" with CB (but not represented as such!). One then shows that the resulting conditions on the angles are contradicted by the fact that the angle BE *is contained in* the angle BD. This last condition seems to work only because the "straight line" CBE has been imagined as broken and CBD not. It does not seem possible to get rid of this representation (for example, exchange the representation of CBD with CBE or represent both of them as broken lines) and maintain the argument as it stands (Fig. 3).

Another famous example of *reductio* is given by prop. I.27 which states that "if a straight line falling on two straight lines makes the alternate angles equal to one another, then the straight lines are parallel to one another". In the course of the

---

[25] After this first rejection, Proclus also proposes a *reductio ad absurdum* of the fact that two lines may have a common segment by showing that this contradicts the fact, which he demonstrated before, that a circle will be cut in two by its diameter (*In Eucl.* 216).

[26] In what follows, one should keep in mind that proofs by *reductio* are widespread in the *Elements*. By modern standards, their ratio in the total number of proofs seems even very high (around one fourth). Although not all of them involve absurd representations (see note 58 below), the latter constitute a significant number of them, especially in *Elements* Book III where they are numerous (nearly 50 % of the proposition are proven by *reductio*), see Vitrac (2012).

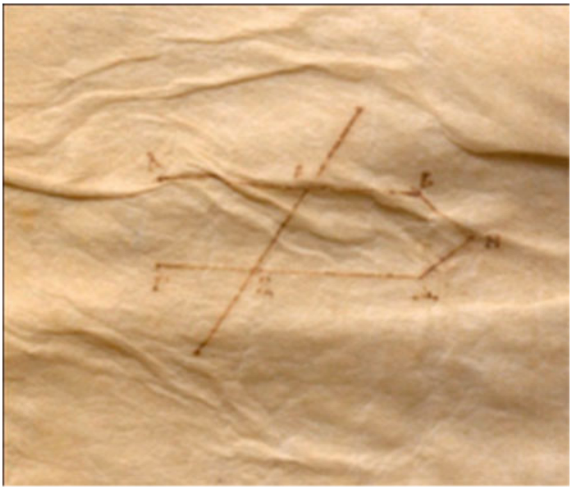


Fig. 4 *Elements* I.27 from the Bodleian copy (MS D'Orville 301)

proof, one assumes that the two straight lines meet and that they hence form a triangle with the one falling on them. The most common way to represent this situation is that presented in Fig. 4. In this representation, we see that considering a "triangle" whose sides are represented by broken lines, as I did in my alternative diagram to I.16 and as Zeno did in his objection to I.1, does not seem inadmissible.

At this point, one could object that this is not necessarily related to Proclus' view on the role of geometric imagination, since he is just here collecting objections without endorsing them or assuming that they are legitimate.[27] On this issue, let me first recall that "objection" is presented by him as a technical term of geometric discourse, on a par with "lemma", "case", "diorism", etc. (*In Eucl.* 212). According to the characterization given here, "objection" involves accepting counter-arguments attacking either the demonstration or the construction *without proof.*[28] In this sense, it is already unclear what an "illegitimate" objection would be and what grounds Proclus may have to reject an objection. It seems that any counter-argument based on an alternative diagram has to be accepted by the geometer, with the burden of proof lying on him. As we just have seen, there are good reasons for this principle of tolerance: considering the very functioning of *reductio ad absurdum* in the ancient geometric context, there seems to be no way of rejecting a diagram as illegitimate once and for all. A geometer who would answer to Zeno's argument by

[27] Think of the objections related to the fact that there may not be "enough room" around a given diagram (*In Eucl*. 225, 16; 275, 7; 289, 21), which seems to come from an "empiricist" interlocutor.

[28] "An 'objection' (*enstasis*) prevents an argument from proceeding on its way by opposing either the construction or the demonstration. Unlike the proposer of a case, who has to show that the proposition is true of it, he who makes an objection does not need to prove anything; rather it is necessary [for his opponent] to refute the objection and show that he who uses it is in error (*In Eucl.* 212).

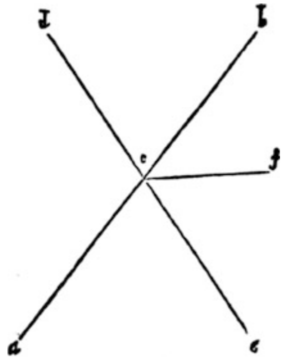


Fig. 5 Proclus's lemma in I.15 (Barozzi 1560, p. 172)

saying that what he represented as a triangle (or straight lines) could not be a triangle (or straight lines) would likewise have to block "regular" Euclidean proofs such as I.14 or I.27. On the other hand, it is clear that not all the alternative diagrams could be admissible, if we want the geometer to do something other than painstakingly answer any absurd configurations which could cross the mind of his objector.

I will come back to this issue below. Let me just note for now that even if it were the case that Proclus did not assume some of these objections as legitimate, it so happens that he also appeals in his own name to situations in which broken "straight lines" are involved. Commenting on propositon I.15, for example, he proposes to prove its converse by a *reductio* involving exactly the same type of diagram as the one used in I.14 (*In Eucl.* 302–305). In the diagram below CF is supposed to lie "on a straight line" with CD. Accepting this kind of representation naturally leads to accepting Zeno's objection as a serious one. What Zeno is doing is just transporting this type of representation, used by the geometer in some proof, in others (Fig. 5).

Another example of straight line represented as "non straight" and used by Proclus is given in the course of the commentary on I.4. There he criticizes the fact that Euclid assumed without proof that two lines cannot enclose a space (this is a crucial assumption in proofs by superposition such as I.4 or I.8 and was even incorporated in the "common notions" in some versions of the *Elements*).[29] Proclus undertakes to provide the missing argument and proceeds *ad absurdum* by relying on the following diagram (Fig. 6). In this case, straight lines are represented by arcs,

[29] Note that the problem is not unrelated to Zeno's objection. If we accept triangles with sides represented as "unstraight", we could make two triangles coincide on two of their sides and the angle contained by them without coinciding on the third. This will ruin the demonstrations "by superposition". The admissibility of such a diagram seems to be attested in the discussion about the curious "four sided" triangle, to which Proclus alludes in *In Eucl.* 329. Barozzi (1560, p. 189)

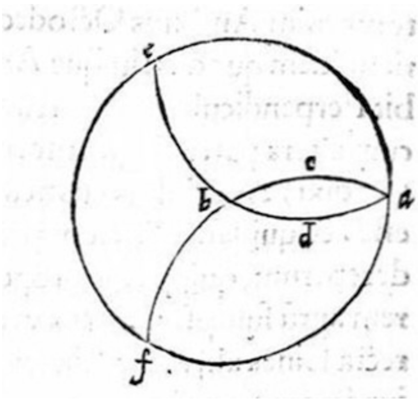
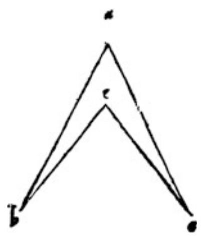


Fig. 6 Two "straight lines" which enclose a space (Barozzi 1560, p. 136)

a situation which is once again to be found in Euclidean proofs (see, for example, III.2)[30] and raises similar issues as broken lines do.

What I would like to emphasize with these examples is the following. If we pay close attention to the practice described by Proclus, be it in ordinary Euclidean proofs which he comments or in objections raised by him or by others, we see that there is nothing exceptional about the situation sketched in the preceding section. The very functioning of a *reductio* proof involves a clear distinction between two regimes which are not fully transparent to each other: geometric imagination on the one hand, which provides diagrammatic configurations which are admissible *prima facie*, and discursive reason on the other hand, which analyses these configurations.[31] The same holds for objections and, as I have shown above, there is an interesting interplay between the two situations: it does not seem possible to reject certain obviously "absurd" objections at first glance precisely because we need

(Footnote 29 continued)



With such a diagram, we could easily demonstrate the contradictory of I.4 by supposing that two triangles coincide on two sides (and the angle containing them) and then realize that one is nonetheless *contained* in the other. This possibility is ruled out by the fact that it would presuppose admitting two straight lines enclosing a space.

30 Diagram in Saito (2008).

31 I deliberately mimic Proclus' terminology (distinction between *phantasia* and *dianoia*, realm of the *logoi*). However, it should be clear at this point that "conceptual analysis" may involve information taken from the diagrams.

them to be admissible, at least hypothetically, in order to conduct regular *reductio* proofs.

All of these cases provide us with a geometric imagination which is therefore at the same time necessary, autonomous *and* opaque.[32] This could be understood in a quite straightforward way: imagination produces configurations which the conceptual process cannot recognize as legitimate. This is precisely why it may reject them after analysis.[33] We saw a more positive side of this opacity in the first section (it provides a possibility for representing infinite or indivisible entities), we now see its more negative side (although very useful for certain kind of proofs) and the dangerous game of objections without proof into which the geometer enters by allowing room for it in its practice.

If we take "figures" in a broad sense (including straight lines, points, angles, etc.) to be either the objects of ancient geometry or faithful representations of these objects,[34] the diagrams provided in our examples cannot be "figures" (or collections of "figures"): there is no such thing as a "triangle" formed by two straight lines and a third making equal alternate angles with them (*El.* I.27), no "circle" which cuts another in more than two points (*El.* III.10), no "straight line" joining two points of a circle and falling outside it (*El.* III.2), no "straight lines" enclosing a space or

[32] Considering the above example, it does not seem possible to claim, as Nikulin (2008, p. 160) does when presenting Proclus' concept of imagination, that "geometrical figures must be perfect, i.e. adequately represent their corresponding properties. Thus, a straight physical line is never straight, and a bodily circle is never round, whereas a geometrical straight line cannot be anything else but straight, and a circle nothing but perfectly round, which follows from their definitions. Therefore, sense perception, or *aisthesis*, cannot be the faculty responsible for the adequate representation of geometrical figures. Discursive reason, however, conceives geometrical objects in their properties as *logoi* which are not extended. This means that there has to be a distinct cognitive faculty capable of representing geometrical objects as figures, i.e. as extended and perfect.".

[33] This discrepancy between imagination and *dianoia* is paradoxically rendered obvious by recent reconstructions of Euclid's theory as a system in which the conceptual and the diagrammatic regimes are supposed to evolve in a perfect parallelism. In M. Panza (2012), for example, it is stated that the determination of the centre of a circle in *El.* III.1 is "in tension" with the rule of EPG (Panza's reconstruction of Euclidean Plane Geometry), because EPG provides "no possibility of constructing a circle without having previously constructed its centre, unless a rule for circles analogous to R.0 is admitted" (i.e. unless we accept that circles can be given as such, without any underlying construction). The problem is that we *need* the centre of a circle *not to be given* to perform the *reductio* proof in III.5 and III.6. In fact, the diagrams of III.5 and III.6 are simply not compatible with what Panza presents as a rule of construction for admissible diagrams of circles: "If two points are given, then two and only two concrete lines, each of which represents a circle having its centre in one of the given points and passing through the other, can be drawn" (p. 89). In Mumma (2006), many *reductio* proofs of book III concerning intersection and tangency (such as III.2 and III.13) are rejected from the outset, since the properties which they describe are consequences of the rule of construction of circles (see for example p. 26 where the convexity condition on representation of circles is stated). Same thing in the system proposed by Miller (2001), where III.2 and III.10 are presented as rules of formation for admissible diagrams (what Miller calls "nicely well-formed" diagrams, see Definition 2.1.5 p. 20 and Miller 2007, p. 26, Sect. 2, Definition 5).

[34] Calling a representation "faithful" when it preserves what are considered as the *characterizing* properties of the *representatum*.

having a common segment (without coinciding), etc.[35] This is precisely what *reductio* proofs prove. It is therefore legitimate to distinguish the figure, in the sense mentioned above, from the diagrammatic configuration presented to our imagination in order to conduct the proof.[36] The spatial proxy in which the *latter* occur is what Proclus calls geometric imagination and what he identifies with geometric space (usually two-dimensional space, since he deals here with plane geometry). One difficulty for the modern reader is that we tend to put geometric space and its various determinations in an objective position, fully captured by conceptual determinations (precisely because we tend to identify the consideration of space in mathematics with the fact of taking it as a proper object of study, see Sect. 1). This tendency can result in a considerable amount of confusion when attempting to comprehend a practice in which this was not done. Moreover, as I will try to indicate in the conclusion of this paper, it is not clear that our modern "structural" geometry relates to space *only* as an object or a system of objects. Quite on the contrary, one main feature of modern usages of space may well be that spatial configurations are used very generally not only as objects of study per se but as ways of acquiring information about objects of a non-geometric nature (in the sense in which, for example, a given non-geometric structure can be "equipped" with a topology).

*

Another interest of Proclus' testimony is to present us with a context of "objections and replies" which seems, from the examples collected and the names mentioned by him, quite widespread. The continuity between this dialogical context and the practice of *reductio* is nicely expressed in the discussion of I.7 where the Euclidean proof *ad absurdum* gives rise to the following discussion: "maybe perhaps some persons, notwithstanding all these scientific restrictions, will be bold enough to object and say that what our geometer calls impossible is possible" (*In Eucl.* 262, 5–6). The core of the objection is simply to present another diagram in which the conditions expressed in the Euclidean proof are not satisfied. Symmetrically, Proclus often replies to objections by showing that they lead to absurdity

[35] Note that this has nothing to do with the usual distinction between an imperfect physical drawing, either actually drawn or imagined, and a figure which would be its ideal counterpart. In our case, there is simply no ideal counterpart.

[36] See similar remarks by Manders, mentioned in the next section, against the "semantic" role spontaneously ascribed to diagrams and contradicted by the practice of *reductio*. See also Netz's remarks about the "make believe" elements contained in these proofs: one natural way to describe the situation is that we have a spatial configuration *pretending* to be such and such (a circle, a straight line, a parallel…) and conceptual analysis ruling out this hypothesis. Interestingly enough, Arab mathematicians had two different words for designating these two entities: the figure as representation and the figure as object, see Crozet (1999). This is related to an issue which I will not tackle in this paper, but which is another entrance in the question of spatial representations, that of the possibility of various configurations ("cases") representing one and the same geometric proposition.

(see the answer to Zeno *In Eucl.* 216, but also the answer to the objection to I.7 just mentioned *In Eucl.* 262–263).

More generally, this situation seems to be related to a common practice amongst practitioners in which one imagines a problem solved or a theorem proved and represents it in a diagram without knowing whether it is even possible to solve or to prove it—one of the meanings of "analysis" in a broad sense. In this regard, it is very striking that Proclus embeds his commentary on *reductio* in the following context:

> Every reduction to impossibility takes the contradictory of what it intends to prove and from this as a hypothesis proceeds until it encounters something admitted to be absurd and, by thus destroying its hypothesis, confirms the proposition it set out to establish. In general, we must understand that all mathematical arguments proceed either from or to the starting-points, as Porphyry somewhere says. (…) Those that proceed to the starting-points are either affirmative of them or destructive. But those that affirm first principles are called 'analyses' (…); when they are destructive, they are called 'reductions to impossibility', for it is the function of this procedure to show that something generally accepted and self-evident is overthrown. (*In Eucl.* 255–256)

Here we might also recall Pappus' famous description of "problematic analysis" in which he evokes the fact that, even in the process of analysis *stricto* sensu, one can stumble upon impossibilities.[37] An interesting piece of evidence for the continuity between these various aspects is given by "analyses" leading to cases under which a problem reveals itself "impossible" to solve. Although we do not have many testimonies on this practice in Ancient Greek Geometry (since we do not have many testimonies on analyses anyway), we do have some.[38] They tend to indicate an important role attributed to impossible configurations in the determination of the conditions under which a problem would be solvable. This corresponds to one of the senses of "diorism" whose purpose, according to Proclus, was "to determine when a problem under investigation is capable of solution and when it is not" (*In Eucl.* 66, 22).[39]

[37] Pappus, *Collection* 7.2.24–12: "We assume the proposition as something we know, then, proceeding through its consequences, as if true, to be something established, if the established thing is possible and obtainable, which is what mathematicians call 'given', the required thing will also be possible, and again the proof will be the reverse of the analysis; but should we meet with something established to be impossible, then the problem too will be impossible" (translation Jones 1986, pp. 82–83). Note that according to this description, only a *possible* geometric configuration can be considered as "given".

[38] Examples of this kind of analyses can be found in Apollonius (see, for example *Conica* II. 54) or in Eutocius commenting Archimedes (*On the Sphere and the Cylinder* II.4). On many occasions, including I.7 already mentioned, Proclus insists on the precision of the conditions expressed in the Euclidean propositions and mentions to this effect the impossibility encountered when they are not specified (*In Eucl.* 260–261).

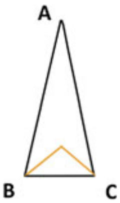
[39] See also *In Eucl*. 202, 3–8 where Proclus explains that this is the place where geometry asks questions such as: "does the object exist as defined?" This passage is important to counter a widespread view on ancient diagrams according to which they are supposed to attest to the existence of the objects: if I can ask if the object exists as characterized in a proof, it may happen that the answer is 'no' (if not, why ask?). I shall come back to this issue later on.

I emphasize these aspects, because they indicate that the opacity of geometric space is not the result of some metaphysical views here, but of a certain practice. Opacity is a natural outcome of ignorance and ignorance a condition of discovery and progress. When we pose a question without knowing the answer, it is normal procedure to imagine the question solved and see what will result from there. It may happen that we can reduce opacity into knowledge, but it may happen that we cannot. In other words, we have represented something which is *stricto sensu* impossible to conceive. This is still a very useful situation, since it can help us either to show that the contradictory proposition of an assumed theorem is true, or to rule out a problem as impossible to solve, or to specify conditions under which a more specific proposition/problem may be true/solvable. This general setting may help us to explain why geometers may have been led to accept objections *without proofs*. It is normal standard in geometric practice to propose situations in which we do not know in advance if they admit a solution or under which conditions they do.[40] This is part of the game and a condition of progress. Hence it seems a normal aspect of ancient geometric practice that one can produce a diagram without *knowing* if this diagram is consistent with the conceptual determinations given by the geometric discourse. This spatial configuration which is not already known in every respect to us is another way to designate what I have pointed to repeatedly as a form of "opacity".

It is worth noting that Proclus expresses no reservation at all regarding the practice described in the present section. As I explained before, he takes various objections seriously and accepts that an objection does not even need the support of a proof. When explaining the process of *reductio*, he raises no criticism whatsoever against this way of proving (*In Eucl.* 255–256). He even presents it as the most natural way to prove converse theorems.[41] This may sound puzzling at first, since

---

[40] Note that it seems to remain true in a modern setting. One can ask, for example, what is the shape of the right angle triangle built on the bisectors of a given isosceles triangle. It also looks isosceles, but is it *really*?



With a little reflection, one may come to the conclusion that such a triangle is, in fact, "impossible". This problem, which I took from a study in Mathematics Education, works in Ancient *and* "structural" presentations of Euclidean Geometry, see Richard (2000).

[41] See especially the commentary on *El*. I.19: "It was obviously from a desire to avoid complexity in the order of demonstration that the author of the *Elements* avoided this method of proof [scil. an alternative direct proof mentioned by Proclus], preferring to proceed by division and reduction to impossibility, because he wished to establish the converse of the preceding theorem without anything intervening. (…). It is preferable to prove a converse theorem by the reduction to impossibility while preserving continuity than to break the continuity with the preceding demonstration. This is why he almost always proves a converse by reduction to impossibility" (*In Eucl.* 321.9–20).

Proclus is also well known for having emphasized the role of geometric constructions as a testimony for the *existence* of geometric objects.[42] This is often taken as being a distinctive feature of the ancient epistemology of mathematics, as opposed to "modern" approaches in which the symbolic means no longer give us evidence for the existence of objects.[43] Moreover, Proclus strongly correlated the constructive aspect of geometric proofs with their explicative power and did not hesitate to criticize Euclid in that regard.[44] This question played a very important historical role in the debate over the nature of mathematical explanation, especially at the beginning of the Early Modern Age, and does not seem without relation to (the first?) strong rejections of *reductio* proofs.[45] It also plays a pivotal role in the widespread parallel drawn by modern commentators between Proclus' and Kant's forms of "productive imagination" (sometimes along with Descartes).[46]

Hence the need to remember that criticisms against *reductio* are *not* to be found in Proclus (whereas they play a crucial role in Descartes and Kant, for example). Moreover Proclus' "constructivism" should be balanced by the fact that it appears only in very specific examples. In the discussion on the fact that geometric proofs *can* be causal (202, 9–25), Proclus explicitly states, without expressing any form of discontent, that this is *not* the case for the proofs *by absurdum* so widely used by geometers.[47] When mentioning the fact that Euclid used "both proof founded on causes and proof based on signs", he hastens to add: "but all of them impeccable, exact *and appropriate to science*" (*In Eucl*. 69, 10–13, my emphasis).[48]

An interesting passage on the role of "porism" makes it clear that one should not conflate the role of geometric imagination in general with that of *construction*, if we understand the latter in terms of "geneses" expressing causal processes:

---

[42] See the famous comment on the fact that in the *Elements* the problems concerning construction of triangles precede the first Theorem (I.4), which Proclus comments in this way: "For unless he had previously shown the existence of triangles and their mode of construction, how could he discourse about their essential properties?" (*In Eucl.* 233–235). In her paper critically discussing the widespread "existential" interpretation of constructions in Euclid, Harari (2003, p. 5) recalls that "the main evidence in supporting the existential interpretation is found in Proclus' commentary on the first book of Euclid's *Elements*, where he accounts for the sequential priority of problems over theorems in existential terms".

[43] See Detlefsen (2005).

[44] See the famous discussion on *El.* I.32, mentioned in *In Eucl.* 206.12–26, and the related issue in the commentary of *El*. I.16 and I.17 (*In Eucl*. 309–312), cf. Harari (2008).

[45] Mancosu (1996).

[46] "The part played by imagination is Proclus' main addition to the Platonic theory, an addition which anticipates, it need hardly be pointed out, Kant's doctrine of schematism of the understanding" (Morrow 1992, p. lix). See also Bouriau (2000).

[47] "It is true that, when the reasoning employs reduction to impossibility, geometers are content merely to discover an attribute" (by contrast to establishing the reason for a given fact, *In Eucl.* 202.19–21).

[48] Note also what he says further: "if you add or take away any detail whatever, are you not inadvertently leaving the way of science and being led down the opposite path of error and ignorance?" (*In Eucl*. 69.27–70.1).

> Bisecting an angle, constructing a triangle, taking away or adding a length—all these require us to make something. But to find the centre of a given circle, or the greatest common measure of two given commensurable magnitudes, and the like—these lie in a sense between problems and theorems. For in these inquiries, there is no construction (γενέσεις) of the things sought, but a finding of them.[49] Nor is the procedure purely theoretical; for it is necessary to bring what is sought into view and to exhibit it before the eyes (δεῖ γὰρ ὑπ᾽ ὄψιν ἀγαγεῖν καὶ πρὸ ὀμμάτων ποιήσασθαι τὸ ζητούμενον). Such are the porisms that Euclid composed and arranged in three books. (302, 3–13)

This offers a nice setting not only to develop a more nuanced view of Proclus's concept of imagination, but also to understand why it appears at once in an active and a passive role.[50] Imagination allows the representing and exploring of the conceptual realm of objects with the resources proper to it. But at the same time it may offer some resistance to conceptualization. This last feature seems incompatible with the kind of productive imagination which was put forward by later philosophers—and is still quite widespread in various "constructivist" readings of Ancient geometry.

3 Philosophical Issues

Reflecting on the case of the *reductio* proof in Ancient Geometry, Ken Manders has claimed that traditional philosophical questions about the nature of geometric objects were ill posed because they assumed a semantic role of the diagrams which is simply *incompatible* with the practice of ancient geometers:

> Artifacts in a practice that gives us a grip on life are sometimes thought of in semantic terms —say, as representing something in life. There is, of course, an age-old debate on how geometrical diagrams are to be treated in this regard. Long-standing philosophical difficulties, on the nature of geometric objects and our knowledge of them, arise from the assumption that the geometrical text is in an ordinary sense true of the diagram or a 'perfect counterpart'. These difficulties aside, a genuinely semantic relationship between the geometrical diagram and text is incompatible with the successful use of diagrams in proof by contradiction: reductio contexts serve precisely to assemble a body of assertions which patently could not together be true; hence no genuine geometrical situation could in a serious sense be pictured in which they were. (Manders 2008b, p. 84)

From this "simple minded objection", Manders concluded that "the problem of the relationship between diagram and geometric inference here turns out to be one of standards of inference *not reducible in a straightforward way to an interplay of ontology, truth, and approximate representation*" (Manders 2008b, p. 86, my emphasis). This went along with a program of "inferential analysis of diagram-based geometrical reasoning" which resulted in a very nice and clear-cut result. Indeed, some inferences appear to be licensed by the diagrams and by the diagrams

[49] Note that the first example, the finding of the centre of a circle (*El*. III.1) is precisely the one which Panza found to be "in tension" with his reconstruction of Euclidean Plane Geometry (see note 33).

[50] On this issue, see Claessens (2012, especially p. 6, where the relevant literature is mentioned); Nikulin (2008, p. 164).

only (Manders coined the attribution on which they rely "co-exact"), whereas other inferences are licensed by the discourse and by the discourse only ("exact").[51] The main question is then to understand how the two inferential regimes (diagrams and text) can adjust so that one can have a good control over their interplay (especially as regards the range of admissible variations in the diagrams since the "co-exactness" criterion is not precise enough to rule out many apparently inappropriate configurations). We know that ancient geometry was not only a success, but also an efficient piece of machinery for producing results which are still considered, for good or bad reasons, to be true. But we do not know exactly why. This is one important aspect of what Manders has called the problem of "diagram control".[52]

Reviel Netz, in his seminal study on the shaping of deduction in Greek Mathematics, also put some emphasis on the *reductio* proof in order to discard a naïve "semantic" view of diagrams. But he reached a slightly more dramatic conclusion: "We seem to have reached a certain impasse. On the one hand, the Greeks speak as if the object of the proposition is the diagram. Verbs signifying spatial action must be taken literally. On the other hand, Greeks act in a way which precludes this possibility (quite regardless of what their ontology may have been!), and the verbs signifying spatial action must, therefore, be counted as metaphors" (Netz 1999, p. 54). He proposed to solve this puzzle by emphasizing the "make believe" element entailed in this kind of proof:

> The proof, of course, proceeds with the aid of a diagram. But this is a strange diagram: for good geometric reasons, proved *in this very proposition*, such a diagram is impossible. Euclid draws what is impossible; worse, what is patently impossible. For, let us remember, there is reason to believe a circle is one of the few geometric objects a Greek diagram could represent in a satisfying manner. The diagram cannot be; it can only survive thanks to the make-believe which calls a 'circle' something which is similar to the oval figure in Fig. 7. By the force of the make-believe, this oval shape is invested with circlehood for the course of the *reductio* argument. The make-believe is discarded at the end of the argument, the bells of midnight toll and the circle reverts to a pumpkin. (Netz 1999, p. 55)

[51] Exact attributes "are those which, for at least some continuous variation of the diagram, obtain only in isolated cases". The latter "are those […] which are unaffected by some range of every continuous variation of a specified diagram" (Manders 2008b, p. 92). For a detailed and critical discussion of these criteria, see Panza (2012).

[52] The whole Sect. (4.1) is entitled "Euclidean diagrams: artifacts of control or semantics?" (pp. 82–87). It is introduced in the following manner: "At its most basic, a mathematical practice is a structure for cooperative effort in *control* of self and life. In geometry, this takes many forms, starting with the acceptance of postulates, and the unqualified assent to stipulations—and as it appears, for now, to conclusions—required of participants. Successes of control may be seen in the way we can expect the world to behave according to the geometer's conclusions; the way one geometer centuries later can pick up where another left off; the way geometers can afford not to accept contradiction. When the process fails to meet the expectations of control to which the practice gives rise, I speak of *disarray*, or occasionally, *impotence*. Such occurrences are disruptive of mathematical practices; they tend to reduce the benefits to participants and to deter participation. At best, they motivate adjusting artefact use, modifying the practice to give similar benefits with less risk of disarray.".

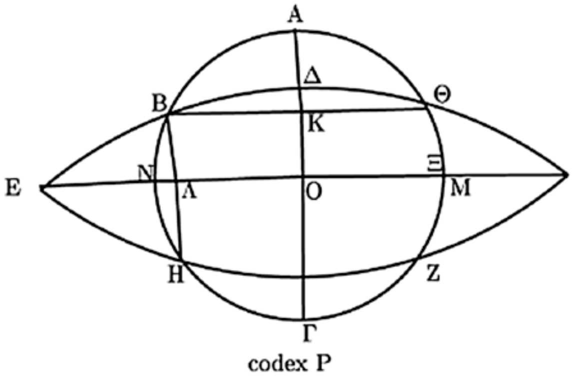


Fig. 7 Euclid's *Elements* III.10 (diagram from Saito 2008, p. 59)

But the "make believe" element, although undeniable and often mentioned when dealing with the *reductio* context, gives us no clue about the rules of the game.[53] Remember my alternative diagram to I.16: when are we to accept such a representation and when are we to reject it? There seem to be some fictions harder to swallow than others.

It is not sufficient to answer that "impossible" diagrams are only admissible in the framework of *reductio* (as does the "make believe" argument), since this is just begging the question: why not forge new proofs with new "impossible" diagrams and prove in this manner some propositions *incompatible* with Euclid's ones? This is the real question raised by my alternative diagram: by mimicking Euclid's reasoning, one could now show by *absurdum* that the exterior angle is *not* greater than one of either the interior and opposite angles. One just needs to suppose that it is greater, draw the alternative diagram following Euclid' instructions (but with the representation of a broken line) and reach the absurd conclusion that ECD is at the same time greater than ECF (equal to BAE) and contained in it. If one objects that we do not have the right to retrieve this last information from the diagram, we will reply that this is precisely the kind of information that Euclid retrieves from his own configuration. If one objects that we do not have the right to imagine a straight line in such a way, we will reply that this is what Euclid does in I.14. There is clearly a problem of "diagram control" here and this problem occurs *inside* the general regime of fictitious configurations (Fig. 8).

An hypothesis often mentioned is that such "impossible" diagrams may have been *temporarily* admissible precisely because they allowed the *excluding of* forms of representation and therefore progressively the limiting of the range of admissible variations. Diagram control and the *reductio* proof would then go hand by hand.[54]

---

[53] In a subsequent paper, Netz himself emphasized the fact that the role of imagination in ancient mathematics is quite widespread and not limited to impossible diagrams. In all of these cases, the "make believe" elements are of first importance (Netz 2009).

[54] "Diagram control theory invokes our ability, using geometrical constructions, to produce reasonably accurate physical diagrams, and so limit the diagram appearance outcomes to be considered by physical diagram production rather than discursive argument. Conversations with

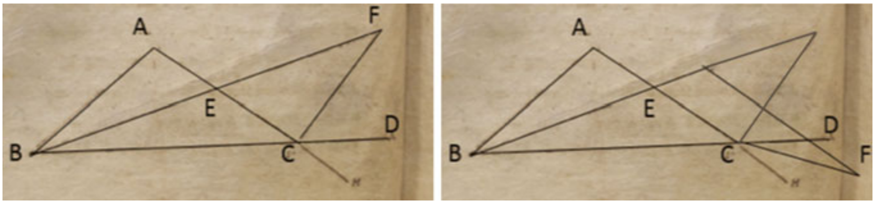


Fig. 8 *Elements* I.16 (*left*) and an alternative diagram with broken lines (*right*)

But we already know by the example of Zeno's objection that this does not seem to be the case: if we exclude the diagram in I.14 and not before, we will solve the difficulty for I.16, but not for I.1 (Zeno's objection); and if we reject it as early as I.1 (for example, following the reasons developed by Proclus), we would have to reject it in I.14. Helped by the larger view presented in the preceding section, we can easily generalize this dilemma. We just have to play the same trick as Zeno's by retrojecting one of the "impossible" diagrams taken from a *reductio* in a previous proof where it is still supposed to be, according to the above mentioned view, admissible. If one then objects that what it shows is that the diagram was already non admissible at that stage, the later proof from which it was taken will become itself non-admissible.

Take, for example, III.13 which states that "A circle does not touch another circle at more than one point whether it touches it internally or externally". It relies on a diagram of the following type (Fig. 9) in which we suppose circles touching in two points.

We can plug this diagram (for example for the case of circles touching internally) into a previous proposition such as III.6 which states that "if two circles touch another, they do not have the same centre". The result will amount to blocking what produces the contradiction in the known Euclidean proof, that is to say that ZE is shown to be at the same time equal to ZB (by the fact that the two circles have the same centre Z) and less than ZB (by inclusion in the diagram: this condition is satisfied in the diagram on the left, but not in the alternative diagram on the right, see Fig. 10).[55]

I leave it to the reader to play this game with other propositions. We have already seen that we can shortcut with broken "straight lines" (or "four sided triangles") the extant proof of I.4 or I.8, which are of pervasive use in the *Elements*; the same for I.16, which is needed for I.17 and I.18, themselves used in I.19 and

(Footnote 54 continued)
specialists suggest this is the basic tool of ancient practice, *with reductio argument* for the exclusion of putative alternatives as backup" (Manders 2008b, pp. 70–71, my emphasis).

[55] There may, of course, be discussions about whether to exclude the alternative diagram, for example by emphasizing that ZB and ZE are chosen in a particular position. But, even in this case, the demonstration would be different than the one given in the *Elements*.

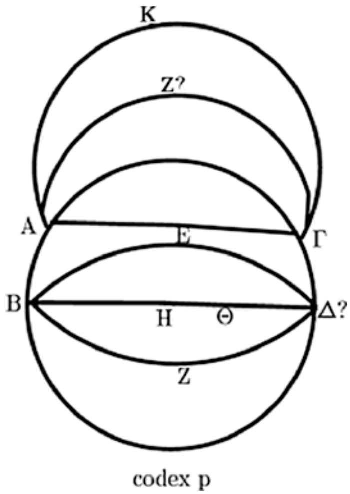


Fig. 9 *Elements*. III.13 (diagram from Saito (2008), p. 62)]

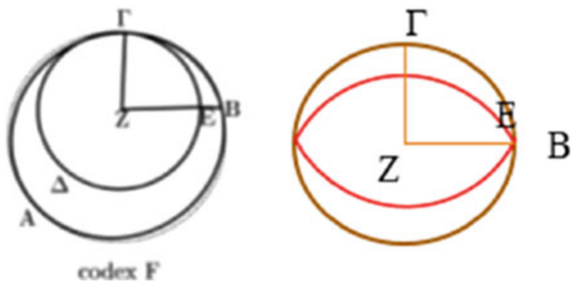


Fig. 10 *Elements*. III.6 and an alternative diagram derived from III. 13

I.20, etc. Many propositions of Book III rely on these results and when they do not, they are subject to the problems presented above in the case of III.6, which is easily generalized. Bit by bit, the whole edifice of the first books of the *Elements* appears in danger of collapsing, if not by contradiction, at least by multiplication of "impossible" diagrams to rule out at each step.

Of course, there is no real danger here. One will always be in the position to reject one of the diagrams presented so far through conceptual analysis. This is the very reason why we called them "absurd" or "impossible": they exhibit features which contradict the characteristic properties of the geometric configurations under study (in III.6, for example, it is easy to indicate by inclusion of one segment in another that we have *radii* in the inside circle whose distance to the centre have to be unequal). This is what Proclus first undertakes to show when he wants to answer an objection. But something is worth stressing at this point: if the result of conceptual analysis leading to this rejection is then considered to be a *rule* for the admissibility of diagrams, then many of the *reductio* used by Euclid will hence have to be *rejected* according to that very standard.[56] If, on the contrary, some other norm *allows* the introduction of the diagram, it will also allow it in situations *in which it has to be rejected*. This may be, according to my account, one of the reasons why objections were taken seriously. This is also why it seems so difficult, if at all possible, to determine a set of rules that precisely fix the range of variation for diagrams in ancient Geometric practice. If no information were retrieved from absurd diagrams, this would be unproblematic. Unfortunately, as we have now seen in many examples, this is not the case.

Neglecting however the quest for a complete system of rules, there may be an easy way out of the preceding dilemma. Indeed one can always introduce an absurd diagram and answer to possible criticisms that one *knows* ("by concepts") that this diagram is absurd. In the case of *reductio*, the objection against the diagram would then not be sufficient to block the proof. Indeed we could easily reply that this absurdity is precisely what we want to establish. In other words, we would *agree* with our interlocutor on the "absurdity" of the diagram, which is just an illustration of the absurdity of the premise. If the objector still maintains that we don't have the right to assume such an absurdity, what she or he will be criticizing will not be the representation, but the structure of the proof. But this seems to imply that the "control" over diagrams will involve in a substantial manner some form of semantic regulation (an "illustrative" role for diagrams). In particular, it would impose further constraints on the "make believe" component: the absurdity of the diagram will be admissible if and only if it is directly related to (if it "illustrates") the claim which we want to refute. Yet it so happens that this condition is *not* fulfilled in my alternative diagram to I.16 (even when the proposition is recast into the form of a *reductio*) or in Zeno's objection to I.1. In both cases, the absurdity of the diagram is not related to the particular claim in play. We hence have no reason to accept it if it enters in conflict with conceptual analysis and this is precisely what Proclus intends to show.[57]

[56] Not all of them however, since they do not all rely on "impossible" diagrams in the sense mentioned here. I.6, for example, represents two equal segments as unequal, which is *normal* practice. Even I.27 does not seem to be inadmissible since we do not really need the absurd diagram (with broken lines) and can just draw two straight lines which meet in a "regular" triangle (as is drawn in one of the diagrams in the Bodleian copy, MS D'Orville 301). These examples should not be put on the same footing as I.14 or III.2 (*pace* Panza (2012), p. 82, n. 53).

[57] Let me make this argument more explicit. In the case of Zeno's objection, Proclus' first reaction is to reply that the diagram contradicts the condition imposed on what it is to be a straight line. As I have tried to argue above, this reply does not justify an immediate rejection of the diagram and this is no surprise since we need this diagram to be admissible in other cases (typically I.14). The question is then: how is it that the diagram of I.14 is not ruled out on the same grounds? My interpretation is that the diagram remains admissible in I.14 after conceptual analysis, because what it shows is precisely what the proposition *claims*: it is not possible to draw a line satisfying the conditions on angles and not being "in a straight line" with the given one. In any *reductio* of that form, we will be forced to represent a "straight line" by a non-straight one: this is no objection, this is the claim!.

Of course, one will always be in a position to emphasize the "make believe" elements and claim that there is no real absurdity involved in this kind of representation. Of course, one is just *pretending* that what is apparently not a circle or a straight line in the diagram is a circle or a straight line. This does not seem different from claiming that what does not appear as an indivisible entity or an infinite one is. Since points without part and lines without breadth are pervasive in plane Geometry, what we are describing here is just the *regular* functioning of the geometric imagination. These are strange representations, since they exhibit features incompatible with what they represent, but this is simply the *general* problem of geometric representation. Moreover, the "inferential analysis of diagram-based geometrical reasoning" helps give a precise contour to a fact which is widely accepted: in geometry, we do not need representations to be faithful, but trustworthy. For that purpose we only need diagrams to carry certain kind of information. There is no more difficulty in the fact that they exhibit properties incompatible with the property of the objects they represent than in the fact of writing "red" in green ink and "green" in red ink. So far so good. But I would like to emphasize that this general view leaves the problems raised above *untouched*. What seems to have escaped attention until now is that *the same diagram* must be admitted in some proofs and rejected in others. In other words, diagrams of the sort that we are dealing with are *not* trustworthy. As a consequence, the problem of "diagram control" does not seem to be solvable by fixing a range of variation once and for all (since the same variant will have to be admissible in some proofs and not in others); neither will it be solved by stating a set of fixed rules of construction.[58]

More generally, many questions are raised by our detour through Proclus about the role of these "constructions" in Euclidean Geometric practice. If we take *El.* III.6 seriously, we can draw a circle without knowing where its centre is. If we take *El.* III.2 seriously, we can join two points on a circle by a straight line without knowing where the straight line stands relative to the given circle. How would these claims be consistent with a constructive reading of lines and circles?[59] More generally what do I "construct" when imagining two "circles" intersecting in more than two points or touching internally in more than two points? What do I "construct" when representing a "straight line" which does not coincide with the line making angles equal to two right angles at a given point on another straight line? Many modern philosophical interpretations of Ancient Geometry seem too imbued with a view in which constructions are essential because they are supposed to give evidence for the possibility of concepts and/or existence of objects. But this is far from obvious as soon as one realizes that it was standard practice in Ancient

[58] This is also why we cannot be fully satisfied with overly general rules such as the Euclidean postulates, which allow "impossible" diagrams without specifying the cases in which they have to be rejected. If the first postulate allows us to represent a "straight line" by a broken line, as it seems, or if the third postulate allows us to represent a circle with an eccentric "centre", they allow "constructions" which will have to be *prohibited* in some other proofs.

[59] Remember that postulate III asks us to draw a circle *from any centre* and radius: Καὶ παντὶ κέντρῳ καὶ διαστήματι κύκλον γράφεσθαι.

Geometry (and not only in Ancient Geometry!) to imagine a situation which is not known in advance to be even *possible* (i.e. "constructible", if possibility and construction of objects are equated). In my view, a recognition of what I have designated (following Proclus) as the "opacity" of space and the autonomous functioning of the geometric imagination offers a more promising account.

Another way to designate this opacity would be to stress that ("spatial") images and ("geometric") concepts do not evolve in a perfectly parallel manner and that this is not something that we have to fix. It is true that this discrepancy may sound strange at first. As I have recalled above, the usual expectation would be that spatial images ("diagrams") are, if not faithful, at least, trustworthy. Unfortunately, this does not seem to be the case, at least in a straightforward way. But much of this strangeness may vanish if one realizes that the coupling of heterogeneous entities that relate one to the other, but do not evolve in perfect parallelism is a common situation in semiotic systems.[60] In this regard, it is very interesting to note that, up to now, nobody seems to have succeeded in coming up with a complete system of rules fixing the range of the tremendous variations occurring in the writing and/or pronouncing of one and the same phoneme in a given language. The complexity of automatic recognition systems are good evidence for that. They still rely massively on the statistical method and on a control operated a posteriori by the knowledge of the linguistic "content"—as opposed to systems of rules fixing a priori the variation of graphemes or phonemes in and of themselves.[61] This is not the place to enter into the fascinating questions related to these systems of co-variation, but let me just emphasize here that stabilization does not seem to occur in symbolic systems *only* through a complete and fixed system of parallel rules (although there *are* some structural rules on both levels). It also usually involves in a substantial manner the *interplay* between the different levels. As I have tried to indicate in this section, this seems to be the kind of "control" which is at stake in Ancient Geometric practice. The fascination for "formal language", in which this interplay is carefully shortcut, and the (false) conviction that it offers a model for any symbolic systems used in mathematical sciences may explain why we still have difficulty applying this description to mathematical symbolic systems.

This way of approaching Proclus by laying emphasis on the interplay between image and concept is of particular interest when faced with the epistemological issues which we have encountered so far. Let me first summarize them. On the one hand, we have the fact that geometric texts talk about something which, on their

---

[60] Beginning with the coupling of acoustic images and concepts in natural languages, one of Saussure's deep insights when launching the "structuralist" approach to language, see Maniglier (2006).

[61] Crettez and Lorette (1998). Even in very standardized systems such as typographical ones, we encounter not only important variations for the same grapheme, but also graphical signs which can be very similar for different graphemes (such as  and , for example). Symmetrically, one could note the similarity of graphical signs used in one and the same system for representing different graphemes (such as  and ).

surface grammar, *is* the diagram;[62] the *reductio* proofs, however, make it clear that the surface grammar is misleading here and that geometric texts cannot talk about the actual configuration which they exhibit; at the end of the argument, as Netz puts it, "the bells of midnight toll and the circle reverts to a pumpkin". Call this the semantic problem or Netz's "impasse". On the other hand, one could declare the approach in terms of "objects" to be misleading and consider that one has to focus mainly on the inferences carried with the help of our two resources: texts and diagrams. In the absence of semantic rigidity, what we need now is a good adjustment of two inferential regimes which are presented as evolving in a parallel and complementary way. The problem, which I have been pointing out regularly in this paper, is that the variation of diagrams does not appear to be *intrinsically* regulated: one and the same diagram has to be accepted in one proof and rejected in another; moreover, this seems to depend on the "content" of the proposition in play, that is to say… on a "semantic" regulation. Call this a (dramatic) variant of the problem of "diagram control".

Proclus's position falls so to speak in between these two options. It may even help us to rephrase them in a more positive way. Let me, for example, restate the first dilemma in a Proclean manner: contrary to a lazy and narrow Platonist interpretation, it is not possible to say that Geometry deals only with ideal figures, the diagram being just a dispensable auxiliary. This is made obvious, amongst other reasons, by cases in which one has to represent a geometric situation to which *no* ideal entities correspond in the conceptual setting (be it an infinite straight line, a circle intersecting another circle in more than two points, two straight lines enclosing a space, etc.). However, and for the very same reason, it is not possible to avoid the *distinction* between entities characterized by concepts and diagrams. Remember that at the end of the argument, the circle reverts to a pumpkin. In other words, the circle, as presented by the definition in the text (and propositions exhibiting such and such of its properties), has to be something *different* from what was represented as such in the diagram. There is no impasse here if one accepts that mathematical knowledge stabilizes itself in the *interplay* between these two regimes which Proclus calls "discursive reasoning" and "imagination".[63] This means to accept that imagination has a form of autonomy and does not limit itself to

[62] This aspect is documented in great detail in Netz's book. This relation is the basis of the "ontological" issues then engaged in order to assess the link between concrete diagrams and possible abstract objects.

[63] The interest of a semiotic approach is to detach the role of imagination from its geometric origin and reveal its more general nature. When, for example, one supposes in abstract algebra that the centre of a p-group is trivial (reduced to identity) in order to show that it leads to absurdity, the usual proof proceeds as follows: one decomposes the group in its conjugacy classes and obtains an equation of the form $p^k = 1 + p^i + p^j + \cdots + p^m$ (p^k being the order of the group). This latter formula is an "impossible diagram" (since it represents a p which is supposed to divide both side of the equation) in much the same sense as the one we encountered in this study (see van der Waerden 2003, p. 153 for this classical proof in abstract algebra). The opacity of geometric space transfers immediately to the opacity of symbolic writings, as was remarkably seen by Leibniz who called both of them "characters" and associated them with "symbolical" or "blind" knowledge.

illustrating some purely conceptual process. As regards the second problem, Proclus clearly emphasizes the dynamical aspect internal to the two regimes and the fact that they provide complementary systems of inferences, *but* he also points out another aspect of particular importance: the fact that imagination occasionally provides situations which are opaque to knowledge. As I have tried to document above, this is another very important role of diagrams: they not only help us to prove propositions, but also to represent situations which we do not know in advance to be possible. In this sense, the control cannot stem from an internal system of rules fixing the kind of information which can be retrieved from diagrams. It also needs the regulation of conceptual knowledge.

4 Conclusion

As a conclusion, I would like to sketch, as announced in the title of this paper, what I take to be the actuality of some of Proclus's insights on geometric space. Before doing so, let me summarize the principle results of this study. We have seen that geometric imagination (*phantasia*), which Proclus identifies with geometric space or a form of "receptacle" on which discursive thinking (*dianoia*) projects its conceptual determinations, is not limited to a passive role of illustrating or picturing. It also has an autonomous activity, which manifests itself in the form of a possible opacity to conceptual knowledge. The metaphor comes directly from Proclus who mentions that sight can paradoxically recognize what escapes its power: darkness. Between complete obscurity, evoked in the case of the diagrammatic representation of actual infinity, and the bright light of concepts stand many other situations in which some parts of our representations are clear and some are not, many forms of *chiaroscuro*, so to speak. Moreover, not only can we picture darkness, but we can picture *with* darkness in order to make what has to be visible visible. What I have tried to do in the second section of the paper is to document this art of *chiaroscuro* in Proclus and show that it seems to correspond to a standard practice in Ancient Geometry. This allows us to provide to Proclus' conceptions not only with an immediate context, but also with a first actuality. When related to geometric practice, Proclus conception of space appears as a fruitful framework that can be used to solve certain difficulties encountered in philosophical reconstructions of ancient geometry, because they don't pay enough attention to this specific functioning of geometric imagination. What I have tried to emphasize in the third section is that this may even help us to overcome certain epistemological "impasses" in which recent studies seem to be stuck.

This overall picture conforms to two features of Proclus' philosophy which I had no space to develop in this paper, but which are of tremendous importance to situate his thought. First, Proclus, along with Jamblichus, disagreed with other members of the Neoplatonic School, such as Porphyry, on the role of logic in its relation to mathematics. The latter considered, in an Aristotelian vein, that logic is a universal science, attached to the structure of predication and governing the rules of reasoning, which has therefore to be learned *before* any particular science, such as mathematics. The former argued that mathematics is the proper place to learn the general art of reasoning ("dialectic"), which extends far beyond what logic in the above sense can teach us.[64] This is related to the fact that mathematics involves other elements of discursive thinking than mere deduction, such as definitions, divisions (of cases) or analyses.[65] A second distinctive feature of Proclus' philosophy concerns where he diverges from Jamblichus, that is: on where to situate the universal mathematical science. Whereas the former conceived of it, following a Neopythagorean tradition, in close connection to arithmetic, Proclus and his pupils considered it to be closely connected to Geometry. This is made clear in Marinus' commentary on Euclid's *Data* where he states that this treatise does not belong to any particular mathematical theory, but to the "universal mathematics"; he then emphasizes the fact that Euclid has presented other aspects of this general mathematics in Book V of the *Elements* (dealing with ratios and proportions), but *under the guise of a geometric presentation.*[66]

As I have tried to argue elsewhere, we have here different visions of what it means to give "foundations" to mathematics.[67] It would not seem exaggerated to state that the first direction was the leading option in terms of foundations from the middle of the nineteenth century to the middle of the twentieth century. It is not surprising that it accompanied a debate focusing on the disagreement between those who held that mathematics starts with logic and those who held that it starts with (basic) arithmetic and systems of numbers. Although there were always mathematicians who resisted that general tendency (and even more mathematicians not interested in foundational issues!), not many of them protested that geometry was the proper place for foundations. Things began to change significantly in the 1960s, when the unexpected relationship between geometry and logic began to emerge. In a philosophical "manifesto" entitled "Logic as a geometry of cognition" Jean-Yves Girard, a leading protagonist of this evolution, has called it—following a suggestion

---

[64] See Jamblichus *De com. Math.*, Chap. 29; Proclus, *In Eucl*. Chap. XIV and 69, 8 sq; for a commentary: O'Meara (1989), pp. 47–48 and Chap. 8.

[65] One may find surprising that Proclus credits Euclid's *Elements* for exhibiting forms of analysis, since according to the standard picture of the treatise, it seems the prototype of the *synthetic* method. But I hope to have given elements to better understand this claim: in a broad sense, analysis designates any way back to the principles. This is not what Proclus calls *analysis* strictly speaking in *In Eucl*. 255–256, but this broader sense is clearly stated in other places such as *In Eucl.* 8.9 and 57.19. In these passages, analysis is characterized more loosely as the method of proceeding from complex to simple, from things we seek to know to things better known. As regards division of cases, it should be noted that it is another entry into the question of the discrepancy between spatial representation and geometric object, which I could not deal with in this paper. Interestingly enough, this path was followed by Arab mathematicians who distinguished several meanings of "figures", see Crozet (1999).

[66] Marinus, *Com. in Euclidis Data*, 254, 5–27 and Rabouin (2009).

[67] Rabouin (2009), Chaps. I–III.

by Samuel Tronçon—a "geometric turn" (by contrast to the "linguistic turn").[68] I cannot resist quoting the incipit of this provocative paper:

> I. 'Les Grands cimetières sous la lune'
> To place philosophy again at the centre of scientific activity, to rehabilitate philosophy of science, what a program! In order to do so, we propose to reactivate the central tool of *logic* by extracting it from the narrow path of the 'linguistic turn'; this reactivation would operate with the tool of *geometry*, a 'geometric turn' so to speak (p. 15).[69]

This is not the place to comment upon this evolution, which has many facets and parallel developments. But it is certainly a task for the philosophers of our time to understand what it might signify. Supporters of the "logico-arithmetical view" had a very nice story to tell about their construction, a story going all the way back to Aristotle and justifying the foundation in terms of building from the more general to the more particular (whether in terms of sciences dealing with general forms, as opposed to science dealing with specific domains of objects, or in terms of domains of objects simpler than others and needed for their construction). First logic, then arithmetic, then geometry—the pending question being when exactly does mathematics start in this overall picture (think of the ambiguous status of Set Theory). But what kind of story, if any, could support the universality of geometry?

Proclus is not very explicit about this issue, but his general strategy makes it clear that in his eyes, geometry exhibits universal structures which not only can be useful for dealing with continuous magnitudes *and* numbers, but can also enrich the tools considered by logic (tools which remain invisible as long as one only considers the structure of predication and of direct deduction). This directly contradicts the picture according to which the "above" science is completely independent from the sciences "below". This is, so to speak, a universality "from below", from the point of view of what is *transversal* to the whole science considered. The point which I would like to insist on is that this view seems to be related to the issues tackled in this paper. Indeed, it amounts to considering spatial configurations not only as related to domains of objects, as is the case in geometry *stricto sensu*, but also as useful *tools* for studying other kinds of mathematical objects, *including* geometric objects themselves! In this picture, space acts in an ambivalent position: either as a framework in which one studies objects or as a means with which we study them. Call it the passive and the active role of geometric space. Moreover, as a tool, space need not to be fully transparent to one of the identified domains of objects.

It is worth noting that this kind of story about the universality of space was heavily emphasized in recent times by mathematicians involved in the "geometric turn". Alexander Grothendieck, when commenting on his idea of introducing "generalised spaces" declared:

---

[68] Girard (2007).

[69] My translation, except for the title of the section which is taken from a book by Georges Bernanos and which I left in French.

> La notion d'‘espace’ est sans doute une des plus anciennes en mathématique. Elle est si fondamentale dans notre appréhension ‘géométrique’ du monde, qu'elle est restée plus ou moins tacite pendant plus de deux millénaires. C'est au cours du siècle écoulé seulement que cette notion a fini, progressivement, par se détacher de l'emprise tyrannique de la perception immédiate (d'un seul et même ‘espace’ qui nous entoure), et de sa théorisation traditionnelle (‘euclidienne’), pour acquérir son autonomie et sa dynamique propres. De nos jours, elle fait partie des quelques notions les plus universellement et les plus couramment utilisées en mathématique, familière sans doute à tout mathématicien sans exception. Notion protéiforme d'ailleurs s'il en fut, aux cents et mille visages, selon le type de structures qu'on incorpore à ces espaces (Grothendieck (1985–86), p. 52).[70]

Interestingly enough, it so happened that the concept of *topos* revealed itself as one of the forms in which logic and geometry appear as intimately related one to each other.[71] My intention, once again, is certainly not to comment upon these highly technical questions, which I am far from fully understanding, nor to claim that we have here the “right” candidate for foundations in mathematics. The mere fact that we have already two very different proposals (Girard, as he explains in his paper, is influenced by non-commutative geometry, not by topos theory),[72] and in fact two *amongst many others*, should make it clear that “foundations” is a *locus* of an internal debate in mathematics—a debate which structurally appears to have no definitive winner—and not a place for philosophers to validate (let alone dictate!) what should be the “right” set of choices. My only aim is to give a context where the questions raised in this paper could find interesting prolongations. In the quote from Grothendieck, one is surprised by the presence of many ideas, which are not so common in philosophical views on space, but are presented as widely accepted by mathematicians: the fact that space is *universally* present in mathematics; the fact that it has an *autonomy* and a *proper dynamic*; the fact that it is *proteiform*, precisely because it is not limited to such and such a structure under study, but changes faces by “incorporation” with other structures. All of these insights played a crucial role in the arguments developed in this paper.

---

[70] The new idea of space or topos is then clearly presented as unifying the realm of continuous magnitudes and numbers: “Cette idée englobe, dans une intuition topologique commune, aussi bien les traditionnels espaces (topologiques), incarnant le monde de la grandeur continue, que les (soi-disant) ‘espaces’ (ou ‘variétés’) des géomètres algébristes abstraits impénitents, ainsi que d'innombrables autres types de structures, qui jusque-là avaient semblé rivées irrémédiablement au ‘monde arithmétique’ des agrégats ‘discontinus’ ou ‘discrets’.” (Grothendieck (1985–86), p. 54).

[71] “A startling aspect of topos theory is that it unifies two seemingly wholly distinct mathematical subjects: on the one hand, topology and algebraic geometry, and on the other hand, logic and set theory. Indeed a topos can be considered both as a ‘generalized space’ and as a ‘generalized universe of sets’. These different aspects arose independently around 1963: with A. Grothendieck in his reformulation of sheaf theory for algebraic geometry, with William F. Lawvere in his search for an axiomatization of the category of sets and that of ‘variable‵ sets’, and with Paul Cohen in the use of forcing to construct new models of Zermelo-Fraenkel set theory.” (Mac Lane and Moerdijk 1992, p. 1).

[72] He also makes it clear that he is not criticising the linguistic turn, which was so fruitful for logic, but is proposing to complete it.

Another way to state the "proteiform" nature of the notion of space would be to simply say that we do not know exactly what space is. Contrary to what is too often assumed by philosophers (by relying on such and such examples: metric spaces, manifolds, topological spaces, *topos*…), space does not have "a" fixed structure.[73] It is not transparent to concepts. As we can see in Grothendieck's passage, this has nothing to do with some form of metaphysical assumption, but with the very functioning of the geometric imagination which pervades all of mathematics. As soon as we use space in order to know some other structure, we open two important possibilities: first of all (and this was the general case in ancient Geometry), we may us space in order to know… spatial entities (think of my example with the real line above); this leads immediately to an irreducible *duality* between space as represented and space as tool of representation; second, the fact that we use space in order to know other mathematical objects structurally involves a form of opacity: in the case when we do not obtain a fully transparent representation, there is no way to know if the difficulty is due to the tools or is a structural impossibility, a problem coming from some restricted condition or, on the contrary, some excessively loose stipulations in the data of the problem, etc.[74]

In this regard, I would like to conclude this paper by mentioning a very interesting inquiry undertaken by the mathematicians John Gratus and Timothy Porter in a paper entitled: "A Geometry of Information".[75] Let me just quote the very first sentence of their study: "Spatial representation has two contrasting but closely related aspects: (i) representation *of* spaces and (ii) representation *by* spaces".[76] The paper is very interesting for what it proposes as a unifying structure for these "spaces" of various sorts. But it is also interesting as symptom: in 2005, it was still possible to consider that mathematical space has a dual nature, neatly expressed by the distinction between "representation *of* spaces" and "representation *by* spaces";

---

[73] A very basic example may render this clearer: if we take the usual "real line", we may have the impression of being faced with a simple form of one-dimensional metric space. But this depends on the choice of what we take as basic open sets. If, for example, we consider the topology based on open sets of the form $[0, a]$ for any real a in the unit interval $[0, 1]$, we obtain a simple example of a topology which is not separated (hence not metric, any metric space being separated). This last feature contradicts what were often considered as the characterizing properties of "spaces" (being *partes extra partes*). Moreover, it raises the question: what is *the* spatial form of the real line?

[74] One can think, for example, of the story told by Grothendieck in the passage quoted above. Its background was the way in which algebraic geometers stumbled upon difficulties in their use of usual topologies and were led to introduce more exotic ones, such as Zariski's topology. At the end of the process, it is topology itself which revealed itself too narrow and had to make room for "topos" and "sites".

[75] Gratus and Porter (2005).

[76] This duality is then explained in the expected manner: "The first is, classically, based firmly in geometry, and topology and assumes some 'space' is given, whilst its aim is to study the 'attributes' of the space - essentially its geometry and topology, or more precisely those parts that are amenable to study by the usual tools of geometry and topology! The other aspect represents some configuration by a space. This 'configuration' may be a formal situation modelling some relationship between some objects and attributes, or perhaps a physical context such as the space of physical configurations of a molecule".

but, more interestingly, one could realize on those grounds that it still has no unified conceptualization. Although it would be wildly exaggerated—if not totally ridiculous—to claim that Proclus' philosophy could help us answer these questions, which are, of course, the prerogative of mathematicians, it can nonetheless aid us in understanding *why* they are still pertinent.

Acknowledgments I would like to thank Marco Panza who pushed me (quite literally!) To present my views on proclus to the workshop "diagrams in mathematics", which he organized with Reviel Netz in 2008 at Stanford University. It gave me an initial occasion to develop some clumsy remarks about proclus's relevance and exchange views with a wonderful audience including Michael Friedman, Sébastien Gandon, Ken Manders, John Mumma or Ken Saito. The written version was read by Sébastien Gandon, John Mumma, Reviel Netz and Marco Panza, who helped me to improve it by many valuable comments.

References

Ariew, R., & Descartes, R. (Eds.). (2000). *Philosophical essays and correspondence*. Indianapolis: Hackett-Publishing.

Barozzi, F. (1560). *Procli Diadochi Lycii ... In primum Euclidis elementorum librum commentariorum ad vniuersam mathematicam disciplinam principium eruditionis tradentium libri IIII.* Padua: Gratiosus Perchacinus.

Benatouïl, T. (2010). Les critiques épicuriennes de la géométrie. In P. E. Bour, L. Rebuschi, & L. Rollet (Eds.), *Construction, Festschrift for Gerhard Heinzmann*. London: College Publication.

Bouriau, C. (2000). *Aspects de la finitude, Descartes et Kant*. Pessac: Presses Universitaires de Bordeaux.

Breton, S. (1969). *Philosophie et mathématique chez Proclus*, suivi de *Principes philosophiques des mathémathiques d'après le commentaire de Proclus aux deux premiers livres des Éléments d'Euclide* (transl. G. De Pesloüan). In N. Hartmann (Ed.) Paris: Beauchesne.

Charles-Saget, A. (1982). *L'Architecture du divin*. Paris: Les Belles Lettres.

Claessens, G. (2012). Proclus: Imagination as a symptom. *Ancient Philosophy, 32*(2), 394–395.

Cleary, J. J. (2000). Proclus' philosophy of mathematics. In D. J. O'Meara & G. Bechtle (Eds.), *La philosophie des mathématiques de l'Antiquité tardive* (pp. 85–101). Fribourg: Éditions universitaires de Fribourg.

Crettez, J.P., & Lorette, G. (1998). Reconnaissance de l'écriture manuscrite. In *Traité informatique*, *Technique de l'Ingénieur,* (pp. 1–15). http://www.techniques-ingenieur.fr/base-documentaire/archives-th12/archives-documents-numeriques-gestion-de-contenu.

Crozet, P. (1999). A propos des figures dans les manuscrits arabes de géométrie: l'exemple de Sijzi. In Y. Ibish (Ed.), *Editing Islamic manuscripts on science* (pp. 131–163). London: al-Furqan.

Crozet, P. (2010). De I'usage des transformations géométriques à la notion d'invariant: la contribution d'al-sijzī. *Arabic Sciences and Philosophy, 20*(1), 53–91.

Descartes, R. (1969–75) [AT], *Oeuvres de Descartes* (New ed., 11 Vols.). In C. Adam & P. Tannery (Eds.). Paris: Vrin.

Detlefsen, M. (2005). Formalism. In S. Shapiro (Ed.), *The oxford handbook of philosophy of mathematics and logic* (pp. 236–317). Oxford: Oxford University Press.

De Vittori, T. (2009). *Les notions d'espace en géométrie, de l'Antiquité à l'Âge Classique*. Paris: L'Harmattan.

Domski, M. (2012). Newton and Proclus: Geometry, imagination, and knowing space. *The Southern Journal of Philosophy, 50*(3), 389–413.

Girard, J.-Y. (2007). La logique comme géométrie du cognitif. In J.-B. Joinet (Ed.), *Logique, dynamique et cognition* (pp. 15–29). Paris: Publications de la Sorbonne.
Grant, E. (1981). *Much ado about nothing: Theories of space and vacuum from the middle ages to the scientific revolution*. Cambridge: Cambridge University.
Gratus, J. & Porter, T. (2005). A geometry of information, I: Nerves, posets and differential forms. In R. Kopperman, M.B. Smyth, D. Spreen & J. Webster (Eds.), *Spatial representation: discrete vs. continuous computational models*, no. 04351 in Dagstuhl Seminar Proceedings, IBF Germany: Schloss Dagstuhl. http://drops.dagstuhl.de/opus/volltexte/2005/126/.
Grothendieck, A. (1985–86). Récoltes et Semailles. http://www.math.jussieu.fr/~leila/grothendieckcircle/RetS.pdf.
Harari, O. (2003). The concept of existence and the role of constructions in euclid's elements. *Archive for History of Exact Sciences, 57*, 1–23.
Harari, O. (2008). Proclus' account of explanatory demonstrations in mathematics and its context. *Archiv für Geschichte der Philosophie, 90*(2), 137–164.
Jones, A. (1986). *Pappus of Alexandria: Book 7 of the collection*. New York: Springer.
Koyré, A. (1957). *From the closed world to the infinite universe*. Baltimore: The John Hopkins Press.
Lernould, A. (Ed.). (2010). *Etudes sur le Commentaire de Proclus au premier livre des Eléments d'Euclide*. Lille: Presses du Septentrion.
Mac Lane, S., & Moerdijk, I. (1992). *Sheaves in geometry and logic: A first introduction in topos theory*. New York: Springer.
Mancosu, P. (1996). *Philosophy of mathematics and mathematical practice in the seventeenth century*. New York: Oxford University Press.
Manders, K. (2008a). Diagram-based geometric practice. In P. Mancosu (Ed.), *The philosophy of mathematical practice* (pp. 65–79). Oxford: Clarendon Press.
Manders, K. (2008b). The Euclidean diagram. In P. Mancosu (Ed.), *The philosophy of mathematical practice* (pp. 80–133). Oxford: Oxford University Press.
Maniglier, P. (2006). *La vie énigmatique des signes. Saussure et la naissance du structuralisme*. Paris: Leo Scheer.
Miller, N. (2001). *A diagrammatic formal system for Euclidean geometry.* Ph.D. Dissertation, Cornell University.
Miller, N. (2007). *Euclid and his twentieth century rivals: Diagrams in the logic of Euclidean geometry*. Stanford: Centre for the Study of Language and Information Press.
Morrow, G. R. (1992). *A commentary on the first book of Euclid's element*. Princeton: Princeton University Press.
Mueller, I. (1981). *Philosophy of mathematics and deductive structure in Euclid's 'Elements'*. Cambridge Mass: The MIT Press.
Mueller, I. (1987). Mathematics and philosophy in Proclus' commentary on book I of Euclid's elements. In J. Pépin & H. D. Saffrey (Eds.), *Proclus lecteur et interprète des anciens* (pp. 305–318). Paris: Edition du CNRS.
Mumma, J. (2006). *Intuition formalized: Ancient and modern methods of proof in elementary Euclidean geometry.* (Ph.D. Dissertation, Carnegie Mellon University).
Mumma, J. (2012). Constructive geometrical reasoning and diagrams. *Synthese, 186*, 103–119.
Netz, R. (1999). *The shaping of deduction in Greek mathematics: A study in cognitive history*. Cambridge: Cambridge University Press.
Netz, R. (2009). Imagination and layered ontology in Greek mathematics. *Configurations, 17*(1), 19–50.
Nikulin, D. (2002). *Imagination and geometry: Ontology, natural philosophy and mathematics in Plotinus, Proclus and Descartes*. Aldershot: Ashgate.
Nikulin, D. (2008). Imagination and mathematics in Proclus. *Ancient Philosophy, 28*(1), 153–172.
O'Meara, D. J. (1989). *Pythagoras revived.* Oxford: Clarendon Press.
Panza, M. (2012). The twofold role of diagrams in Euclid's plane geometry. *Synthese, 186*, 55–102.

Proclus (1873). *Procli Diadochi in primum Euclidis elementorum librum commentarii ex recognitione Godofredi Friedlein.* Leipzig: Teubner (reprint Olms, 1992).

Rabouin, D. (2009). *Mathesis universalis. L'idée de 'mathématique universelle' d'Aristote à Descartes*. Paris: P.U.F.

Rashed, R. (2001). *Les Mathématiques infinitésimales du IXe au XIe siècle* (Vol. IV). London: Al-Furqân Islamic Heritage.

Richard, P.R. (2000). Diagnostic sur la structure et la qualité de preuves inadmissibles, *La lettre de la preuve*. *International newsletter on the teaching and learning of mathematical proof*, November/December (published online).

Saito, K. (2006). A preliminary study in the critical assessment of diagrams in Greek mathematical works. *Sciamus, 7*, 81–144.

Saito, K. (2008). Diagrams in Greek mathematical texts. http://www.greekmath.org/index.html.

Saito, K., & Sidoli, N. (2012). Diagrams and arguments in ancient Greek mathematics: Lessons drawn from comparisons of the manuscript diagrams with those in modern critical editions. In K. Chemla (Ed.), *The history of mathematical proof in ancient traditions* (pp. 135–162). Cambridge: Cambridge University Press.

Vitrac, B. (2005). Quelques remarques sur l'usage du mouvement en géométrie dans la tradition euclidienne: de Platon et Aristote à 'Umar Khayyam. *Fahrang, 18*, 1–56.

Vitrac, B. (2012). *Les démonstrations par l'absurde dans les Éléments d'Euclide : Inventaire, formulation, usages.* http://hal.archives-ouvertes.fr/docs/00/68/55/20/PDF/Les_dA_monstrations_par_l_absurde.pdf.

van der Waerden, B. L. (2003). *Algebra*. New York: Springer.

Subject, Space, Object: The Birth of Modernity

Franco Farinelli

According to Hans Blumenberg, there are two inventions whereby the modern age cleanly breaks with the past: Florentine linear perspective and absolute space.[1] They are the same thing in many ways since one is unthinkable without the other and vice versa. But this is of little import at the moment. Of immediate concern, although it has heretofore escaped notice, is that the first sighting of the New World was an avowedly perspective view in the Florentine sense of the term. Indeed, as in ancient myths, the real significance of the discovery of America resides in what at first sight appears to be a useless, even misleading complication, if not a glaring incongruity or logical contradiction. An entry in what remains of Columbus's blog[2] notes the complicated event of that sighting under the date of 11 October. At 2 h past midnight, the crew of the *Pinta* sighted land, and Rodrigo de Triana claimed the promised reward. Yet he would never receive it because Columbus asserted he had already seen it about 10 pm the evening before when, standing on the poop deck, he had glimpsed in the same direction a flickering light like that of a wax candle. Much excited talk ensued. As calculated later, Columbus's position was about 56 miles, nearly 90 km, from land, a distance which seemingly would have made it hard to have seen anything of the kind. Even the versions of Las Casas and Fernando Columbus, which mention people walking along the coast carrying a sort of lantern, are less than convincing. In the event, the reward went to Columbus, and one tradition has it that poor Rodrigo, resentful and embittered, converted to Islam

F. Farinelli (✉)
Dipartimento di Filosofia e Comunicazione, Università di Bologna,
Via Azzo Gardino 23, 40122 Bologna, Italy
e-mail: franco.farinelli@unibo.it

[1] Blumenberg (2009, pp. 324–5).

[2] Varela and Gil (1992, pp. 23–4 e nn. 26–8).

V. De Risi (ed.), *Mathematizing Space*, Trends in the History of Science,
DOI 10.1007/978-3-319-12102-4_6

and emigrated to Africa. Yet what are we to make of this story? Indeed, is it possible to find a sense whose 'real content' coincides perfectly with its 'truth content', or, as Benjamin[3] might have put it, the content whose import is what comes to light with time because it is exemplary and acknowledged as the key to what is to follow?

Christopher means *Christo ferens* (as Columbus will unfailingly sign his name from 1502 on), i.e. not he who bears Christ but bears to or by virtue of Christ.[4] In order to solve the enigma of what the name meant to Columbus and, hence, to grasp the sense we are seeking, it is enough to recall that in every circular medieval world map, such as those that adorned cathedral walls or even fifteenth-century missals, the figures of the Earth and of Christ formed a single whole, and that, as the host becomes the real body of the Redeemer through the Eucharist, the map—the host of the Earth—turns into his real body. All modernity develops through, and ends in, this transformation, albeit progressively losing cognizance of it at the same time. For Columbus, however, the map consciously remained what it had been throughout the Middle Ages—a prophecy.[5] According to the principle that Baudrillard will call the "precession of simulacra"[6] Columbus, loyal executor of the injunction incorporated in the "Map of the Ocean" by Toscanelli of Florence which he carried with him,[7] could not but oppose Rodrigo's assertion and claim as his the real first sighting of the New World. He alone had to be the first to proclaim the fulfilment of the prediction, to personally validate the precedence of the prophetic image over reality, thereby sanctioning the derivative nature of the latter with respect to the former. In effect, had Rodrigo been right, any assimilation of knowledge to acknowledgement of the cartographical model (the fundamental but almost always implicit proposition of modern epistemology) would have been placed in doubt. As King Ferdinand and Queen Isabella will write to Columbus on 16 August 1494 in acknowledging his feat: "What you had announced to us has been achieved, as if you had seen it before speaking to us."[8] *Christo ferens* is thus the one who leads to Christ by virtue of Christ, that is, the map, and where the only word, the name of the Saviour, stands for both means and end.

This is precisely the meaning of the passage in which Heidegger asserts that the "Modern World is the triumph of the world resolved into image", that is, according to "the configuration of the representing production",[9] the most icastic of definitions of cartographical depiction. With this we know exactly how to understand what Jacob Burckhardt intended at the beginning of his Zurich lectures in referring to the Renaissance as a more or less direct retrieval of the "image of Ancient Rome".

---

[3] Benjamin (1962, pp. 157–8).

[4] Varela and Gil (1992, pp. L–LI).

[5] Farinelli (2009, pp. 20–6).

[6] Baudrillard (1981).

[7] Farinelli (2011a, p. 7).

[8] Todorov (1992, p. 27).

[9] Heidegger (1968, p. 100).

Before becoming something ideal, as it was for Burckhardt and even for his interpreters,[10] there is something material about this image, a tangible code as all codes which practices are built upon are in the end, i.e. processes that make it possible to act on what is real through what is symbolic.[11] The image of Rome is the material figure that Rome produces of itself at the Empire's apogee, the genitive case being understood in its double sense of possession and specification: it is the spatial map that Claudius Ptolemy, the last heir to Greek science,[12] showed us how to construct. Historians of art have long kept his memory alive: it is from the Quattrocento rediscovery in Florence of his *Geographia* that Renaissance linear perspective derives,[13] which (except for a decisive difference discussed shortly *infra*) is nothing but projection—the Ptolomeic procedure of "depiction of the world on a plane that has proportional measurements and a correspondence to what is depicted as round, or in spherical form", in the words of an authoritative fifteenth-century commentator.[14] The image in perspective is much more than a way to access a new kind of "truth" through representation.[15] Nor is it simply the figure before which the modern agent "discovers himself".[16] Just as it is not the only artefact marking the birth of modern science.[17] Understanding what Panofsky has called "artificial perspective"[18] does not involve breaking down its isolated elements and individual effects. A total institution, "total social phenomenon", in the precise sense that Marcel Mauss attributed to the expression, i.e. one capable of moving society in its entirety,[19] modern perspective produces a *habitus*, exactly according to the meaning Pierre Bourdieu sees in the term: that of a system of "durable arrangements", of a "structured structure intended to function as a structuring structure" of practices and representations at once objectively regulated but free of any conscience and control of the operations involved.[20] In short, an ethic within which physical stature and ideal conception, practice and theory (in the term's original visual meaning), are inseparable. It is on the basis of this connection that Florentine perspective apodictically states its constrictive and totalitarian lesson, one capable of colonising all modernity starting from the relationship between *civitas* and *urbs*, between its civic component and urban element.

No one understood all of this—the subversive import of the modern perspective gaze—better or more deeply than Pavel Florenskij. Nor is it just a matter of

[10] Ghelardi (1991, pp. 129, 134, 135).

[11] Schiavone (2005, pp. 8–13).

[12] Aujac (1993, p. 7).

[13] Edgerton (1975, pp. 93–), Veltman (1980, pp. 403–7), Edgerton (1991, pp. 151–).

[14] Ptolemy (1621, p. 8).

[15] Damisch (1987, p. 170).

[16] Belting (2010, p. 12).

[17] Gioseffi (1980, pp. 81–91).

[18] Panofsky (1961, pp. 35–114).

[19] Mauss (1991, p. 172).

[20] Bourdieu (2000, pp. 256 e 393, n. 39).

phenomenology or of his insight into the spatial mechanism built on a quantitative standard that is connected to it. He also apprehended its more general consequences. It would be hard to find a simpler, clearer explanation of the mechanism: the construction rests on the invariance of the bi-ratios between a straight line on which lay several points, another straight line corresponding to the plane of representation and all of the points intersecting with a bundle of straight lines that start from a single point outside the first two.[21] Thus having unveiled "the basis of the artist's perspective", the consequences he sees deriving from its use become clear in a flash: the onlooker, motionless, "paralysed as though poisoned by curare", is no longer a "living person" and the image becomes "a morbid optical illusion largely deprived of any humanity".[22] The absence of humanity within the mechanism of perspective corresponds in effect to the triumph of the quantitative and brings us back to Heidegger, to America and to Columbus. Heidegger writes that as soon as the quantitative takes on a quality of its own, it is transformed through the gigantic into the incalculable "invisible shadow cast over everything when man has become *subjectum* and the world image".[23] Yet when does all this occur? When is subjection (being-already-at-hand) of man to the quantitative. i.e. spatial, image of the world establish itself, and where does it do so?

The answer admits of no doubt: in Florence, under the Portico degli Innocenti which Filippo Brunelleschi, the founder of the architectural cycle of Humanism, put his hand to from 1419 to 1422.[24] This work cannot be understood without bearing in mind the prodigious efforts of the Florentines beginning in the 1300s to transform a county seat, whose territory was still semi-feudal, into a city-state of national import[25] by converting it into a single space having a plurality of places. It is as part of the spatial ordering connected to Florence's mercantile ideology of the early Quattrocento that the Ospedale degli Innocenti is for Brunelleschi a *speculum*. The idea is that all the building's horizontals and volumes are meant to speak a "clean and clear" idiom with the "rigour of the abacus, with the categorical imperative of the double-entry": a building within which an unheard-of façade arcade achieves the intercision of the first real visual pyramid of modern perspective with the entire city plan.[26] It is the building itself that seems to signal immediately and unequivocally Brunelleschi's manifest awareness of the ontological nature of the procedure perspective entails. That is, it his conscious sense of the fact that perspective (projection) transmutes not only linearity but essence, not incidence but substance, not just form but the nature of everything that falls within its ambit, of everything that it captures. And it begins with the city and its citizens.

---

21 Florenskij (1995, pp. 244–48).

22 Florenskij (1983, pp. 83, 124–).

23 Heidegger (1968, p. 100).

24 Morolli (2010, pp. 80–5).

25 Saalman (1980, pp. 473–4).

26 Morolli (1979, p. 174).

Walking up the steps separating the building from the piazza, one enters the Arcade precisely at the point that Brunelleschi intended: under the false side door opposite the small square window that was, up to 1875, what would today be called a night safe in which to put new-born babies "whose fathers and mothers had fled from the duties of Nature", as a nineteenth-century biographer of Brunelleschi has put it.[27] To place oneself within the Arcade already means to submit, to become subjected to, to acknowledge oneself as *subjectum*. It is like Odysseus clutching the sheep's underbelly in the cave of Polyphemus since here too one is surmounted by a material structure, literally underneath it and, hence, on which one physically depends. And the rigour mortis that the subject must immediately take on so that the trick of perspective works, and that Florenskij marvelled at so much, is itself analogous to the "mimesis of the dead"[28] that the Greek hero and his companions must obey to save themselves from the giant. Then too it is comprehensible, for in both cases, in the Cyclops's cavern and under the ribbed vaults of the Ospedale's unusual façade, the same thing occurs—the birth of space[29]—the subsuming of the face of the Earth to a single standard of linear measurement. Once in the prescribed position, you need only look ahead to be immediately summoned to decide a question that would have been inconceivable in the Classical or Middle Age: whether to believe what you touch or what you see. Jesus says in the Gospel to the Apostle Thomas who doubts the Resurrection of the Redeemer: "Reach hither thy finger, and behold my hands; and reach hither thy hand, and thrust it into my side". As Glenn Most acknowledges: since the imperative of seeing takes primacy here over the placing of the hand, Jesus is not simply telling Thomas to extend his finger and then only look at his hands without touching them. Rather, he is telling him that by thrusting his finger into the hole left by the nail he will be able to see the hands of Jesus for what they are, the hands of a man who was killed by mortal wounds. The same holds true for his side.[30] So Thomas sees with his hands in the sense that for Thomas (as for all of Euclidian geometry) the tactile, muscle intuition decidedly prevails over the visual, nor is there ever a contradiction with it. Exactly the opposite is what begins happening under the Arcade of the Innocenti because one is unable to decide whether the lines delimiting its floor are parallel or not. In this regard, our touch says yes but for the first time our sight says the opposite.

If we run our fingers along the sides of a straight ruler, we have a clear, incontrovertible sensation that the parallel lines will never meet. But if we appraise parallelism by sight, "by looking along a column, say," the sensation is exactly the opposite: the parallel lines end up converging so long as they are long enough.[31] Just substitute for the column Ivins mentions the first example of modern architecture in perspective, Brunelleschi's Arcade, to understand the revolutionary

[27] von Fabriczy (1979, p. 270).

[28] Horkheimer and Adorno (1966, p. 71).

[29] Farinelli (2003, pp. 12–3, 120–23).

[30] Most (2009, p. 48).

[31] Ivins (1985, pp. 31–7).

import of cognitive knowledge Florentine linear perspective inaugurated. It's as if the supposed convergence point of parallel lines, the so-called vanishing point, swallowed up and caused to disappear the entire sensory experience of Antiquity, only to have it reborn in a form that is essentially visual and already incorporates every principle of dematerialisation and virtuality bruited about today. It's a formidable re-nascence, one that Masaccio, in his coeval *Holy Trinity* of the Brancacci Chapel, the first painting rendered in perspective, associated with the supreme, most defining of re-nascences, that of Our Lord, whose nearly intact and immaculate side wound stands for the opening of the modern age. It marks in other words the end of the identity of the tactile experience and beginning of the visual while proclaiming at the same time the impossibility of the latter to account for the now secret clockwork of the world reduced to image. Masaccio's *Trinity* is above all a long catalogue of gazes ranging from the altogether human and trusting ones of the donors kneeling in the foreground to that of St John the Evangelist contemplating the mystery of the Cross and that of the Virgin Mary, oblique, suspended on the threshold between the worldly and the otherworldly. Of course, what most impresses is the gaze of the Godhead, the figure dominating all the others because it has nothing natural to it, being all but devoid of expression. This is precisely the gaze of perspective: there is almost nothing human about it any more, fixed as it is on something at the edge of the visible, "almost to infinity", as the vanishing point, the invisible spot that holds the entire artifice of perspective together, will seem to Alberti. The gaze of the *Trinity*'s God, which stops at the very limit of what Florenskij will call "humanity", is exactly that of the Columbus who claimed to see land. Both are fixed on a tiny point that appears and disappears, that continually oscillates between the statute of existence and subsistence, or the *vanishing point* as it is still technically called today in the English-speaking world. "Where it seems to me I fix a point": this is how Alberti[32] describes the practice by which perspectivists direct their visual pyramid wherever they please, the point in question being its vertex. And it is exactly this very same Florentine gaze that the Genoese Christopher Columbus, perspectivist in his turn, will impose that night on the New World so that it becomes the gaze of the entire modern West.

Walter Ong has written that printing is a technique for making sound permanent by transforming it into silence.[33] Similarly, it can be said that perspective is a technique for making the subject stable by transforming it into object. The vanishing point represents the observer in the image because it reveals to the latter within itself a symbolic reference capable of immobilising one's gaze.[34] The result is that the self-same observer is reduced to a point, a mathematical entity,[35] and it is on the basis of this abstraction that the equivalence between the terminals of the cognitive procedure, between subject and object, is accomplished and rendered

[32] Alberti (1950, pp. 31–2).

[33] Ong (1956, p. 228).

[34] Belting (2010, p. 21).

[35] Aiken (1998, p. 95).

irreversible by making the former equal to the latter. The entire device rests upon the ability to calculate the interval separating them, and this ability is the archetype and incubator (the chance to say so did not come about by chance) of every subsequent, modern form of reification, according to the meaning Lukács assigns the word. In this connection even Florenskij's phenomenology of perspective evinces its limits since it fails to grasp the essential premise of the trick: before being immobile, the subject must be alone, reduced to a singular essence. Beneath the Arcade's false door in front of the window there's room for one person only, perspective's illusion being for one individual at a time. Indeed, in front of it, the city's inhabitants are in practice forced to disband as civic collective and acknowledge themselves to be a myriad of single, solitary, isolated citizens equal to one another by virtue of their geometrical objectification and atomic abstraction. As Lukács puts it "…all forms of objectuality and all the forms of subjectivity in bourgeois society corresponding to them" are determined by the "…essence of the structure of commodities".[36] And the "secret" of this form, in Marx's words,[37] by which is meant that relations between people reveal themselves as things and relations between things, originates under that Arcade. In effect, as a function of the spatial model's success, the Arcade inexorably leads to, even compels, that initial separation between the social body and the individual body. Prey to Brunelleschi's mechanism, the subject takes on, per force of the geometrical model's dictate, the nature of that object from which for the first time the subject is clearly separated and irremovably dissociated. It is from this perspectival subject that descends what four centuries later will appear to Frederick W. Taylor, the inventor of scientific management, as a "trained gorilla" endowed with minimum mobility and under complete control: the assembly-line worker. He was forced, as Gramsci noted to "reduce production operations to a mere physical, machine-like motions", thereby cutting "the psycho-physical bond of qualified professional work", which instead required a worker's intelligence, imagination and active participation. Whence "puritan" movements like prohibition to ban alcohol, which became a veritable "function of State" as government was charged with the task of contrasting the physiological collapse of a work force pushed to its limits by the new production method.[38] Under the Arcade of the Innocenti, incubator of Taylorism and Fordism, is the very vision that becomes a government task in that, without the constraints implicit in the perspectival gaze, the modern State would never have been able to exist, to the point of standing Gramsci's formulation on its head and considering the former as a function of the latter.

For the trick of perspective to work best a monocular vision is advisable. It is almost as if the constraint imposing the subject's individuality and, hence, the no longer collective but singular (albeit serial because artificially conditioned) nature of his or her gaze had some sort of precise physiological corroboration in the use of

[36] Lukács (1974, pp. 19–20).

[37] Marx (1967, pp. 104–).

[38] Gramsci (1975, p. 2165).

one eye instead of both. It is in this link between physiology and subjectivity that the Florentine perspective device reveals the political character of its task. What happens to the gaze under the Arcade is exactly what happened to the newborn placed in the compartment of the wooden wheel behind the window: both being swallowed up by the vanishing point, they change their nature and mode of being. The vertex of the first modern perspectival artifice coincides with an aperture that physically took a newborn from one world and put it in another, marking its material passage from one condition to the other, from anonymity of natural descent to the identity of social existence—an existence that in passing through the short dark turn of the wheel now acquired both a name it had not been given as a simple creature in a purely biological sense before that passage and the visibility required by full membership in the human consortium, the *civitas*, wherein to be called Innocenti meant being acknowledged not as simple citizens but as outright "children of Florence". It was a veritable rebirth, a real re-nascence of the infant, through which the nature of the world that the technique of perspective constructs really was a compensation of nature itself, the compensation of a recovering and correcting (emendating and rectifying) by the *polis* of the immediate result of the physical and biological processes that had put it into the world. The same holds for the gaze. In comparing the fixed gaze to the mobile, circular glance, Edward Casey[39] notes that the former tends to become disembodied in the object of its stare, to repress the corporeal basis of its seeing, thus sacrificing its original nature. This is exactly what Leon Battista Alberti precociously realised apropos of the perspective gaze, and it was that very discovery that led him to change his emblem to a winged eye completely separated from the rest of the body.[40] In the words of Florenskij, which are evidently similar to those of Gramsci *supra*: all the "psycho-physiological processes" of the act of seeing are in this way eliminated in the sense that the eye, rendered immobile and impassive "like an optical lens" does not have "the right to move despite [the fact] that the essential condition of vision is activity", so that the latter is degraded to a mere "outwardly mechanical process" that is not "accompanied by memory, or by spiritual effort, or by analysis".[41] At this point it is indeed easy to see in the perspective eye's kind of performance the model of studied efficiency that in the early Twentieth Century will lead under the criteria of Taylorism and Fordism to the appearance of a "new man",[42] one who is functional to the mass production and consumption of goods these criteria call for. It is as if this subject, in his entire constitution and aptitude, was the finished product of the progressive colonisation of his physical and psychic apparatus that had begun five centuries earlier starting with sight.

The explanation resides in the homology between the principles of Florentine perspective vision and the statutes underpinning the construction of modern

[39] Casey (2007, p. 154).

[40] Smith (1994, pp. 453–55).

[41] Florenskij (1983, p. 125).

[42] Harvey (1997, p. 158).

territorial nations, a construction that begins in the ideal and material sense with the building of the Ospedale degli Innocenti. To produce the illusion, the gaze induced by Brunelleschi and the other "perspectivists" must be continuous, uniform and isotropic, according to the criteria that defined for Euclid the geometric nature of extension. That is, the gaze cannot linger or stop here and there during the flash-like path towards the vanishing point. It always has to be the same at any given moment, i.e. fixed along the rectilinearity and in velocity, rules that along with orthogonality are the fundamental principles of the logic of space, and to be orientated in the same direction. Similarly, all modern nation-building, with rare exceptions, complies with these precepts: it must lay claim to a continuous area,[43] hence encompassing a territory all of a piece, of a single plot of the Earth's surface; it must be organised according to equality, uniformity or identity of its constituent elements, which here means the cultural values and ideals of the subjects, and then citizens, of the nation; and all of its parts must be functionally orientated in the same direction, which the capital stands for since it tends to be located at the centre of the territory.[44] Given the homologous relation between disembodied gaze and body of State, the perspective *habitus* acquires the nature of a veritable "organizational closure", in the words of the autopoietic theory, that is, a circular concatenation of processes that, taken as a whole, constitute a self-calculating complex of subject and object capable of acquiring coherence through its own operating in space but not through the intervention of outside stimuli.[45] These are the stimuli that, in origin, the semi-closed structure of the Arcade is intended to eliminate.

Florentine linear perspective in this sense has the task of acting as a mode of regulation, by which is meant the interiorised rules upon which are built the social processes functional to capitalist accumulation,[46] that is, the set of mediations that maintain the distortions produced by such accumulation within limits compatible with social cohesion.[47] Indeed, all the evidence converges to indicate in Brunelleschi's invention the archetypal form for modernity of this mode of production, the form on which all subsequent ones depend. Take, for example, the postulate, which the neoclassical theory of liberal inspiration holds so dear, of the homogeneity of the economic system, the "dream country of [general] equilibrium" longed for by Walras. It is posited on at least two conditions: (i) that the features of the system are in the heads of all individuals, who act as a single "representative individual" according to the hypothesis of rational expectations, and (ii) that the coordinating of individual actions is guided by an explicit or implicit planner.[48]

[43] "Of course, the idea of continuity …is one of the oldest in philosophy. Yet only the modern age has made the philosophical term 'continuity, continuum' a cliché and known to everyone, only in the modern age has the spring from which this idea gradually gushed seeped into the minds of entire generations" (Florenskij 2007, p. 16).

[44] Farinelli (2009, pp. 3–105).

[45] Varela (1979, pp. 55–6).

[46] Lipietz (1986, pp. 18–19).

[47] Aglietta (2001, pp. 12–25).

[48] Aglietta (2001, p. 10).

These assumptions are evidently inconceivable without hypostatising the perspectival subject that was developed (as it is still the case to say so) at Florence during the Quattrocento and was the forerunner of *Homo oeconomicus*; it is still the model of a great deal of Western thought and still reflects its Florentine progenitor, explicating, making explicit the latter's mental attitudes in its nascent state. The solitude of perspective's subject, which signals the authentic advent of modernity, represents in the thought of Marcel Gauchet the form of the end of religious, or holistic, society, the completion of the shift from ontological unity wherein God and the world are the same indissoluble thing to the expression of transcendence, i.e. of the definitive duality between the former and the latter.[49] Yet it is a form that already contains a decisive variant capable of endangering the entire construction, of undermining the foundations of transcendence itself. Its model, like that of Ptolemaic projection, develops vertically from top to bottom. By contrast, the perspective model runs along the horizontal—the world and God (the infinite terminus behind the vanishing point) are coplanar, as the subject and object of vision that are opposite one another under the Arcade and that constitute the origin of modern subjectivity and objectivity. This marks the end of the Middle Ages' Augustinian "pilgrim city", which arose from the regime of equivalence between things and living beings but from the point of view of the latter, from their copenetration, from the internal dialectic between *civitas* and *urbs*, between subjects and objects. In its place the contours of the modern spatial city begin to take shape, a city founded on separation, on the distance between the former and the latter, a distance that is ontologically irreducible as it is linearly measurable with maximum precision. Even the last urban smile thus fades away towards the late Cinquecento in the sense that buildings cease to have a real face like living beings in their representations and take on the impenetrability of the façade devoid of any expression, ceasing in short to have any affinity at all with humans, as for example the cartographical frescoes in the Vatican begin to attest to.[50] If the Commonwealth, Leviathan, is the "mortal God" for Hobbes in the mid-1600s, then Ptolomeic space, the logic of both the Commonwealth and the relationship between subject and object, is the lay God. Yet the Commonwealth for Hobbes is exactly the same as *Civitas*, understood as a "multitude united in one Person".[51] In other words: space is the agent of the transformation of the city in something truly monstrous because that something is Leviathan, the monster that, "spread upon gold and in mire", fears no other power on the face of the Earth, as is written in the Book of Job. It is argued that in the modern era the principle of territorial sovereignty has imposed a spatial solution to the primary ontological problem concerning the relationship between universality and particularity.[52] In actual fact, the opposite is true: territorial sovereignty is the form that space imposes on power and that the latter, in order to be,

[49] Gauchet (1985, p. 57).

[50] Farinelli (2011b, pp. 73–8).

[51] Hobbes (1981, p. 227).

[52] Walker (1993, p. 64), Di Martino (2010, p. 20).

and to be acknowledged as such, was compelled to take on, beginning from the city. This is how power begins to reproduce itself in a geometrical, indeed a geographical, way since its sphere of application is the entire face of the Earth. This is the birth of modernity, with nothing innocent attaching thereto. Under the Arcade of the Innocenti the *polis*, the very expression of power, consciously and programmatically reproduces itself for the first time according to a geometrical model, that is, through the control brought to bear by the form of the *urbs*, the material city, over the nature of the *civitas*, the city as symbolically shared heritage that is individually experienced.

References

Aglietta, M. (2001). Regolazione e crisi del capitalismo. In M. Aglietta & G. Lunghini (Eds.), *Sul capitalismo contemporaneo*. Turin: Bollati Boringhieri.

Aiken, J. A. (1998). The perspective construction of Masaccio's Trinity Fresco and medieval astronomical graphics. In R. Goffen (Ed.), *Masaccio's 'Trinity'*. Cambridge: Cambridge University Press.

Alberti, L. B. (1950). *Della pittura*. In L. Mallè (Ed.). Florence: Sansoni.

Aujac, G. (1993). *Claude Ptolémée*. Paris: CTHS.

Baudrillard, J. (1981). *Simulacres et simulation*. Paris: Galilée.

Belting, H. (2010). *I canoni dello sguardo. Storia della cultura visiva tra Oriente e Occidente* (M. Gregorio, Trans.). Turin: Bollati Boringhieri (Original edition: *Florenz und Bagdad. Eine westöstliche Geschichte des Blicks*. Munich: C. H. Beck, 2008).

Benjamin, W. (1962). Le affinità elettive. In R. Solmi (Ed.), *Angelus Novus. Saggi e frammenti*. Turin: Einaudi (Original edition: Goethes Wahlverwandtschaften. R. Tiedemann & H. Schweppenhäuser, (Eds.), *Walter Benjamin. Gesammelte Schriften*. Bd. I, 1, Frankfurt: Suhrkamp 1974, pp. 125–201).

Blumenberg, H. (2009). *Uscite dalla caverna* (M. Doni, Trans.). Milan: Medusa. (Original edition: *Höhlenausgänge*, Frankfurt am Main: Suhrkamp, 1989).

Bourdieu, P. (2000). *Esquisse d'une théorie de la pratique*. Paris: Seuil.

Casey, E. S. (2007). *The world at a glance*. Bloomington and Indianapolis: Indiana University Press.

Damisch, H. (1987). *L'origine de la perspective*. Paris: Flammarion.

Di Martino, A. (2010). *Il Territorio: Dallo Stato-Nazione alla Globalizzazione. Sfide e prospettive dello Stato costituzionale aperto*. Milan: Giuffré.

Edgerton, S. Y. (1975). *The renaissance rediscovery of linear perspective*. New York: Basic Books.

Edgerton, S. Y. (1991). *The heritage of Giotto's geometry. Art and science on the eve of the scientific revolution*. Ithaca and London: Cornell University Press.

Farinelli, F. (2003). *Geografia. Un'introduzione ai modelli del mondo*. Turin: Einaudi.

Farinelli, F. (2009). *La crisi della ragione cartografica*. Turin: Einaudi.

Farinelli, F. (2011a). Why America is called America. In S. Daniels, D. Delyser, J. N. Entrikin, & D. Richardson (Eds.), *Envisioning landscapes, making worlds. Geography and the humanities*. London, New York: Routledge.

Farinelli, F. (2011b). Bononiensis Ditio. Un racconto di tre città. In F. Ceccarelli & N. Aksamija (Eds.), *La Sala Bologna nei Palazzi Vaticani. Architettura, cartografia e potere nell'età di Gregorio XIII.* Venice: Marsilio.

Florenskij, P. (1983). *Beyond vision: Essays on the perception of art*. In N. Misler & W. Salmond (2002, Trans.). London: Reaction Book (La prospettiva rovesciata. In N. Misler (Ed.), Rome: Casa del Libro Editrice. Original edition: *Обратная перспектива* [*Obratnaja perspektiva*], 1920).

Florenskij, P. (1995). *Lo spazio e il tempo nell'arte*. In N. Misler (Ed.), Milan: Adelphi. (Original edition: *Анализ пространственности и времени в художественно-изобразительных произведениях* [*Analiz prostranstvennosti i vremeni v chudožestvenno-izobrazitel'nych proizvedenijach*] Moscow: Izdatel'skaja gruppa Progress, 1993).

Florenskij, P. (2007). *Il simbolo e la forma. Scritti di filosofia della scienza*. In N. Valentino & A. Gorelov (Ed.). Turin: Bollati Boringhieri.

Gauchet, M. (1985). *Le désenchantement du monde. Une histoire politique de la religion*. Paris: Gallimard.

Ghelardi, M. (1991). *La scoperta del Rinascimento. L' "Età di Raffaello" di Jacob Burckhardt*. Turin: Einaudi.

Gioseffi, D. (1980). Filippo Brunelleschi e la "svolta copernicana": la formalizzazione "geometrica" della prospettiva. Gli inizi della scienza moderna. In G. De Angelis d'Ossat (Ed.), *Filippo Brunelleschi. La sua opera e il suo tempo* (Vol. I). Florence: Centro Di.

Gramsci, A. (1975). *Quaderni dal carcere* (Vol. III). Turin: Einaudi.

Harvey, D. (1997). *La crisi della modernità* (M. Viezzi, Trans.). Milan: il Saggiatore. (Original edition: *The condition of postmodernity: An enquiry into the origins of cultural change*, Oxford: Blackwell, 1990).

Heidegger, M. (1968). L'epoca dell'immagine del mondo. In P. Chiodi (Ed.), *Sentieri interrotti*. Florence: La Nuova Italia. (Original edition: F.-W. von Herrmann (Ed.). (1950). Frankfurt: Klostermann. Die Zeit des Weltbildes [1938]. *Holzwege*. Heidegger, M. *Gesamtausgabe*, vol. 5).

Hobbes, T. (1981). *Leviathan* [1651]. In C. B. Macpherson (Ed.). Harmondsworth: Penguin Books.

Horkheimer, M., & Adorno, T. W. (1966). *Dialettica dell'illuminismo (*R. Solmi Trans.). Turin: Einaudi. (Original edition: *Dialektik der Aufklärung* [1947]. In Adorno, T. W. *Gesammelte Schriften* (Vol. 3). R. Tiedemann (Ed.) Frankfurt: Suhrkamp, 1981).

Ivins, W. (1985). La rationalisation du regard. *Culture Technique, 14*, 31–377.

Lipietz, A. (1986). New tendencies in the international division of labor: Regimes of accumulation and modes of regulation. In A. J. Scott & M. Storper (Eds.), *Production, work, territory. The geographical anatomy of industrial capitalism.* Boston: Allen & Unwin.

Lukács, G. (1974). *Storia e coscienza di classe* (G. Piana, Trans.). Milan: Sugar. (Original edition: *Geschichte und Klassenbewusstsein. Studien über marxistische Dialektik*, Berlin: Malik, 1923).

Marx, K. (1967). *Il Capitale* (Vol. I, D. Cantimori, Trans.). Rome: Editori Riuniti. (Original edition: *Das Kapital*, Hamburg: Meissner, 1867; 1890 (Vols. 23–25, 4th ed.) in *Marx-Engels-Werke*. Berlin/DDR: Dietz Verlag, 1962).

Mauss, M. (1991). *Saggio sul dono*. M. Aime (Ed.). Turin: Einaudi. (Original edition: *Essai sur le don. Forme et raison de l'échange dans les sociétés archaïques* [1925, 4th ed.], Paris: Les Presses universitaires de France, 1968).

Morolli, G. (1979). I Cantieri. In F. Borsi, G. Morolli & F. Quinterio (Eds.), *Brunelleschiani*. Rome: Officina Edizioni.

Morolli, G. (2010). Gli Innocenti, Brunelleschi e gli altri. In S. Filipponi, E. Mazzocchi, & L. Sebregondi (Eds.), *Il Mercante, l'Ospedale, i Fanciulli. La donazione di Francesco Datini, Santa Maria Nuova e la fondazione degli Innocenti*. Florence: Nardini.

Most, G. W. (2009). *Il dito nella piaga. Le storie di Tommaso l'incredulo* (D. La Rosa,Trans.). Turin: Einaudi. (Original edition: *Doubting Thomas*, Cambridge, MA-London: Harvard University Press, 2005).

Ong, W. (1956). System, space and intellect in renaissance symbolism. *Bibliothèque d'Humanisme et Renaissance, 18*, 222–239.

Panofsky, E. (1961). La prospettiva come 'forma simbolica'. In G. D. Neri (Ed.) Id., *La prospettiva come "forma simbolica" e altri scritti*, pp. 35–114. Milan: Feltrinelli. (Original edition: Die Perspektive als 'symbolische Form'. In K. Michels & M. Warnke Id., *Deutschsprachige Aufsätze* (Vol. 2). Berlin: Akademie-Verlag. 1998 [1st. edition in *Vorträge der Bibliothek Warburg* 1924–1925. F. Saxl (Ed.) Leipzig-Berlin, 1927]; English edition: *Perspective as symbolic form* (C. S. Wood, Trans.). New York: Zone Books, 1991).

Ptolemy [Claudius Ptolemaeus] (1621). *Geografia cioè Descrittione universale della terra partita in due volumi, nel primo de' quali si contengono gli otto libri di Cl. Tolomeo, nuovamente con singolare studio rincontrati, et corretti dall'eccell.mo sig. Gio. Antonio Magini padovano* (L. Cernoti, Trans.). *The Geography.* Padua: Paolo et Francesco Galignani.

Saalman, H. (1980). Vecchie e nuove prospettive su Brunelleschi. In G. De Angelis d'Ossat (Ed.), *Filippo Brunelleschi. La sua opera e il suo tempo* (Vol. 2). Florence: Centro Di.

Schiavone, A. (2005). *Ius. L'invenzione del diritto in Occidente*. Turin: Einaudi.

Smith, C. (1994). L'occhio alato: Leon Battista Alberti e la rappresentazione del passato. In H. Millon & V. Magnago Lampugnani (Eds.), *Rinascimento. Da Brunelleschi a Michelangelo: la rappresentazione dell'architettura*. Milan: Bompiani.

Todorov, T. (1992). *La conquista dell'America. Il problema dell' "altro"* (A. Serafini, Trans.). Turin: Einaudi. (Original edition: *La conquête de l'Amérique*: *la question de l'autre*, Paris: Éd. du Seuil, 1982).

Varela, F. (1979). *Principles of biological autonomy*. New York-Oxford: North Holland.

Varela, C., & Gil, J. (Eds.). (1992). *Cristoforo Colombo. Gli scritti* (P. L. Crovetto, M. Cipolloni & D. Carpani, Trans.). Turin: Einaudi. (Original edition: *Cristóbal Colón. Textos y documentos completos*, Madrid: Alianza Editorial, 1992).

Veltman, K. H. (1980). Ptolemy and the origins of linear perspective. In M. Dalai Emiliani (Ed.), *La prospettiva rinascimentale. Codificazioni e trasgressioni*. Florence: Centro Di.

von Fabriczy, C. (1979). *Filippo Brunelleschi. La vita e le opere* (Vol. II). Florence: Uniedit.

Walker, R. B. J. (1993). *Inside/outside. International relations as political theory*. Cambridge: Cambridge University Press.

On Natural Geometry and Seeing Distance Directly in Descartes

Gary Hatfield

As the word "optics" was understood from antiquity into and beyond the early modern period, it did not mean simply the physics and geometry of light, but meant the "theory of vision" and included what we should now call physiological and psychological aspects. From antiquity, these aspects were subject to geometrical analysis. Accordingly, the geometry of visual experience has long been an object of investigation.

From antiquity to the seventeenth century, the understanding of the geometry of visual stimulation underwent significant changes. Across these changes, some aspects were conserved (e.g., the notion of visual angle) and some were radically altered (e.g., the character and location of optical stimulation). Focusing on the geometry of distance perception, I survey the development of theories relating visual stimulation to visual experience, culminating in a close examination of Descartes' theory of spatial perception and his doctrine of "natural geometry."

Kepler's discovery of the retinal image around 1600 afforded and required new thoughts about spatial vision. Kepler's discovery, when paired with Renaissance practices of making perspective images, revealed a new and basic fact about how vision works: it starts from an optical image with properties like those of a perspective drawing. There is, in effect, a perspective image on the retina, and this fact can be demonstrated by a suitable anatomical preparation, such as that depicted here (Fig. 1).

Kepler's discovery generated a deeply held theory that persists today: that what we "really see" is a two-dimensional image like that on the retina, with depth and distance added through judgment, interpretation, or ingrained habit. Many theorists held such a position. But Jonathan Crary's influential 1990 book universalizes this

G. Hatfield (✉)
Department of Philosophy, University of Pennsylvania,
Cohen Hall, Rm. 433, Philadelphia, PA 19104-6304, USA
e-mail: hatfield@phil.upenn.edu

V. De Risi (ed.), *Mathematizing Space*, Trends in the History of Science,
DOI 10.1007/978-3-319-12102-4_7

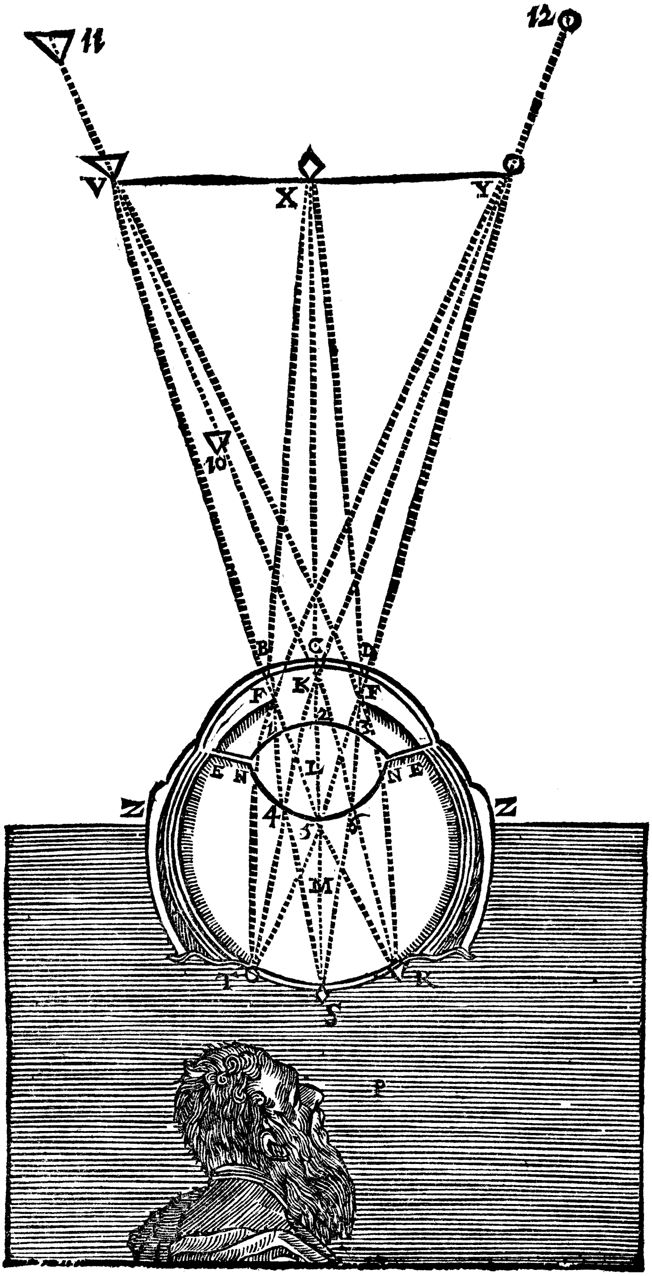


Fig. 1 A bovine eye in an anatomical preparation that allows the retinal image to be observed by an onlooker. From Descartes' *Dioptrics* (originally published in 1637). Reproduced from Descartes (1692a), p. 58 (author's collection)

position, in contending that the *camera obscura*, a device for projecting scenes into two-dimensional perspective images, served as a model of vision itself throughout the early modern period, to around 1810. On this model, which Crary (1990, pp. 43–50) attributes explicitly to Descartes, we "really see" only two dimensions.

Focusing on Descartes' *Dioptrics* and *Treatise on Man*, I challenge the adequacy of this understanding of the history of visual theory. It is true that most early modern visual theorists posited a two-dimensional sensory core as the immediate product of retinal stimulation as transmitted into the brain (Hatfield and Epstein 1979; Hatfield 1990, Chap. 2). But some accounts also described an immediate (primitive) experience of distance and depth. Descartes in particular held that the binocular eyes are embodied physiologically and psychologically (Fig. 2) and that

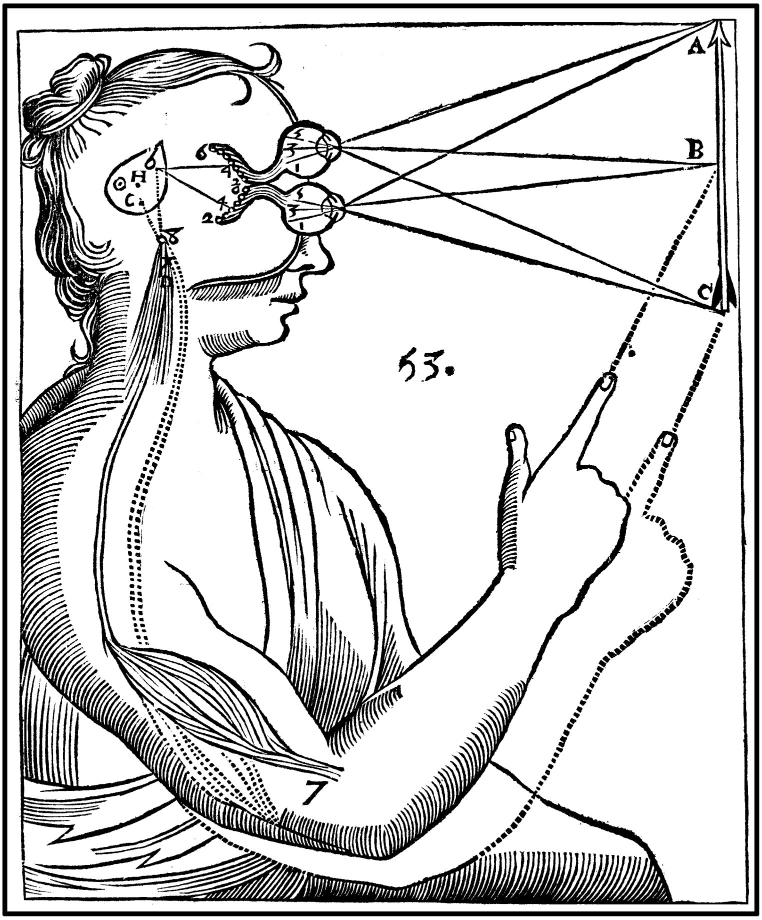


Fig. 2 Binocular eyes interacting with a pointing arm, illustrating embodied vision in Descartes, from the *Treatise on Man* (originally published in 1664). Reproduced from Descartes (1677), p. 74 (author's collection)

bodily processes produce a phenomenally immediate experience of depth and distance. In Descartes' theory, our visual experience is not "really" of a two-dimensional image but includes an experience of depth and distance that is directly produced by bodily mechanisms. The *camera obscura* is a model of the eye but not of vision. The eye does not see: it enables the embodied mind to see. Using two eyes, the embodied mind directly experiences objects at a distance. The notion that visual distance is phenomenally immediate was not peculiar to Descartes. Accordingly, accepted narratives about distance not being seen directly, stemming from a famous line in Berkeley's *New Theory of Vision*, will need to be rethought.

My account of the phenomenal and physiological immediacy of distance in Descartes depends upon a particular reading of how he theorized the role of the ocular muscular "cues" to distance. Separate muscle systems control the accommodation of the lens to provide clear vision at various distances and the convergence of the eyes to fixate a common focal point at a distance. I contend that Descartes established a psychophysiological account of distance perception, according to which brain states associated with accommodation and convergence directly produce an experience of visual distance in the mind through an "institution of nature." This interpretation can allow that Descartes acknowledged other processes operative in producing the experience of things at a distance in visual experience, including cognitively based operations that rely on (perhaps unnoticed) judgments. My aim is not to establish the psychophysiological account as the only account in Descartes, but to develop this aspect of his position in greater detail than heretofore.[1] In so doing, I intend to draw attention to a thread in early modern visual theory that extends from Descartes to nineteenth-century nativism.

The interpretation of Descartes' natural geometry that I advance here has implications for our understanding of his sensory epistemology and the role of reason or the intellect in the perception of primary qualities such as size, shape, and

[1] I have discussed Descartes' "psychophysical" (or psychophysiological) mechanisms for yielding sensations of distance and position (or direction) versus his "psychological" (or "judgmental" or "cognitive") accounts of distance, size, and shape perception in Hatfield and Epstein (1979, pp. 375, 378); Hatfield (1986, pp. 56–60); Hatfield (1990, pp. 38–39); and Hatfield (1992, pp. 356–57). These discussions show that, especially in the Sixth Replies in the *Meditations*, Descartes offered a traditional "unnoticed judgment" account of the sort stemming from Ibn al-Haytham's optics but that, in the *Dioptrics* and (in greater detail) the *Treatise*, he also theorized that the perception of position and distance arises directly from brain mechanisms. In Hatfield and Epstein (1979, p. 378), we speculate that he might have imagined mechanisms in which distance and visual angle are brought together to yield size and shape perception; below, I indicate how this might be done. Vinci (1998, Chap. 3), reconstructs material from the *Dioptrics* and Sixth Replies to show how visual angle (retinal size) and distance information might be combined physiologically to yield size constancy (my speculation differs). Wolf-Devine (1993) discusses my psychophysical versus psychological distinction using the terms "mechanical" and "homuncular," suggesting that, in speaking of unnoticed judgments, Descartes engaged in homuncular reasoning, as if a little person is looking at a copy of the retinal image and employing "reasoning and judgment" (p. 67) to perceive situation, distance, size, and shape (1993, Chap. 4). Although Descartes sometimes uses homuncular language, I resist assimilating his psychological (unnoticed judgment) account to a homunculus.

position. Without denying that intellect and judgment can be involved in the sensory perception of primary qualities and must be involved for the perception of object-kinds, I propose that for Descartes the primary qualities of objects can come to conscious perception without the intervention of reasoning or habitual calculation. This includes the basic operation of the "natural geometry" of the visual system, which I interpret as a mechanical process whose result affects the mind according to laws of psychophysiological correspondence (an institution of nature). This interpretation requires a reassessment not only of Berkeley's criticisms of Descartes on natural geometry but also of historical and theoretical accounts that regard a two-dimensional sensory impression as the only psychophysiologically fundamental visual impression in the early modern period and, ultimately, the nineteenth century and beyond.

1 Distance Perception: Kepler and Before

In elaborating my thesis about Descartes' theory, I need first to sketch some background about the history of visual theory in regard to spatial perception and distance perception.

As Lindberg shows in his *Theories of Vision from al-Kindi to Kepler* (1976), the most basic questions about how the eye works and the relation between light and vision are not easy or trivial. Some principles were accepted early on: we normally see in straight lines and light has a role to play. With spatial vision, the relation between eye and object was analyzed through the visual pyramid (Fig. 3). In Euclid's optics, the notion of the visual pyramid explained some aspects of spatial perception: right and left, up and down. Euclid was an extramission theorist, holding that we perceive objects by means of visual rays emanating in straight lines from the eye.[2] His analysis simply equates the structure of visual experience with the order of the rays themselves. In particular, he equates perceived size with visual angle, leaving aside distance. Thus, the objects A and B in Fig. 4a would be perceived as having the same size, whereas the same objects when situated as in Fig. 4b would appear of different sizes.

The extramission theorist Ptolemy held that the outstretching visual rays allow us to see the objects where they are, because the rays touch them. Ptolemy held that only light and surface color are intrinsically and primarily perceptible. We see bodies with colors through visual rays, and we discern distance directly through the length of those rays (Ptolemy ca. 160 [1996, pp. 71–72, 81–82]). The perception of the spatial locations and spatial properties of objects is a tidy affair: the rays touch the surfaces of objects in a geometrically precise and orderly manner. In perceiving size, we take distance (the length of the rays) into account (1996, pp. 93–94). Thus, in Fig. 4a, object A would be perceived as being closer and smaller than object B.

[2] Lindberg (1976, pp. 12–14). Euclid is not explicit about the relation between the extramitted visual rays and light rays from a luminous body such as the sun. Subsequent extramissionists such as Ptolemy make clear that external light plays a role in vision.

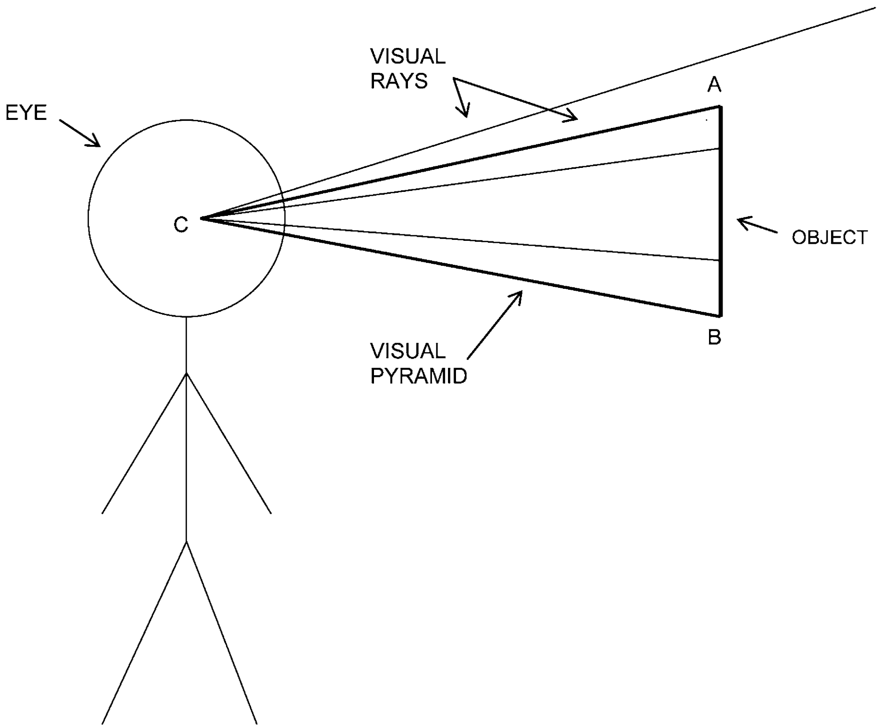


Fig. 3 The visual pyramid in Euclid, illustrating extramission. The lines originating in the eye are visual rays, not light rays. Note that there is no retinal inversion. Visual rays normally extend from the eye until they reach an opaque surface. Euclid is not explicit on whether visual rays originate from a point within the eye (as shown here) or from the eye's surface (see Lindberg 1976, pp. 12–14)

If no illusion is involved (such as occurs when the distance to the moon is misperceived),[3] each object would be perceived at its true distance with its true size. The location and distance of the visual rays would determine the perception of direction, distance, and size.

In the history of visual theory, extramission theories were supplanted by intromission theories, according to which the eye is affected by light that has been reflected from the surfaces of objects. For intromission theorists, distance is no longer "given" by the length of extramitted rays, but must somehow be discerned from the response of the eye and optical system to the incoming light. Moreover, since from every point on a normal, matte object (not a mirror) light is scattered,

[3] The moon illusion was discussed several times by Ptolemy, most cogently in his *Optics* (160/1996, p. 151). There has been some controversy in interpreting Ptolemy's various statements, on which see Sabra (1987). The important point here is that Ptolemy explained the illusion in terms of differing perceptions of distance for an object subtending the same visual angle (as the moon effectively does at the horizon and zenith).

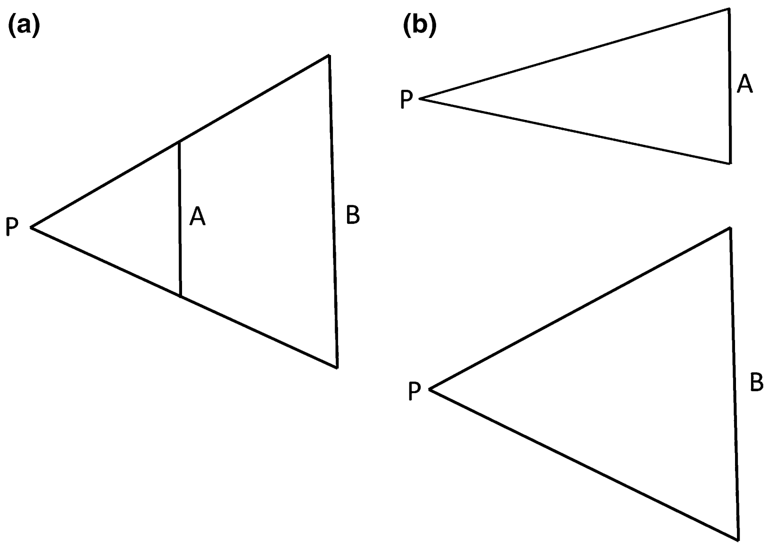


Fig. 4 Visual angle, size, and distance illustrated in the visual pyramid. Part (**a**) illustrates the relation between a common visual angle and objects A and B at different distances. Since A occludes B, one should think of the objects as seen one at a time. Part (**b**) shows the relation between visual angle and objects A and B of different sizes but at the same distance

light from each point bathes the entire front of the eye (see Fig. 1). This dispersal of the light creates a second problem not faced by the extramission theory: how to establish a one-to-one correspondence between a surface or other structure in the eye and the array of visible points lying ordered in the field of vision. Such a correspondence is required in explaining the basic fact that we perceive an orderly visual world. In its absence, the eye could sense only a blur, more profound than that experienced without one's glasses.

Ibn al-Haytham, an Islamic natural philosopher who flourished about 1020, proposed a solution to this problem. He ascribed to the eye a selective sensitivity to the rays perpendicular to the cornea and crystalline humor (Fig. 5). The other rays scattered from the object are not received. Hence, an orderly one-to-one relationship is established. Still, the problem of how we perceive distance remains. Here Ibn al-Haytham distinguished what is directly *sensed* by vision from what is achieved by the visual faculty with the aid of "discernment, inference, recognition," and "prior knowledge" (ca. 1030 [1989], II.3.71). He held that the sense of sight perceives light and color by "pure sensation" (II.3.25). It perceives distance through discernment, inference, recognition, and prior knowledge. That is, light and color are perceived in a *bare sensory act*, whereas distance perception is *cognitive* in nature: according to Ibn al-Haytham, it involves inferences of the same sort as found in conscious reasoning, albeit formed by habit and subsequently reduced to an unnoticed act of recognition (Sabra 1978; Hatfield 2002).

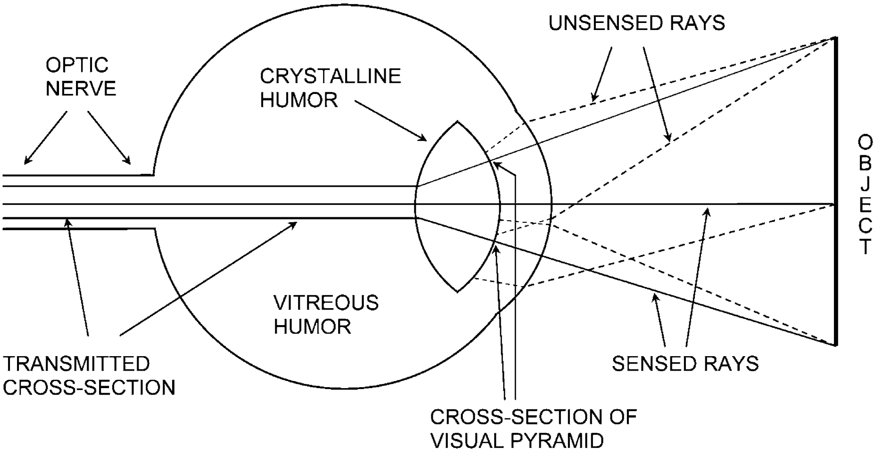


Fig. 5 The visual pyramid in the intromission scheme of Ibn al-Haytham. Light rays are scattered from each point on the object. According to the theory, only those that enter the cornea and crystalline humor at right angles are sensed. The order in which the rays intersect the cornea and crystalline is preserved as sensory stimulation is conveyed, in quasi-optical fashion, down the optic nerve (Ibn al-Haytham 1989, I.6, and Lindberg 1976, Chap. 4). This order corresponds to a cross-section of the visual pyramid (Hatfield and Epstein 1979, pp. 367–68)

Al-Haytham ascribed distance perception to two factors.[4] First and foremost, he held that, through experience, we become able to know the size of various patches of ground and other objects that lie intermediate between us and a distant object. By recognizing their size, we perceive the magnitude of the distance to the thing.[5] This works for short and intermediate distances (II.3.76). Secondarily, if we have prior experience with some objects and how they look at a given distance, we may recognize the form of the object and conjecture its distance by comparing it to the previous perception (II.3.87–88). In both cases, the cognitive acts underlying the recognition are not conscious (II.3.26–42). There isn't space to consider additional details here. The important thing is that these are the only means for perceiving distance mentioned by al-Haytham and that both of them invoke cognitive operations, beyond bare sensation. We may note that, in addition, al-Haytham held that we perceive the size of objects through perceiving visual angle and distance and

[4] Al-Haytham (1989), II.3.68, distinguished between distance per se (bare non-contiguity) and the magnitude of the distance. The two factors that I discuss concern his account of the perception of the magnitude of distance.

[5] These ground intervals are learned in relation to the size of our body, including feet, arms, hands, and paces, as well as other known measures, such as the range of a spear or the flight of an arrow (II.3.151–55).

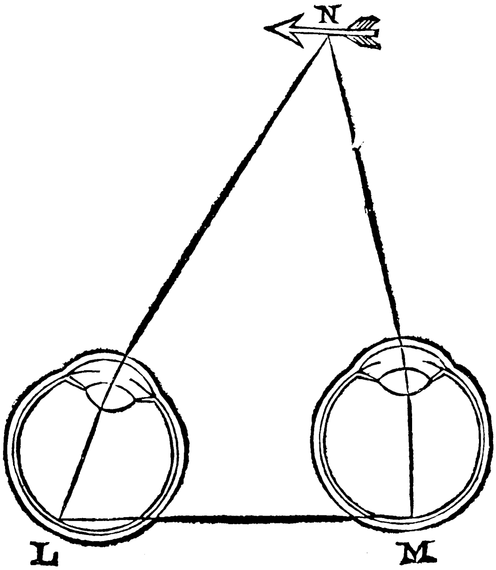


Fig. 6 The convergence of the eyes and the resulting triangle formed by inter-ocular line segment LM together with the angles formed when the eyes fixate point N. Because segments LN and MN are intended to converge on the focal point of the two eyes, the lines should be straight, not broken as in the diagram, and eye M should be shown rotated slightly to the left. From Descartes' *Treatise* (1677, p. 46, author's collection)

"imagining" the size subtended by that angle at that distance (II.3.145–48).[6] Al-Haytham's theory, including these fundamentals, passed into the Latin west.[7]

Famously, Kepler reworked Ibn al-Haytham's answer to the problem of establishing an orderly relation by proposing that the crystalline humor and the cornea act as a lens to focus the scattered bundles of light into an optical image on the retina (Fig. 1). This solved the problem of order and led some theorists to conclude that we see by means of receiving and experiencing this image. As for distance perception, Kepler offered a new means, at least for nearby objects. He proposed that the convergence of the two eyes on a focal point allows us to perceive the distance to that point through geometrical relations (Fig. 6). According to Kepler, the *sensus communis*, or common sense, becomes aware of the distance between the two eyes by "becoming accustomed to it," that is, by a process of learning; the common sense also perceives "the turning of the eyes toward each other" and "takes note of the angles" (Chap. 3, proposition 8). Kepler invokes "the geometry of the

---

[6] As with distance perception, al-Haytham writes as if these acts of discernment occur rapidly and are unnoticed, so that we appear to perceive distance and size immediately (e.g., he speaks of how the sizes of objects "appear" [II.3.147]). However, he does not provide an explicit account of how these visual appearances are generated by habitual acts of discernment.

[7] On the reception of al-Haytham's theory in the Latin west, especially by Roger Bacon, John Pecham, and Witelo, see Lindberg (1976), Chap. 6.

triangle" in explaining how we are able to "grasp" or "estimate" the distance: "given two angles of a triangle, with the side between them, the remaining sides are given" (ibid.).[8] He does not actually say that the perceiver reasons from angle-side-angle to compute the distance; it is consistent with his wording that perceivers learn the various individual values specified by this relation, that is, learn to associate visual distances with specific angle-side-angle relations. But, either way, whether achieved by calculation or through habitual learning, Kepler's is a cognitive account.

As it happens, many subsequent authors, including the Irish philosopher George Berkeley, believed that optical writers commonly held that distance is calculated by perceiving lines and angles and reasoning about them, and this has become a commonplace interpretation of pre-Berkeleian accounts of convergence. However, Descartes' work supports another interpretation of how the geometry of convergence yields a perception of distance.

2 Perception of Distance and Position (Direction) in Descartes

Descartes' *Dioptrics* of 1637 and *Treatise on Man* of 1630–1633 contain an extensive and conceptually intricate discussion of the perception of distance in a post-Keplerian optical framework.[9] Descartes was familiar with the Latin optical tradition, including Roger Bacon and especially Witelo. He knew the work of Kepler, his "first master" in optics (AT 2: 86).

In his discussion of distance perception, Descartes partly invoked previous accounts, including a known-size account that bears a relation to al-Haytham's second factor, and a mention of atmospheric perspective as when distant objects are seen less distinctly and more dimly due to atmospheric attenuation of the light. But he focused the most attention on convergence of the eyes and accommodation of

[8] Kepler also conjectured that a one-eyed person can discern distance using a triangle in which the apex is a point on the object and the base is a diameter of the pupil, a means of discerning distance that he considered to be only one tenth as powerful as convergence of the two eyes (1604/2000, Chap. 3, proposition 9). This act of discernment occurs with "the assistance of the higher faculties" (which might be learning or reasoning) and involves being aware of the width of the retinal image in relation to the width of the pupil, variations in these two magnitudes serving to specify unique triangles with the apex on the object (Chap. 3, proposition 14), thereby solving what is subsequently called the "inverse problem" for distance perception. Turbayne (1970, p. 146), treats this triangle as equivalent to the accommodation of the eye, which it clearly is not. It is now known that the lens accommodates itself in a way that correlates with these triangles (at near distances). Kepler knew that some mechanism of accommodation is needed—he posited that the distance or focal length between a lens of constant power and the retina is changed by altering the shape of the eye (1611, prop. 64)—but he did not specify that change as a cue for distance.

[9] At least part of the *Dioptrics* was completed in 1630 (AT 1: 179), and he refers to the work in the *Treatise*, which was composed in 1630–33, revised during the 1640s, and first published in the original French in 1664. In this chapter, Descartes' work on optics, entitled *Dioptrique* in French, is referred to in English as *Dioptrics*.

the lens as cues for distance. These are not found in al-Haytham (although, in principle, he could have invoked convergence).

In the *Dioptrics*, prior to his account of distance, Descartes offers an account of how position or direction is perceived. This account is important for two reasons. First, it introduces the notion of an institution of nature with respect to a spatial property, visual direction (as well as tactual direction). Second, Descartes suggests subsequently that the perception of size and shape can be explained as consequences of perceiving direction and distance. I consider these two points in subsequent sections.

The *Dioptrics* account of "the seeing of distance"(O 105, AT 6: 137) mixes psychophysiological explanation with more traditional cognitive explanations. In doing so, Descartes also introduces a distinction between *seeing* distance (or direction) and merely *imagining* it: we see distance by means of the cues of accommodation, convergence, distinctness or indistinctness of the image, and strength or weakness of the light. However, if we already imagine the size of an object and its position, or its distinctness or the force of the light from it, then we are able "not actually to see, but to imagine its distance" (O 107, AT 6: 138–39). From Descartes' examples, of looking at a mountain or ships at sea, it seems that the act of imagining distance occurs when objects are far away. But rather than pursuing this distinction, I want to focus on the seeing of distance.

For the seeing of distance, Descartes uses a variety of terms in the *Dioptrics* and the *Treatise*. He speaks of "seeing," "sensing," "perceiving" and "being acquainted with" distance, but also of "judging" it.[10] My inclination, when he speaks of "seeing" distance, is to suppose that he means that we experience distance in a phenomenally immediate manner; that is, we find ourselves confronted by a visual world in which objects are phenomenally presented as being at various distances. However, that by itself would not imply that the experience of distance is produced by a psychophysiological process that occurs without any mediating process of habit or judgment. Indeed, a much-analyzed passage from the Sixth Objections gives reason to allow that some of Descartes' instances of apparently seeing distance in a phenomenally immediate manner involve implicit judgments or habits

[10] The nouns or verbs are: *vision*, *voir*; *sentir*; *apercevoir*; *sçavoir* (modern *savoir*) and *connoistre* or *cognoistre* (modern *connaître*); and *juger*. Celia Wolf-Devine (2000, pp. 512–13) spends some time explaining away Descartes' intermittent use of *savoir* and *connaître*, assuming that the latter but not the former has the connotation of "to be acquainted with." This may be a case of improperly projecting back more recent French usage of those words (as, respectively, "an intellectual kind of knowledge" and "to be acquainted with" [2000, pp. 512–13]) onto an earlier time in which these connotations were not dominant. Thus, in earlier usage, *savoir* could have the meaning of "being acquainted with," "having consciousness of," or simply "perceiving," and, conversely, *connaître* could have the modern sense of *savoir* as intellectual knowledge, especially in regard to concrete objects (Rey 1992, pp. 475, 1887). We are left to interpret Descartes' uses of these terms in their sentential contexts.

that go unnoticed (because they are rapid), in which case the perception of distance only seems to be unmediated. Descartes in fact attributes the account of the Sixth Replies to the *Dioptrics*.[11] That work does include what seems to be an uncontroversial case of judgment underlying the perception or seeing of distance in the case of distinctness and strength of light, where this does not involve merely imagining the distance, but seeing it by judging (O 106–7, AT 6: 138). But it also invokes psychophysiological immediacy.

The notions that distance might be "judged" according to distinctness of details and strength of light or "imagined" in accordance with known size are not new. They belong in the family of those things perceived by "discernment, inference, recognition, and prior knowledge" in Ibn al-Haytham, which included distance. According to Ibn al-Haytham and Kepler, all visual qualities except light and color are perceived by a kind of cognitive operation.

It is my thesis that Descartes, in discussing physiological mechanisms by which distance might be "seen" or made present by perceptual "acquaintance," created a new way of conceiving distance perception. He offered physiological mechanisms that directly cause the experience of distance in us according to an institution of nature. Descartes thereby reconceived the line between purely sensory products in vision and products that rely on cognitive operations such as judgment or learning. He not only accepts distance as phenomenally immediate—that is, by sight we experience things as at a distance in a manner that is as phenomenally primitive as our experience of their color—but he also makes distance, in some cases, the product of psychophysiological mechanisms that are as unmediated cognitively as is the production of a color sensation. Distance, for Descartes, is in the first instance a matter of what Ibn al-Haytham called "pure sensation." I say "in some cases" and "in the first instance" because I also allow judgmentally mediated instances of seeing distance, as in the Sixth Replies. In the remainder, my focus is primarily on Descartes' historically novel psychophysiological account as explicated in his full accounts of vision in the *Dioptrics* and *Treatise*.

[11] Descartes says in the Sixth Replies: "And indeed that size, distance, and shape can be perceived one from another by means of calculation alone I demonstrated in the *Dioptrics*" (AT 7: 438; CSM 2: 295*). This sounds as if the *Dioptrics* offered only a cognitive account of the perception of those properties, which is contrary to fact. Descartes may have had several reasons for this oversimplification, the most obvious being that, in responding to the objector's question about error being corrected by reason or the senses, he focused on the distinction between unnoticed judgments (which may frequently contain errors) and studied judgments (which do so less frequently) and so referenced only cognitive factors from the *Dioptrics*. On this point, see Hatfield and Epstein (1979, p. 378). Vinci (1998, pp. 129–30), emphasizes the role of unnoticed judgments in explaining illusions. Wolf-Devine (1993, pp. 84–88), examines the relation between the Sixth Replies and *Dioptrics*. In any event, the *Dioptrics* and *Treatise* carry greater weight, as they give Descartes' full account of distance, size, and shape perception. My discussion in this chapter is distinctive for making full use of the *Treatise*.

3 Natural Geometry

Descartes' famous passage in the *Dioptrics* on natural geometry has been incorporated into various interpretations of his work according to which, in seeing distance by natural geometry, the mind *does* geometry involving lines and angles. Some interpreters hold that, accordingly, the space that we see is objectively geometricized and rationalized, which provides support for Descartes' conception of matter as pure extension having the properties of size, shape, and motion. These readings occur in differing narratives with differing aims. They share the idea that the mind does geometry in seeing distance by means of convergence and perhaps also in seeing size by combining distance with visual angle (or retinal image size—these can be treated as equivalent).

Before turning to authors who engage Descartes' actual texts in detail, I want to mention an outlook toward Descartes and geometry that takes a more abstracted view. This outlook occurs in Erwin Panofsky's examination of perspective as a symbolic form. Panofsky regards perspective itself as an attempt to "rationalize" visual space, by which he means the space of visual experience and not merely the paths that physical light takes in forming an image when focused by a lens. He links the trend toward "objectifying" and "rationalizing" visual space with Descartes, and leaves it to Kant to "formalize" this relation (Panofsky 1927 [1991, p. 66]).

This line is taken up more recently in Crary's reading of Descartes as modeling vision through the monocular *camera obscura*, according to which Descartes treats vision as detached from the body and as positing a single, ideal point from which to view an objective world. In interpreting Descartes' laboratory demonstration of the retinal image (Fig. 1), Crary writes: "The aperture of the camera obscura corresponds to a single, mathematically definable point, from which the world can be logically deduced by a progressive accumulation and combination of signs…. It is an infallible metaphysical eye more than it is a 'mechanical' eye" (1990, p. 48).

Crary is correct that Descartes appealed to a rational eye of the mind. However, contrary to Crary's account, this eye was detached from the senses and wasn't modeled on the *camera obscura*. Descartes' use of the bovine eye in his laboratory demonstration was part of his description of the visual system as an embodied sensory system, as even a cursory reading of the *Dioptrics* makes clear.[12] Descartes indeed did seek to ground his metaphysics and physics apart from sensory evidence and the mechanical eye. But he did not do so by rendering the eye or its analogue in the *camera obscura* as a rationalizing device. Rather, in justifying his geometrical concept of matter, he turned away from the senses to discover a non-sensory source

[12] Massey (2007), Chaps. 2, 4, is a helpful corrective to Crary by a fellow art historian. She challenges Crary's claim that acceptance of geometrical perspective must yield a "disembodied" eye. She notes that the *camera obscura* implies a viewer but that, as conceived by Descartes, the perspectival image on the retina in a living being involves no viewer. Indeed, Descartes specifically denies that the retinal image or its physiological correlate is itself seen, "as if there were yet other eyes in our brain with which we could perceive it" (O 101*, AT 6: 130).

of knowledge, which could then provide a framework for analyzing the senses.[13] Descartes regarded the senses as fallible but practically reliable sources of information about the nearby environment. He turned to the pure intellect, (allegedly) devoid of sensory content, for the metaphysical grounding of his physics.

Readings of Descartes' theory of perception by Nancy Maull and Margaret Wilson offer a fuller consideration of the place of reasoning in the "natural geometry" passages. Maull (1980) compares Descartes' accounts of color perception and distance perception. She rightly contends that, in Descartes' theory, color is immediately perceived through the effect of a physiological process on the mind. It is a brain-caused sensation. But, in her interpretation, distance is never perceived immediately or by sensation; it is always inferred (1980, pp. 30–33). In the domain of spatial properties, we immediately perceive a two-dimensional pattern, and we reason to the distance of objects. We then confuse this reasoning with sense perception, so that distance may seem to be immediate but is not. Accordingly, natural geometry involves a judgment using an algorithm. For Descartes, "judgements of distance… are always mediated, often by natural geometry" (Maull 1980, p. 34). Maull places this account into a larger, proto-Kantian reading of Descartes, according to which, in perceiving the world, we "reverse" the rules of perspective "to apply natural geometry and to form perceptual judgments about three-dimensional objects" (1980, p. 36). On her reading, this (intellectualized) account of perception is Descartes' way of justifying the applicability of mathematics to nature. That is, because we are doing geometry in seeing, we can expect that geometry will be an apt description (subject to limitations on sensory acuity) of the world as experienced.[14]

Wilson (1993) does not share Maull's view that Descartes sought to justify the application of geometry to nature through his natural geometry of distance perception. She considers the *Dioptrics* and Sixth Replies in examining Descartes' account of the perception of primary qualities (including especially size, shape, and position) as opposed to secondary qualities. Color and other so-called secondary

---

[13] For an account of Descartes' "fleshless eye of the mind" and its intelligible world, in contrast to vision as a sensory capacity, see Hatfield (1986).

[14] Three additional interpreters agree that Descartes' account of the senses was supposed to justify his geometrical approach to nature. Schuster (1980) uses the *Dioptrics* to elaborate a justificatory program that he imputes to the *Rules*, a program that moves via an "optics-psychology-physiology nexus" from two-dimensional patterns in the imagination to the mathematization of external bodies. He sees Descartes' mature positions emerging from the perceived failure of the o-p-p nexus to ground this mathematization. Turbayne (1970, pp. 160–61), interprets Descartes as having the perceiver perform (perhaps unnoticed) acts of geometrical reasoning via the "natural geometry." He aligns this view with Kant. Atherton (1990) discusses the differences between the Sixth Replies and the *Dioptrics*, contending that in the former Descartes posits "psychologically real" geometrical calculations that are part of his scheme to justify the geometrization of nature (1990, p. 21, citing Maull 1980); regarding the latter, she first says that it is "not at all clear" that Descartes intends the "natural geometry" to involve psychologically real calculations but concludes that "geometrical calculations enter into both theories" (1990, p. 32). By contrast, Arbini (1983) reduces the philosophical freight that Descartes expected his account of the senses to carry but nonetheless interprets natural geometry as a mental operation (pp. 321, 325).

qualities are perceived as bare sensations that "arise from the mind-body union" (1993, p. 173). She finds two ways in which primary qualities are apprehended. First, they are grasped as "abstract or general ideas with their source in the intellect"; this is what Descartes has in mind when he speaks of clear and distinct ideas of matter in the *Meditations* and *Principles*.[15] Second, in sensory perception we perceive the shape, size, position, and distance of particular bodies through "complicated constructions based to some degree on a rather opaquely specified sensory given" (1993, p. 173). This "sensory given" is the sensory core specified in the Sixth Replies as the second grade of sense, and it includes light and color as well as shape, size, and position but not distance. She reads the *Dioptrics* as (perhaps) adding to this sensory given the further factors of distinctness of shape, the muscular cue of accommodation (and presumably convergence as well), and, as background knowledge, the known size of objects (1993, p. 171). These factors enter into intellectual judgments of distance, which include "surveyor-like" reasoning. According to Wilson, in Descartes' scheme such reasoning "*must* be ascribed to the realm of mind, not body" (1993, p. 171). She analyzes this reasoning as unnoticed yet potentially conscious, taking her cue from the much-analyzed discussion of the third grade of sense in the Sixth Replies.

Wilson sees a role for the institution of nature in these matters but, as I read her, only for the sensory given, including the state of the ocular musculature in accommodation and convergence. These cues about the state of the lens and the positions of the eyes (their directions) then enter into unnoticed judgments that yield the perception of distance. As she says in summing up, for Descartes "All distance perception… involves intellectuality" of the sort found in the third grade of sense. In doing natural geometry, the mind undertakes surveyor-like reasoning that becomes habitual and unnoticed. The results of this habitual reasoning support the usefulness of sensory perception for knowing the size, shape, and position of particular bodies, but she does not assign such reasoning the primary role in justifying Descartes' claim that primary qualities are better known in objects than are secondary qualities. That justificatory work, for Wilson, is grounded in the intellectual apprehension of ideas of primary qualities "generally" or "abstractly" (1993, p. 164).

Although there is some textual support for supposing that Descartes interpreted natural geometry to mean genuine calculation, I wish to open a space for another interpretation in which natural geometry relies on psychophysiological mechanisms and does not involve unnoticed reasoning. My case relies on both the *Treatise* and *Dioptrics*.

---

[15] Although Wilson does not emphasize this, it is helpful in thinking of primary qualities as modes of extension to recall that Descartes purported to grasp extension in a purely intelligible manner, independent of sensory ideas (see the beginning of the Sixth Meditation, AT 7: 72–73). Extension considered in this way need only be "abstracted" from sensory qualities by turning away from them, not by "generalizing" from concrete instances. Indeed, in the cited passage, Descartes speaks of the intelligible grasp of particular figures, such as a pentagon or a chiliagon.

Fig. 7 The illustration of the blindman with crossed sticks from the *Treatise*. Hands *f* and *g* together with sticks *i* and *h* form a triangle with apex K. Reproduced from Descartes (1692b, p. 68) (author's collection)

Here is the passage on "natural geometry" in the *Treatise*, describing first a blindman with sticks (Fig. 7) and then the two eyes (Fig. 6):

> Notice also that if the two hands *f* and *g* each hold sticks *i* and *h* with which they touch the object K, then even though the soul is otherwise ignorant of the length of the sticks, nevertheless, because it will be aware of [*sçavoir*] the distance between the points *f* and *g*, and the sizes of the angles *fgh* and *gfi*, it will be able to be acquainted with [*connoitre*], as if by a natural geometry, where the object K is. And in the just the same way, if the two eyes L and M are turned toward the object N, the size of the line LM and of the two angles LMN and MLN will make it aware of [*connoitre*] where the point N is. (TM 11: 160*)

The corresponding passage in the *Dioptrics* says essentially the same thing.

On the face of it, this looks like a cognitive account of natural geometry. The blindman "is aware of" or "knows" the distance between his hands and the angle of the sticks—and it is certainly plausible that the blindman could be aware of his hands and the direction of the sticks. It would seem that he must, then, perhaps by habit, or rapidly and without noticing it, calculate by angle-side-angle the distance from the base-line to K, and hence the location of K. Presumably, the same goes for vision, although the corresponding appeal to the perceiver's awareness of retinal locations is less plausible (as discussed below).

Still, there is reason for at least a moment's doubt, since the passage can be read as saying that, for touch, these lines and angles allow the man "to be acquainted with" or "to be aware of" the location of K. Similarly, with the eyes, their separation and rotation make the mind "acquainted with" or "aware of" point N. The outcome of natural geometry is described more as a phenomenal awareness than the product of judgment. Is this phenomenal awareness a product of judgment or ingrained habit, or might it be produced more directly by brain mechanisms?

Let us consider more closely the visual perception of position. In the *Dioptrics*, Descartes defines position as "the direction in which each part of the object lies with respect to our body" (O104, AT 6: 134).[16] In both the *Treatise* and *Dioptrics*, he starts with the blindman and the sticks. When first introducing the visual perception of position in the *Treatise*, he says that when the eye is directed toward an object, "the soul will be able to tell the position of this object, inasmuch as the nerves from this eye are disposed in a different way than they would be if it were turned toward some other object" (G131–32, AT 11: 159).[17] In the corresponding portion of the *Dioptrics*, he elaborates this account, saying in effect that our awareness of the position of the parts of our bodies arises from "the position of the small points of the brain whence the nerves originate":

> For this position, changing ever so little each time that of the members where the nerves are inserted changes, is instituted by nature not only in order that the mind may be aware of how each part of the body which it animates is placed with respect to all the others, but also so that it may transfer its attention from there to any of the locations contained in straight lines that we can imagine to be drawn from the extremity of those parts and prolonged to infinity. (O 104*, AT 6: 134–35)

This passage is illustrated with a blindman with crossed sticks, who can attend to any point along the direction of the stick or beyond it, as a result of an "institution of nature."

The passage continues: "when our eye or head turns in some particular direction, our mind is informed of this by the change which the nerves, inserted in the muscles that serve these movements, cause in our brain" (O 105*, AT 6: 135). Thus, Descartes explains, in the eye RST (Fig. 8), a part of the brain 7, 8, and 9 "enables the mind to be aware of all the locations along the line RV, or SX, or TY" (O 105*, AT 6: 135).

[16] Descartes defines position (French *situation*) as direction. Position/direction together with distance yield "location," a definite location in space (French *lieu*). He expresses no commitment to *experiencing* direction and distance apart from one another, but the two factors do have different explanatory mechanisms, drawing on different muscle operations.

[17] In the *Treatise*, Descartes offers two accounts of perception of the "position, shape, distance, size" and other qualities, including use of sticks. The first passage closely relates to the discussion in the *Dioptrics* and includes the mention of "natural geometry" (AT 11: 159–63); this passage does not discuss pineal physiology, and indeed prior to this passage the pineal gland has only been mentioned indirectly ("a certain little gland," AT 11: 129). The second passage (AT 11: 181–89) occurs after the role of the pineal gland in sensory and motor processes has been discussed, and it offers a deeper mechanical account of perception of location (position/direction and distance) with the two sticks and of the brain mechanisms involved in perceiving distance by means of accommodation and convergence. I discuss this second account in the next section.

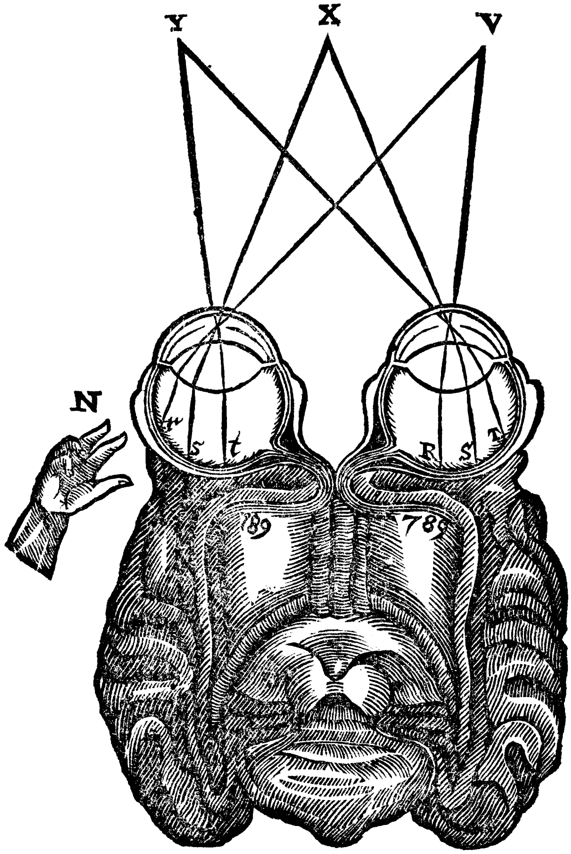


Fig. 8 The binocular visual system from the *Dioptrics*. Note that the portrayal of visual physiology is incomplete by comparison with the fuller account in the *Treatise*, as the role of the pineal gland is not illustrated, although the gland is referred to indirectly in *Dioptrics* IV. Reproduced from Descartes (1692a), p. 65 (author's collection). (The letters Y, X, V have been inserted at the top of the image, in accordance with the original diagram of 1637)

These passages and diagrams (Figs. 7 and 8) can suggest a certain picture of how perception of direction works and how natural geometry works. The passages may seem to suggest that we feel the location of our hands and eyes through muscle sensations mediated by brain states. By feeling the state of the muscles, we know the position of the limb, head, hand, or eye. Following this line of interpretation, we feel the place in the eye where light impinges, and we then project back along the line of sight, that is, along lines RV, SX, or TY.[18] If X is the focal point of convergence, we perceive it at the place where the lines sX and SX cross, as projected back from s and S on the retinas. To get the distance, we take account of

[18] The notion of ray tracing is endorsed by Maull (1980, pp. 30, 36). Atherton articulates the notion explicitly: "Descartes had offered a whole slew of ways in which retinal information is enriched, some of which, those that fit into his intellectualist account, imagine that in seeing, we construct a visual picture by reasoning from the information contained on the retina. Descartes's account of situation perception is like this. The visual system is assumed to be working out the orientation of the objects it sees by tracing them back from the images that appear upside down on the retina" (1997, p. 152).

the distance sS and the angles XSs and XsS, and calculate the lengths of lines sX and SX to find location X.

I don't deny that this reading fits the passages to a certain extent. But I believe the fit is only apparent, because it rests on a misunderstanding of Descartes' theory of muscle function, awareness of limb position, and psychophysiological correspondence. It introduces a psychological projection along lines of sight. True, Descartes says that we are able to "be aware of the all the locations along the line RV, or SX, or TY." But there are other models of what is going on with brain, muscles, and eyes that can account for this.

The first hint of a problem for the projectivist account is that, in the case of the blindman, Descartes says that in order to be aware of the positions of the points along the sticks and continuing in the same direction, "he does not need to be aware of or to think at all of the locations of his two hands" (O 105*, AT 6: 135).[19] One would expect that, in order to project away along a direction from the hands, one would need at least to be aware of where the hands are. But Descartes says that in his account of the perception of positions along these lines, no such awareness is required. What could he have in mind?

The answer lies in the much fuller account of the operation of the muscles, brain, eyes, and limbs given in a second portion of the *Treatise*, which occurs after Descartes has introduced a role for the pineal gland in controlling and perceiving limb position and in the perception of distance. This fuller account clears up matters left ambiguous or unexplained in the *Dioptrics* and the earlier sections of the *Treatise* cited thus far.

4 Limb Position Later in the *Treatise*

The *Treatise* is Descartes' most complete account of his machine psychology. It is written as a fable, in which a hydraulic machine is created that looks like a human being and acts in some ways like one, without the benefit of a soul or mind. We are clearly intended to discharge the fable and to pay attention as Descartes describes in

[19] An interpretive problem arises in comparing this statement with the passage in the *Treatise* (AT 11: 160), which speaks of the soul "knowing" or "being acquainted with" (*sçavoir*) the distance between the two hands. There are several ways to resolve this apparent discrepancy. One might note that in the *Treatise* this passage belongs to the first treatment of distance perception and so doesn't yet make use of pineal physiology to account for how the soul can know all of the locations in question (hands, direction of sticks, crossing point of sticks; visual direction, location, and distance) without calculation but by a psychophysiological "natural geometry" as proposed below. One might read Descartes in the *Dioptrics* as hinting at the brain-mechanism account for perceiving direction and then leaving open the mechanism underlying both that and "natural geometry," since in that work he has not developed his full pineal physiology. On this reading, only the second passage from the *Treatise* reveals the full psychophysiological account. Other options include attributing to Descartes two accounts of natural geometry (cognitive and psychophysiological), or supposing that he changed his mind, that he contradicts himself between the first and second passage in the *Treatise*, or attributing him differing accounts of touch (where the positions of the hands are accessible) as opposed to visual direction and location.

detail mechanisms that might explain all the behaviors of nonhuman animals and many behaviors of human beings, without invoking mind. He also makes provision for ways in which states of the brain will interact with mind, when a mind or soul is added to the machine.

As regards sense and motion, the machine operates via a subtle liquid that rises from the heart in the blood and is distilled out at the base of the brain (AT 11: 128–30). As Descartes ultimately explains (AT 11: 170–97), these “animal spirits” flow from the pineal gland, located at the center of the brain, into various tubes that lead to the muscles. They cause motion by inflating and contracting the balloon-like muscles (Fig. 2). The flow of the spirits is controlled by fibers within each tube, which come from the sense organs. Hence, the motion of the animal or mindless human body is controlled by a kind of sensorimotor feedback loop, mediated by sensory fibers, which change the opening of the tubes, causing a limb to move.

In this scheme, in an ensouled machine sensations from the five senses arise according to how the fibers in the nerve tubes are stimulated by activity at the sense organs, causing in turn a subtle variation in the opening of the brain tubes and so in the flow of spirits from the pineal to the tube opening (Fig. 9a). Various sensations are caused in the mind by variations in the outflowing spirits.[20] If an optical nerve fiber is jiggled one way and produces one sort of pulse in the spirits flowing out, that might produce a sensation of a certain color. If it is jiggled another way, that produces a sensation of a different color. If yet other fibers are jiggled, coming from ears, nose, or mouth, then sounds, odors, or flavors are experienced.

Now, interestingly, Descartes does not account for our perception of the location of our limbs via sensory fibers that come from the muscles and inform us of their state of contraction. Rather, he has what in present-day terminology is called an “outflow” theory of proprioception,[21] or perception of the locations of bodily parts. According to an outflow theory, the brain does not determine the location of bodily parts by kinesthetic information flowing in from the muscles, but rather keeps track of where it “told” the limb to move (the position entailed by a certain motor command), which allows the perceiver to know where his or her limb is. Ascribing such a theory to Descartes might seem like a pun on the outflowing of the animal spirits, but it isn’t. It is significant because Descartes does not speak of muscle feelings but of the perception of limb location itself.

Descartes explains how the direction of the limb is perceived by considering the woman’s arm in the Fig. 2. She has her eyes focused on an arrow. At first, her arm is directed at point B, under the control of the spirits issuing from point *b* on the pineal gland, a point that also sends spirits to optic nerve tubules that have been opened by light reflected from point B and affecting the optic nerves 3–4 and 3–4 in the two eyes. He says:

---

[20] Some readers find it counter-intuitive that the animal spirits should cause a sensation by flowing *out* from the pineal gland. In fact, Descartes is free to propose any “institution of nature” that he likes, as long as there is psychophysiological regularity. His entire system of sensorimotor control is based on the spirits flowing out.

[21] On modern outflow theories, see Gregory (1997, pp. 101–5).

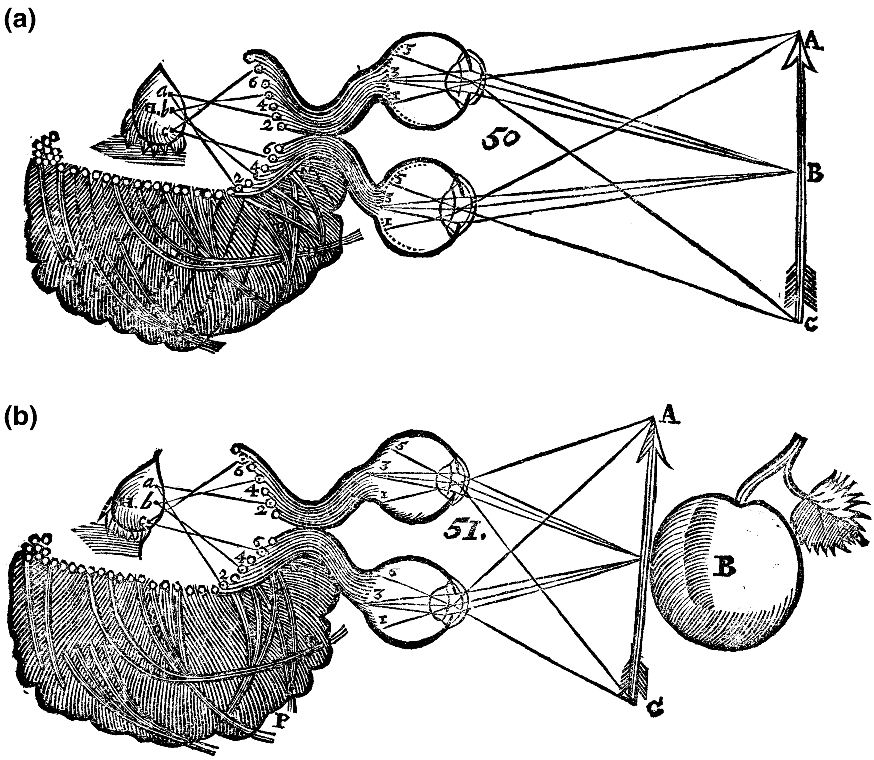


Fig. 9 The visual system in the *Treatise*, showing pineal physiology. Diagrams **a** and **b** exhibit the differences in pineal lean when the visual system is directed toward objects at different distances. Reproduced from Descartes (1692b), pp. 106, 120 (author's collection)

> We can assume that what makes tube 8 turn towards point *b* rather than toward some other point is simply that the spirits that issue from this point tend toward it with a greater force than do any others. And the same thing will cause the soul to sense [*sentir*] that the arm is turned toward object B, if it is already in this machine. (G153, AT 11: 181)

He further explains that if the spirits come from point *c*, they will cause the arm to move and point toward C, again causing an "idea" and a sensing of this movement of the arm. He announces a general principle concerning knowledge of limb locations, dependent on the physiology of spirit flow and tube openings:

> in general we should take it that each tiny tube on the inside surface of the brain corresponds to a bodily part, and that each point on the surface of the gland H corresponds to a

> direction in which these parts can be turned: in this way, the movements of these parts and the ideas of them can cause one another in a reciprocal fashion.[22] (G 154–55, AT 11: 182)

The mind becomes aware of limb location through the outflowing spirits which cause the muscle state that places the limb in that location. Descartes does not speak at all of *feeling* the muscles and their state of contraction; he says simply that as a result of spirit flow, we "sense" the location of the arm or other limb.

5 The Physiology of Distance Perception

A psychophysiological explanation of natural geometry arises from further specification of the variations that can occur in the pineal gland and the sensory experiences they cause in the mind by the institution of nature.

In his physiology of sight, Descartes explains that the convergence of the eyes is controlled by the flow of the spirits. In principle, he might have theorized that we know of the rotation of *each* of the eyes by the outflow theory, and that we use awareness of *where each eye is* to calculate distance by angle-side-angle. However, that isn't his account. Not only does he not speak of feeling where the eyes are, but he instead imagines a mechanism in which the two eyes are treated as a single system, which causes a certain state in the pineal gland, which directly causes the idea of the location (position and distance) of the object of fixation of the two eyes.

This mechanism consists of the lean of the pineal gland. The farther it leans forward, the closer the object is perceived to be. The mechanism is illustrated in Figs. 9a, b. When the arrow is nearer, the gland leans forward (Fig. 9b). Descartes offers this as a general account of how the idea of distance is produced in the mind by the pineal gland. Referring to the figure of the woman pointing (Fig. 2), he offers this account of the idea of distance as regards point B:

---

[22] In these passages, Descartes sometimes speaks of corporeal ideas, which he introduced earlier as "figures … traced in the spirits on the surface" of the pineal gland (AT 11: 176). These corporeal ideas play a significant role in his machine psychology (see Michael and Michael 1989; Hatfield 2012, p. 172), although he subsequent drops the terminology (AT 7: 160–61). In an ensouled machine, they are the brain states ("forms or images") that the soul "will consider directly when it imagines some object or senses it" (AT 11: 177). In discussing the soul's relation to the pineal gland, Descartes often echoes the scholastic Aristotelian terms used to describe the rational soul's (or the rational power of the rational soul's) relation to images in the corporeal imagination. He uses terms such as Latin *convertere ad*, "turn toward" (AT 7: 73) and French *considerer*, "consider" (AT 11: 177); see Thomas Aquinas (1889, I.85.1, ad 2), the agent intellect "turns toward" the phantasm (*conversio supra phantasmata*, I.85.1, *passim*), the intellect "considers" (*considerare*) the common nature in the species or phantasm, and Eustace of St. Paul (1638, p. 298), the understanding "observes" (*speculari*) and "engages" (*versari*) phantasms. We might expect him to fall back on such usage, given the comparative novelty of his psychophysiological institution of nature, in which ideas in the mind are caused by material states that need not resemble the representational content of the idea. In this connection, the passage on grades of sense from the Sixth Replies suggests a picture in which the pineal state creates a two-dimensional image in the mind (see Hatfield and Epstein 1979). The *Dioptrics* and *Treatise* accounts of the perception of direction and distance open up other possibilities.

> You will readily accept this if, in order to understand what the idea of the distance of objects consists in, you assume that as the gland's position changes, the closer points on its surface are to the centre of the brain *o*, the more distant are the places corresponding to them, and that the further the points are from it the closer the corresponding places are. Here, for example, we assume that if *b* were pulled further back, it would correspond to a place more distant than B, and if it were made to lean further forward it would correspond to a place that was closer. (G155, AT 11: 183)

In this way, Descartes tells us, the mind or soul can "sense" (*sentir*) the locations of objects (ibid.). In such cases it does not calculate, judge, or rely on ingrained habit. Rather, a certain order exists among brain states such that, when the eyes are converged on an object at one distance, an idea arises of the focal point on that object at that distance, and when the eyes are differently disposed, the object is sensed as being at a different distance. I take these "sensings" of distance to be phenomenal presentations of a location in a direction at a distance. These are not judgments or opinions, but *experiences* of something at a distance.

In this scheme, the perceiver does not project back along the lines of sight from each eye. Rather, by means of a unified mechanism involving both eyes, the perceiver experiences a single object located in relation to his or her body and head. There is no psychological projection from the two retinas; rather, the direct phenomenal experience of a single object in a location arises as a result of the total pineal state. The lines drawn from the object to the two eyes allow the theorist to determine where the object is perceived (if distance and direction are veridically perceived). But these lines are not represented by the observer or the observer's visual system. Rather, the observer's visual system is constructed in such a way that differing convergences of the two eyes (with the head in a specific position) is yoked to a brain mechanism that directly causes the experience of an object (or point of fixation) at the appropriate place.

This mechanism can also account for the blindman's ability to attend to various locations along the sticks he is holding (G 155, AT 11: 183–84). Without moving the sticks, the gland can lean to one position or another, causing an experience of the distance at that point. The mind need not imagine a geometrical line and project points along it to a location on the stick; rather, the mind senses various points in their spatial locations, and this forms a sequence along the line of the stick itself. As regards sense experience, the locations of the points come first, and any imaginative act of projection would be built on top of that.

In the account of the accommodation of the lens in the *Treatise* and *Dioptrics*, I find a confirmation of this account of not judging but "seeing" distance directly. In the *Treatise*, Descartes describes the accommodation of the lens as being required to form a clear image on the retina (G 128, AT 6: 156). In the *Dioptrics*, he explains how these changes allow the mind to see things as being at various distances, without any cognitive act:

> as we change [the shape of the body of the eye] in order to adjust the eye to the distance of objects, we also change a certain part of our brain, in a way that is instituted by nature to allow our soul to perceive [*apercevoir*] that distance. And this we ordinarily do without reflecting on it, just as when we squeeze some body with our hand, we adjust our hand to

> the size and shape of the body, and thus feel it by means of the hand without having to think of these movements. (O 105–6*, AT 6: 137)

Presumably, to each state of accommodation there corresponds a degree of lean in the pineal gland and an idea or experience of the focal object as being at a certain distance.

6 Perception of Size and Shape

The psychophysiological account might be extended further, to include the perception of shape and size. In the *Dioptrics*, Descartes makes an intriguing claim:

> As to the manner in which we see the size and shape of objects, I need not say anything in particular, inasmuch as it is all included in the manner in which we see the distance and the position of their parts. (O 107, AT 6: 140)

This passage suggests that perception of size and shape can be reduced to the perception of the direction and distance of points. And indeed, perception of direction and distance is sufficient to specify a unique surface in front of an observer, and thus to determine the sizes and shapes of visible surfaces. That is, if for a stationary observer a phenomenal distance is assigned to every phenomenal direction,[23] then phenomenal shape and size are fixed.[24]

---

[23] This assumes that phenomenal directions are orderly, and indeed phenomenal direction is usually perceived with a high degree of accuracy. A question can arise of whether phenomenal direction is keyed to each eye separately or to a single binocular system; above we found that Descartes defines direction in relation to the body, which suggests the latter option, even if his diagramatic constructions use geometrical relations to the two eyes. As mentioned below, there are also questions concerning positions off the point of fixation, which Descartes did not address, although he did hold that clear vision occurs only around the point of fixation, suggesting that scanning is needed to take in whole objects (*Dioptrics*, O 96, AT 6: 123).

[24] See Gogel (1990) for this construction of phenomenal space, including size, orientation, and shape. Using this construction, ovals and diamonds on the retina could make us perceive circles and squares (O 90, 107, AT 6: 113, 140–41), if appropriate distances are assigned according to the set of visual angles determined by the ovals and diamonds, thereby yielding the perception of a circle or square at a slant relative to the eye. Similarly, Descartes' talk of "comparing" distance with "the size of the images that [objects] imprint on the back of the eye" (O 107, AT 6: 140)—an expression that makes no sense, as observers have no direct access to the retina and the soul doesn't seem to simply experience the pineal image—might be translated into the idiom of visual angles and distances, or directions and distances. A larger image on the back of the eye corresponds to a larger angle. In this way, the "sensory core" of the second grade of sense in the Sixth Replies might be rendered as colors in a direction. If we reinterpret the distinction between the second and third grades in line with the psychophysiological account, then the sensory core would include colors in a direction at distances, at least for objects close at hand; if these distances are not determined psychophysiologically, they might still be fixed cognitively (out to a distance of one or two hundred feet [O 111, AT 6: 144]). This revision would not harm Descartes' answer to the Sixth Objections regarding the role of intellectual judgment in solving the conflict between touch and vision. It requires a significant reinterpretation of the notion of bare sensation in Descartes, in line with the richer theory of the *Dioptrics* and *Treatise*.

The determination of phenomenal shape and size follows from phenomenal distance. It occurs whether the phenomenal distance is accurate or not. As in the moon illusion, if the distance is misperceived so that the object appears farther away, it appears larger, and if it is perceived to be closer, it appears smaller (as in Fig. 4a). Descartes acknowledges these points and describes the phenomenon of size constancy, contending that we perceive an object of the same size at differing distances as being "almost equal in size, at least if their distance does not deceive us" (O 107, AT 6: 140). A deception about distance might also cause an illusory shape.

Descartes' psychophysiological account is limited in several ways. As he indicated, accommodation and convergence are only effective for near distances, fifteen to twenty feet for convergence and less for accommodation.[25] Convergence only fixes the distance at the point of fixation. Accommodation and related issues of focus (G 132–33, AT 11: 159–60) may fix distance within a depth of field, but its range is short and Descartes considered it imprecise. Experiences of objects at distances beyond twenty feet presumably must be ascribed to unnoticed cognitive factors such as those described in the Sixth Replies and the *Dioptrics*, and such factors may well also be operative near at hand. But we are left on our own to provide an account of the specifics of how sensory and cognitive aspects interact, although the Sixth Replies discussion of unnoticed judgments provides a starting point that might then be applied to an expanded version of the "purely sensory" second grade of sense.

7 Natural Geometry Again

Accepting that Descartes provided a psychophysiological account of distance perception through the mechanisms of accommodation and convergence, what are the implications for his "natural geometry"? Can it be re-interpreted to fit this account? I want to suggest that, in Descartes' psychophysiological account, the physiological processes underlying convergence constitute geometrical operations that might appropriately be called a "natural geometry."

Consider the two states of the visual system in Fig. 9a, b. When the eyes are converged at the farther distance in Fig. 9a, the pineal gland is upright. When they are converged at the nearer distance, the gland leans forward. We can imagine that, to the various distances in between, there correspond a series of leans of the pineal gland, proportional to the distance (Fig. 10).

This bit of imaginative physiology on Descartes' part contains the seeds of a mechanical interpretation of geometrical relations and their (approximate) proportional values. If we accept this as Descartes' "natural geometry," it seems that he

[25] In the *Dioptrics*, he says that accommodation is only sensitive out to four or five feet and in any case is imprecise, and that convergence is also effective only for a "short distance" (O 110, AT 6: 144). In the *Treatise*, he gives fifteen or twenty feet for convergence, and two or three feet for accommodation (G 138, AT 11: 162).

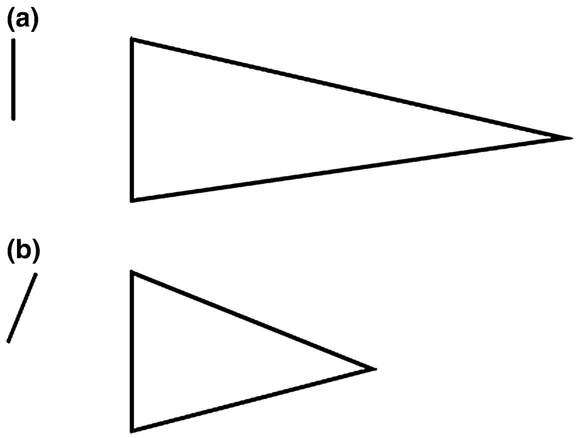


Fig. 10 Proportional relations between the triangles of convergence and the lean of the pineal gland. Diagram **a** shows a farther distance, and corresponds to Fig. 9a; accordingly, **b** shows the nearer distance, as in Fig. 9b

imagined the convergence mechanism as doing the equivalent of geometrical evaluation, without any mental involvement. That is, the mechanism offers a mechanical operation that, as a result of the psychophysiological institution of nature that pairs leans with distances, yields mathematically ordered phenomenal locations for point N (Fig. 6). In effect, the lean of the pineal gland is part of a mechanism that does geometry merely as a result of the turning of the eyes, given the fixed distance between them. It need be only an approximative mechanism in relation to actual distance, since Descartes has acknowledged that we may get distance wrong and that accommodation and convergence are imprecise. But it may be considered a material realization for making a brain state covary with the location of the apex of a triangle.

It seems clear that Descartes intends the various leanings of the pineal, given differing convergences of the eyes, to accord with angle-side-angle relations of the geometry. In this scheme, geometrical relations are incorporated into optical convergence in a mechanical mode.

8 Proportional Compasses

It should be no surprise that Descartes could believe the following to be possible: geometrical proportions effected by mechanical means. During the sixteenth century, various proportional compasses were developed that included trigonometric computations among their functions. These allowed even those with poor mathematical skills to discover the distance to a fixed but inaccessible point by taking observations using the compass as a sextant.

Closer to home, in 1619 Descartes himself invented a mechanical device that solved a much more difficult mathematical problem than locating the apex of a triangle in relation to a base line. He invented a geometrical compass that could find two or more mean proportionals. He designed a second device to allow the trisection of an angle. This work offered mechanical solutions to two of the three classical problems in mathematics that could not be solved by ruler and compass, trisection and the "Delic problem" of doubling the volume of a cube, which is solved by finding two mean proportionals between a pair of given values.[26] (The third problem is squaring the circle).

Descartes was by no means the first to offer solutions to these problems. But his solutions are remarkable for the new devices he designed. He described the construction of devices that provide solutions merely through mechanical manipulation. Consider the trisection of the angle first (Fig. 11). The device consists of four arms hinged at *a*. Segments *af*, *ai*, *ak*, and *al* are equal, and at points *f*, *i*, *k*, and *l* segments *fg*, *ih*, *kg*, and *lh* are attached, themselves being equal in length to *af* and the rest. The rods *fg* and *kg* are attached together to arm *ac* in such a way that point *g* can slide along the arm as the device is opened or closed, and similarly for *ih* and *lh*, attached to arm *ad*. The triangles *afg*, *akg*, and *alh* are always congruent, and the angles *bac*, *cad*, and *dae* are always equal. As Descartes says, none of them can be changed without changing the others (AT 10: 240).[27] Because the three angles are always equal, they trisect the angle *bae*, whatever it is. To trisect any given angle, *a* is placed at the vertex and the arms *ab* and *ae* are made congruent with the sides of the angle. The trisection can then simply be read off. A precise geometrical value arises without mental calculation (for an idealized compass).

Descartes had previously solved the problem of finding two mean proportionals with the compass shown here (Fig. 12), developed in 1619 and pictured in the Geometry of 1637 (O 228–29, AT 6: 442–44). The compass has other geometrical and algebraic properties that interested Descartes, but in his Private Thoughts from 1619 he describes how it solves the Delic problem of doubling the volume of a cube

[26] The word "mechanical" has the broader meaning of a device whose operation is understood in terms of the size, shape, position, and motion of its parts, and that might be realized in a material instrument. It is related Descartes' distinction between properly geometrical curves, which are "exact," and merely mechanically produced curves (such as the spiral), which are not (O 190–91, AT 6: 388–90). Descartes uses this distinction to separate those geometrical constructions that can be made by mechanisms (including ruler and compass) which are clearly understood geometrically from those which are not. He excludes the "merely mechanical" curves from geometry proper. But his contrast between geometrical and mechanical can also apply to the difference between idealized constructions using ruler, compass, and his new compasses, and imperfect diagrams made using those devices. When considered as "solving" trisection and the Delic problem, Descartes' instruments are understood in an idealized geometrical manner, in the same way that ruler and compass allow Euclid to bisect an angle, even if the diagram on paper is imperfect. By contrast, with the physiological mechanism of convergence, the solution for the location of the apex need not be precise mathematically, consistent with the use of actual proportional compasses to survey distances or to provide trigonometric values (which, in being mechanical, is imprecise). Finally, Descartes of course did not solve the problems in their original form, with compass and straightedge alone. It was later shown that they cannot be solved in that way. (On Descartes' compasses, see Shea 1991, Chap. 3, and Bos 2001, Chap. 16.4).

[27] Although the compass itself (considered as an ideal object) can solve the problem, Descartes was interested in its use to produce a (noncircular) curve that allows a further construction; hence, he goes on to describe using the compass the inscribe a curve made by point *g*, which can then serve to construct a trisecting angle by setting a compass at point *f*, open to length *af*, and marking where it crosses the curve made by *g* (see Bos 2001, pp. 237–39).

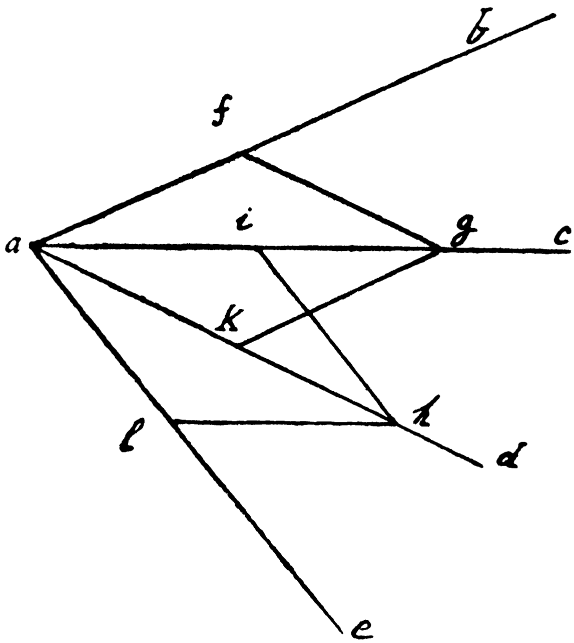


Fig. 11 Descartes' device for trisecting an angle. From Descartes' *Private Thoughts* of 1619. See text for further explanation. Reproduced from Descartes (1897–1913), 10: 240

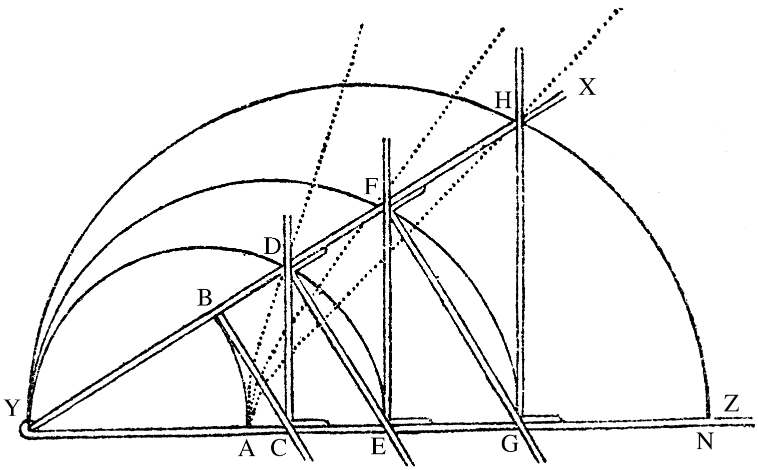


Fig. 12 Descartes' proportional compass, or mesolabe, which, among other things, allows two mean proportionals to be found between known values. See text for further explanation. From Descartes' *Geometry* of 1637. Reproduced from Descartes (1897–1913), 6: 391

(AT 10: 234–40). In the compass, arms YX and YZ are hinged at Y. Segment BC is fixed perpendicular to YX. Segment CD is attached to arm YZ so that it can slide along, pushed by BC. Similarly for DE, EF, and the others. One finds a whole series of mean proportional relations to YB simply by opening the compass! To solve the Delic problem, Descartes would consider YB as the side of the original cube and open the compass so that YE equaled twice YB. YC is the length of a side of the doubled cube.[28]

These mechanical devices for finding proportions suggest the following reading of Descartes' "natural geometry." The two eyes are a constant distance apart. They are considered a single system, which converge on focal point B (Fig. 2). For any given triangle, the location of the apex is thereby given. In the physiology of the visual system, the pineal gland controls this convergence by mechanical means, in such a way that there is a unique lean to the pineal gland for each location of the apex, or at least approximately so.[29] Accordingly, there is a one-to-one correspondence between apex locations and states of the gland. It is then a psycho-physiological law, instituted by "nature" or by God, that each lean of the gland yields a unique experience of the location of the focal object in relation to the head or, taking into account that the pineal also controls head rotation, in relation to the frame of the body. The perception of distance need not be perfect, as befits the shortcomings of a "natural" material system.[30]

This system has the interesting properties that it treats the binocular eye system as a unit and that it does not require a psychological projection back from retinal locations. The lines and angles involving the two eyes and retinal locations are useful *theoretical constructions* for understanding how the system works, but they do not enter into the psychophysiology as elements that are mentally represented or are "felt" psychologically. Rather, the idea of the focal object at a distance is produced directly and immediately by the pineal lean. Accordingly, there is no need to suppose that the subject feels the state of rotation of the eyes or the tenseness of the muscles. The operation of the muscles is simply part of the mechanical system that produces the degree of lean in the pineal gland. The observer need not be cognizant of the rotation of the eyes or the distance between them or the convergent

---

[28] As before, the proportions are only exact in an ideal version; they are approximate in the material compass.

[29] The institution of nature will need to be arranged to allow for the fact that the triangles of convergence are not all isosceles, and also that what is wanted is a solution for the length of the line from the midpoint of the base (between the two eyes) to the apex. Of course, I have not suggested that the two previously described devices of Descartes' design would solve this problem, only that, from them and his presumed knowledge of proportional compasses more generally, Descartes could imagine the type of system needed for a physiological natural geometry of convergence.

[30] For physiological purposes, the co-variation between pineal lean and actual distance, even within twenty feet, need not be exact. Presumably, the divine "institution of nature" that yokes pineal lean to perceived distance is exact, but the accuracy of the resulting experience of distance reflects the imprecision of the material, physiological processes. In the Sixth Meditation, Descartes examines various problems of accuracy in linking a soul to a machine.

lines of sight in order to experience the object at a distance. In this sense, Descartes can say that our visual systems, by means of a natural geometry, can make us *see* and *be aware of* distance in a direct and immediate fashion.

9 Implications

At the beginning of Berkeley's *Essay towards a New Theory of Vision*, there is a famous passage that purports to state a commonly accepted view among philosophers and optical theorists:

> It is, I think, agreed by all that distance, of itself and immediately, cannot be seen. For distance being a line directed end-wise to the eye, it projects only one point in the fund of the eye, which point remains invariably the same, whether the distance be longer or shorter. (Sect. 2)

Berkeley appears to have in mind that since each point on the retina receives stimulation as if in a point, there is no basis in optical stimulation for seeing distance, rather than merely judging it or otherwise coming to experience it based on associative connections with optical cues. But Descartes says that we see distance immediately, so his position belies Berkeley's claim to be asserting a commonplace.

Berkeley was of course aware of Descartes' claim that convergence allows us to perceive distance, but he does not consider Descartes' position, as he understands it, to contradict his assertion that distance cannot be seen—for two reasons. First, he takes Descartes to hold that we are *aware of* the lines and angles in convergence and *compute* the distance by their means. Accordingly, Descartes' position, as Berkeley interprets it, does not contradict him, because that position says that distance perception is mediated by awareness of lines and angles. But, second, Berkeley has it that Descartes' position is obviously false because it says that the perceiver must be aware of the lines and angles posited in optics, and Berkeley finds it manifestly evident that ordinary perceivers are not aware of such entities. As he puts it, "those lines and angles, by means whereof some men pretend to explain the perception of distance, are themselves not at all perceived" (1709, Sect. 12).

Berkeley's interpretation of Descartes' natural geometry has set the agenda for reading Descartes' texts. Whether interpreters have noticed it or not, they have read Descartes' account of convergence as involving muscular feelings, of the sort Berkeley himself posits. The issue has then come down to whether it is plausible that the feeling of ocular muscle rotation could feed into geometrical calculations of the sort accompanying mental-computation readings of Descartes' natural geometry.

But Descartes' natural geometry as I have reconstructed it does not require the perceiver to be aware of or mentally to represent the lines and angles of convergence or of optical projection. So Berkeley's reading of Descartes can be avoided. Further, Descartes treats accommodation and convergence as means by which distance is *seen* and *experienced*.

On this reading, Descartes can be considered a direct realist, in at least one sense of that term. For objects near at hand, our perception of distance is cognitively unmediated. It is as "direct" as the experience of color, arising from a psychophysiological correspondence. This directness is, however, limited to objects near at hand (perhaps up to twenty feet). And indeed, even for perception of spatial properties near at hand (size, shape, distance), unnoticed judgments or ingrained habit may help produce phenomenal states of seeing (as in Descartes' third grade of sense). Accordingly, Descartes' psychophysiological account does not greatly affect his sensory epistemology. Whether one accepts the reading of natural geometry here given or not, Descartes held that the senses are reliable only in restricted circumstances and that they are not a means for perceiving the essences of things (AT 7: 83).

The philosophical significance of this reading pertains to Descartes' philosophy of mind and the greater role envisioned for psychophysiological relations in explaining sensory experience. This reading expands the role of embodiment in Descartes' philosophy. The mind does not inspect the pineal lean and then calculate a distance. Rather, the mind is affected by the body in such a way that it experiences some locally useful but perhaps only approximate facts concerning the distances of things. This reading coheres with an appreciation of a broader role for bodily mechanisms in Descartes' psychology, as propounded in the Sixth Meditation, the Fourth Replies, the *Passions of the Soul*, the *Treatise*,[31] and, more than has generally been appreciated, the *Dioptrics*. Descartes assigned a significant portion of the processes that preserve the body to the body alone. In this way, he conceived both nonhuman animals and human beings as possessed of bodies that are capable on their own (without mental intervention) of behaving for the preservation of the organism.[32] In ensouled beings, these bodily processes can produce sensory mental states that track properties of the environment (such as distance, in visual perception) without having to rely on cognitive operations, even of an unnoticed sort.

Finally, there is the larger picture of the history of theories of vision that has been offered by Crary. According to this picture, subjective phenomena such as afterimages, attention to the binocularity of vision, and embodied perception were distinctively nineteenth-century achievements in the literature of sensory physiology and psychology. One might easily show that subjective phenomena and the subjectivity of visual experience are seventeenth century themes, in deeper ways than Crary imagines. As to binocularity and embodiment, Descartes' theory is one of binocular embodiment.

---

[31] In the Sixth Meditation, Descartes discusses how the machine of the human body is organized so as to produce beneficial sensory states in the mind (CSM 2: 59–61, AT 7: 86–89). In the Fourth Replies, Descartes suggests that many organic processes as well as adaptive behaviors (such as extending the hands in a fall) occur without mental intervention (CSM 2: 161, AT 7: 229–30). The *Passions* and the *Treatise* together develop a significant "machine psychology," on which, see Hatfield (2007, 2012).

[32] In the case of nonhuman animals, I favor an interpretation of Descartes' works in which he completely denies minds, souls, and hence sentience to them (see Hatfield 2012), although of course I don't myself believe that animals are insentient.

Crary's book *Techniques of the Observer* interestingly highlights devices as exemplars of visual theory: the *camera obscura* pre 1810, the stereoscope from about 1830 on. But he seems to believe he can find the theory in the device, rather than needing to interpret the place of the device within the articulated theory. If Descartes held to a *camera obscura* model of vision, then the famous diagram from the *Dioptrics* (Fig. 1, above) would be a representation of how vision occurs rather than a portrayal of an anatomical demonstration, which is how Descartes describes it. Moreover, Descartes firmly rejected the implied homuncular theory that Crary foists upon him, saying: "we must not hold that it is by means of this resemblance that the picture causes us to perceive the objects, as if there were yet other eyes in our brain with which we could perceive it" (O 101*, AT 6: 130).

The preparation of the bovine eye is not a model of how vision works for Descartes. In his oeuvre, the diagrams representing how vision works include Figs. 2, 6, 8, and 9, which all involve binocularity and deep embodiment. As Descartes also said, "it is the mind which sees, not the eye, and it can see immediately only through the intervention of the brain" (O 108, AT 6: 141). The *camera obscura* is a model of the single eye (Fig. 1). For Descartes, vision occurs through the system of two eyes and its psychophysiological relation to the embodied mind (Figs. 9, 10). Rather than making the subject float free in relation to a cameral image, as Crary supposes, Descartes places the subject in the middle of the process, not as a viewer of flat images or a contemplator of geometrical constructions but as a subject who is directly confronted by a world in three dimensions.

The immediacy of distance and depth for Descartes might be compared to the difference between feeling an emotion such as fear, on the one hand, and merely judging that there is reason to be afraid or merely imagining fear, on the other. For Descartes, objects at a distance are *seen* and *felt* to be at a distance, not merely *imagined* to be there. We do not, in his view "really see" a two-dimensional image; what we "really see" includes depth and distance. And we see this depth and distance by brain mediated mechanisms that respond to geometrical structure in a mechanical manner, without need for mental computation or judgment.[33] In essence, Descartes has expanded the physiologically immediate objects of sight to include depth and distance. While not peculiar to sight, distance is as immediate for sight as it is for touch. Which also calls into question the assumed hegemony in the history of early modern visual theory of the Berkeleyan motto that touch educates vision. At the least, they are co-equal in Descartes. But with this difference: vision allows us to see things farther away, beyond the reach of touch. It truly is a marvelous sense.

Acknowledgments I offer this chapter in honor of David Lindberg. I thank audiences at the Max Planck Institute for the History of Science, the Visual Studies Lecture Series at Penn, Indiana University (where an earlier version was given as the Westfall Lecture), Princeton University, and

[33] The unnoticed judgments of the third grade of sense may also deliver a phenomenally immediate experience of distance and size, as in Descartes' discussion of size constancy (O 107, AT 6: 140).

the City University in New York for challenges to and discussion of the ideas presented. Mattia Mantovani offered helpful comments on an earlier draft, as did Greyson Abid and Alistair Isaac on the penultimate version. I acknowledge NSF grant 1028130 in support of Isaac's work.

References

Descartes' works are cited using these abbreviations:

"AT" for the revised Adam and Tannery edition of Descartes' *Oeuvres* (Descartes 1969–75).
"CSM" for the Cottingham et al. translation of Descartes' works (Descartes 1984–85).
"G" for Gaukroger's translation of the *Treatise on Man* (Descartes 1998).
"O" for Olscamp's translation of the *Dioptrics* (Descartes 1965).

I follow the cited translations when possible; emended passages are marked with an asterisk (*).

Al-Haytham, I. (1989). *The optics of Ibn al-Haytham: Books I-III, on direct vision* (Vol. 1) (A. I. Sabra, Trans.). London: Warburg Institute. Originally written *c.*1030.

Aquinas, T. (1889). *Summa theologiae*. S. E. Fretté & P. Maré (Eds.) Paris: Vivès.

Arbini, R. (1983). Did Descartes have a philosophical theory of sense perception? *Journal of the History of Philosophy, 21*, 317–337.

Atherton, M. (1990). *Berkeley's revolution in vision*. Ithaca: Cornell University Press.

Atherton, M. (1997). How to write the history of vision: Understanding the relationship between Berkeley and Descartes. In D. M. Kleinberg-Levin (Ed.), *Sites of vision: The discursive construction of sight in the history of philosophy* (pp. 139–166). Cambridge: MIT Press.

Berkeley, G. (1709). *An essay towards a new theory of vision*. Dublin: Rhamesand Papyat.

Bos, H. J. M. (2001). *Redefining geometrical exactness: Descartes' transformation of the early modern concept of construction*. New York: Springer.

Crary, J. (1990). *Techniques of the observer: On vision and modernity in the nineteenth century*. Cambridge: MIT Press.

Descartes, R. (1677). *L'Homme* (2nd ed.). Paris: Girard.

Descartes, R. (1692a). *Specimina Philosophiae: seu Dissertatio de Methodo, Dioptrice, et Meteora*. Frankfurt: Frederick Knoch (Originally published in French in 1637).

Descartes, R. (1692b). *Tractatus de homine*. Frankfurt: Frederick Knoch.

Descartes, R. (1897–1913). *Oeuvres de Descartes*. C. Adam & P. Tannery (Eds.), (11 Vols.). Paris: Cerf and Vrin.

Descartes, R. (1965). *Discourse on method, optics, geometry, and meteorology.* (P. J. Olscamp, Trans.) New York: Bobbs-Merrill.

Descartes, R. (1969–1975). *Oeuvres de descartes*. C. Adam & P. Tannery (Eds.), (11 Vols.). Paris: Vrin.

Descartes, R. (1984–1985), *Philosophical writings of descartes*. In J. Cottingham, R. Stoothoff & D. Murdoch (Eds.), (2 Vols.). Cambridge: Cambridge University Press.

Descartes, R. (1998). *Treatise on man*, in *The world and other writings* (S. Gaukroger, Trans.), (pp. 99–169). Cambridge: Cambridge University Press.

Eustace of Saint Paul. (1638). *Summa Philosophiae Quadripartita: De Rebus Dialecticis, Ethicis, Physicis, & Metaphysicis* (Cologne: Philip Albert), pt. 3, "Physica.".

Gogel, W. C. (1990). A theory of phenomenal geometry and its applications. *Perception and Psychophysics, 48*, 105–123.

Gregory, R. L. (1997). *Eye and brain: The psychology of seeing* (5th ed.). Princeton: Princeton University Press.

Hatfield, G. (1986). The senses and the fleshless eye: The *Meditations* as cognitive exercises. In A. Rorty (Ed.), *Essays on Descartes' Meditations* (pp. 45–79). Berkeley: University of California Press.

Hatfield, G. (1990). *The natural and the normative: Theories of spatial perception from Kant to Helmholtz*. Cambridge: Mass: MIT Press.

Hatfield, G. (1992). Descartes's physiology and its relation to his psychology. In J. Cottingham (Ed.), *Cambridge companion to Descartes* (pp. 335–370). Cambridge: Cambridge University Press.

Hatfield, G. (2002). Perception as unconscious inference. In D. Heyer & R. Mausfeld (Eds.), *Perception and the physical world: Psychological and philosophical issues in perception* (pp. 115–143). New York: Wiley.

Hatfield, G. (2007). The passions of the soul and Descartes' machine psychology. *Studies in History and Philosophy of Science, 38*, 1–35.

Hatfield, G. (2012). Mechanizing the sensitive soul. In G. Manning (Ed.), *Matter and form in early modern science and philosophy* (pp. 151–186). Leiden: Brill.

Hatfield, G., & Epstein, W. (1979). The sensory core and the medieval foundations of early modern perceptual theory. *Isis, 70*, 363–384.

Kepler, J. (1611). *Dioptrice*. Augsburg: David Franke (Reprint, Cambridge: Heffer, 1962).

Kepler, J. (2000). *Optics: Paralipomena to witelo and optical part of astronomy* (W. H. Donahue, Trans.). Sante Fe: Green Lion Press (Original Latin edition published in 1604).

Lindberg, D. C. (1976). *Theories of vision from al-Kindi to Kepler*. Chicago: University of Chicago Press.

Massey, L. (2007). *Picturing space, displacing bodies: Anamorphosis in early modern theories of perspective*. University Park: University of Pennsylvania Press.

Maull, N. (1980). Cartesian optics and the geometrization of nature. In S. Gaukroger (Ed.), *Descartes: Philosophy mathematics and physics* (pp. 23–40). New Jersey: Barnes and Noble Books.

Michael, E., & Michael, F. S. (1989). Corporeal ideas in seventeenth-century psychology. *Journal of the History of Ideas, 50*, 31–48.

Panofsky, E. (1991). *Perspective as symbolic form*. (C. S. Wood, Trans.) New York: Zone Books.

Ptolemy, C. (1996). *Ptolemy's theory of visual perception: An english translation of the optics*, A. M. Smith (Ed.) Philadelphia: American Philosophical Society. Originally written ca. 160.

Rey, A., ed. (1992), *Dictionnaire historique de la langue française: contenant les mots français en usage et quelques autres délaissés* (2 Vols.). Paris: Dictionnaires Le Robert.

Sabra, A. I. (1978). Sensation and inference in Alhazen's theory of visual perception. In P. Machamer & R. Turnbull (Eds.), *Studies in perception* (pp. 160–185). Columbus: Ohio State University Press.

Sabra, A. I. (1987). Psychology versus mathematics: Ptolemy and Alhazen on the moon illusion. In E. Grant & J. Murdoch (Eds.), *Mathematics and its application to science and natural philosophy in the middle ages* (pp. 217–247). Cambridge: Cambridge University Press.

Schuster, J. (1980). Descartes' Mathesis Universalis, 1619–28. In S. Gaukroger (Ed.), *Descartes: Philosophy, mathematics and physics* (pp. 41–96). New Jersey: Barnes and Noble Books.

Shea, W. R. (1991). *The magic of numbers and motion: The scientific career of René Descartes*. Nantucket: Mass: Science History Publications.

Turbayne, C. M. (1970). *The myth of metaphor* (rev ed.). Columbia: University of South Carolina Press.

Vinci, T. C. (1998). *Cartesian truth*. New York: Oxford University Press.

Wilson, M. D. (1993). Descartes on the perception of primary qualities. In S. Voss (Ed.), *Essays on the science and philosophy of René Descartes* (pp. 162–176). New York: Oxford University Press.

Wolf-Devine, C. (1993). *Descartes on seeing: Epistemology and visual perception*. Carbondale: Southern Illinois University Press.

Wolf-Devine, C. (2000). Descartes' theory of visual spatial perception. In S. Gaukroger, J. Schuster, & J. Sutton (Eds.), *Descartes' natural philosophy* (pp. 506–523). London: Routledge.

Hobbes's Theory of Space

Douglas Jesseph

Philosophy in the seventeenth century is often, and with reason, characterized as a collection of grand systems devoted to an all-encompassing account of the world and its workings. Yet even in an era that featured such systematic thinkers as Descartes, Spinoza, and Leibniz, Hobbes is noteworthy for his "big picture" approach to philosophy. His tripartite *Elements of Philosophy* features the treatises *De Corpore*, *De Homine*, and *De Cive*, which were presented to the public as an exposition of all the philosophy worth knowing. Indeed, one might suspect that by entitling his system *Elements of Philosophy*, Hobbes was attempting to do for philosophy what Euclid's *Elements* had done for geometry. The structure of the three works in Hobbes's *Elements* reflects his conception of the structure of knowledge: beginning with a treatise on the nature of body, Hobbes next proceeded to examine the nature of humans (i.e., animated, rational bodies), and thence to a discourse on the nature of the commonwealth (the artificial body bound together by human covenants). In consequence, Hobbes's *De Corpore*—a dissertation on the nature of body—occupies the foundational place in his system, and its foundational status is due to the fact that Hobbes held to a strict materialism in which only body is real, so that all else must be accounted for in terms of the action of bodies.[1]

A system founded on the nature of material bodies must also require that the concept of space be given central importance. My purpose here is to examine Hobbes's account of the nature of space, with the intent of showing how his theory

D. Jesseph (✉)
Department of Philosophy, University of South Florida,
4202 E. Fowler Avenue, Tampa, FL 33620, USA
e-mail: djesseph@usf.edu

[1] As it happens, these three parts of the Hobbesian system were published out of order: *De Cive* first appeared in 1642, *De Corpore* in 1655, and *De Homine* in 1658.

V. De Risi (ed.), *Mathematizing Space*, Trends in the History of Science,
DOI 10.1007/978-3-319-12102-4_8

of space fits together with his fundamental ontology, his philosophy of science, and his idiosyncratic approach to the foundations of geometry. Toward that end, I will start with an overview of Hobbes's conception of space and time; I then take up some traditional philosophical questions concerning the nature of space and contrast Hobbes's approach with others, notably Descartes and Thomas White. I will conclude by briefly considering how Hobbes's theory of space influences his account of geometry.

1 Hobbes on Space and Time

Hobbes defined philosophy as "such Knowledge of Effects, or Appearances, as we acquire by true Ratiocination from the knowledge we have first of their Causes or Generation; And again, of such Causes or Generations as may be from knowledge first of their Effects" (*DCo* 1.1.1). Reasoning from effects to their (possible) causes is what he termed "analysis," while that which moves from causes to effects is "synthesis;" both are essential to philosophy, and they both require an understanding of the concept of causation.[2] Further, Hobbes held that motion was the only cause of all things (*DCo* 1.6.5), so that the concept of motion is implicated whole of Hobbes's philosophy. In fact, there is no obvious demarcation between first philosophy and natural philosophy in Hobbes's system because the fundamental notions in his first philosophy include seemingly physical concepts such as body and motion.

Hobbes held that the foundations of natural philosophy are to be found by starting with the thought experiment (as we would now term it) of supposing the annihilation of the entire world, with the exception of one individual.[3] According to Hobbes, the solitary thinker who survived the otherwise complete annihilation of the world would retain "Ideas of the World, and of all such Bodies as he had, before their annihilation, seen with his eies, or perceived by any other Sense; that is to say, the memory and Imagination of Magnitudes, Motions, Sounds, Colours & c. as also of their order & parts" (*DCo* 2.7.1). On the basis of this collection of "phantasms" retained from past experience, such a thinker would be supplied with the concept of space, because any recalled phantasm would still contain information about the "order and parts" of the phenomenal world. Recollecting such phantasms and taking these merely to be representations of objects external to his perceiving mind, our thinker will "have presently a Conception of that we call *Space*, an Imaginary

[2] The literature on Hobbes and the concepts of analysis/synthesis is vast. For a general overview see Engfer (1982) and Hintikka and Remes (1974). Watkins (1965, Chaps. 3 and 4) places Hobbes's account in the tradition of the Italian "School of Padua," while Prins (1990) argues against such a connection. See also Hanson (1990) and Talaska (1988) for further accounts of Hobbes on analysis and synthesis.

[3] On the "annihilation experiment" in Hobbes and its background in earlier philosophical traditions, see Leijenhorst (2002, pp. 109–111).

Space indeed, because a meere Phantasme, yet that very thing which all men call so" (*DCo* 2.7.2). Hobbes then took this thought experiment to show that that space is to be properly defined as "*the Phantasme of a Thing existing without the mind simply*, that is to say, that Phantasme, in which we consider no other Accident, but onely that it appears without us" (*DCo* 2.7.2).

This definition has been the source of some considerable confusion and debate among Hobbes's interpreters. It appears to take space to be a "phantasm," or a purely subjective appearance, rather than an objective, observer-independent framework in which objects and events are located. Seth Ward, one of the earliest critics of Hobbes's program for natural philosophy, complained that the definition was "obscure and false" and "seems to be something only an insane person would say" (Ward 1656, pp. 69, 73) because it confuses a subjective experience or phantasm with the objective location of bodies that constitutes space. Some later commentators have seen in Hobbes's definition of space an anticipation of Kant's theory of space as a form of external intuition, derived from a "phenomenological reduction" of experience.[4] However, Hobbes's reasons for adopting his account of space show that his doctrine lacks a significant connection to Kant's. Hobbes emphatically did not take space as some kind of transcendental principle that conditions our perceptions of external objects, but rather as the effect of bodies that interact with our perceptual system. Further, Hobbes took the concept of space to be an empirical in origin instead of an a priori form of perception that is presupposed in any experience. Finally, it is worth noting that Hobbes defined the key term "phantasm" in terms of the motion of material bodies: a phantasm is simply motion communicated from external bodies through the human sensory apparatus, where it elicits a counter-motion.[5] In other words, rather than taking space as a basic concept in terms of which to define motion, Hobbes took motion as basic and defined space in terms of it.[6]

The definition we have been considering identifies what Hobbes termed "Imaginary Space," and if this were all he had to say on the matter he would in fact have been some kind of idealist for whom space is a subjective or mind-dependent

---

[4] Herbert (1987, p. 701) declared that "Hobbes's account of space anticipates in many respects the Kantian phenomenology of space that will appear more than a century later," although the basis for this characterization is highly problematic. See Leijenhorst (2002, p. 107) for the difficulties involved in this interpretation.

[5] Hobbes did not actually define the key term 'phantasm' until the fourth part of *De Corpore*. He first defined 'sense' as '*a Phantasme, made by the reaction and endeavor outwards in the Organ of Sense, caused by an Endeavour inwards from the Object, remaining for some time more or less*' (*DCo* 4.25.2); he then declared a phantasm to be an instantaneous instance of sense: "For a *Phantasme* is the act of Sense, and differs no otherwise from Sense then *fieri* (that is, Being a doing) differs from *Factum esse*, (that is, Being done;) which difference, in things that are done in an Instant, is none at all; and a Phantasme is made in an Instant" (*DCo* 4.25.3).

[6] Brandt (Brandt 1928, p. 253) finds this problematic because "everything is to be motion, hence the idea of magnitude is also 'in reality' motion—but motion cannot be conceived without space and magnitude.".

ordering that individuals assign to items of their experience.[7] However, the Hobbesian doctrine of space does not end with the definition of imaginary space. A key point to consider here is that the main reason for calling something *imaginary* is to contrast an imaginary thing with its *real* or objective counterpart. In Hobbes's treatment of space, this contrast results in the concept of *real space*, which his an objective system of orderings that locates objects in a way independent of the perceptions of any particular mind.

Hobbes announced that "The Extension of a Body, is the same thing with the MAGNITUDE of it, or that which some call *Real Space*. But this *Magnitude* does not depend upon our Cogitation, as Imaginary Space doth; for this is an Effect of our Imagination, but *Magnitude* is the Cause of it; this is an Accident of Mind, that of a Body existing out of the Mind" (*DCo* 2.8.4). This passage clearly intends to distinguish between imaginary and real space: imaginary space is the subjective ordering of objects that arises from external bodies' interaction with the perceiver's sensory system; in contrast, real space is the objective location of external bodies. Indeed, Hobbes's definition of the term 'body' as "*that which having no dependence upon our Thought is coincident or coextended with some part of Space*" (*DCo* 2.8.1) requires the objective ordering of things in real (i.e., not mind-dependent) space.

It is worth noting that when Hobbes defined imaginary space as a "phantasm," the term *phantasm* carries a good deal of spatial content with it. A phantasm, is "made by the Reaction and endeavor outwards in the Organ of Sense, caused by an Endeavour inwards from the Object" (*DCo* 4.25.2), so the very concept of a phantasm presupposes such spatial orientations such as 'inward' and 'outward'. I take this as a tolerably clear indication that imaginary space presupposes the existence of a real space in which objects and events are ordered.

The easiest way to illustrate the distinction between imaginary and real space is to consider the case of spatial orderings. In observing two glasses on a table, a perceiver might see the wine glass as located to the left of the water glass. But such an ordering is clearly subjective because it depends upon the relative positions of the observer and the two glasses. Thus, it makes sense to say that the wine glass is to the left of the water glass in the imaginary space determined by the phantasms of our perceiver. But this ordering does not hold in an absolute or objective sense because there are plenty of reference frames in which perceivers would not see such an ordering. There is, nevertheless, an objective fact of the matter about the relative positions of the two glasses, and this position in real space (together with the spatial location of the perceiver) is what causes the ordering in imaginary space.

Hobbes drew a more informative and explicit distinction between imaginary and real space in his long manuscript treatise critiquing Thomas White's 1643 work *De Mundo Dialogi Tres*. White's dialogues were a response to Galileo's case for Copernicanism as set forth in the *Dialogue on the Two Chief World Systems*, and

[7] See Leijenhorst (2002, pp. 105–122) for a discussion of Hobbes's treatment of imaginary space and its background in medieval discussions of the ontology of space. Grant (1969) is a useful summary of medieval accounts of imaginary space.

Hobbes undertook the defense of Galileo. In his critique of White, Hobbes defined imaginary space as "the image or phantasm of a body," concluding that "the existence of this [space] does not depend on the existence of body but on the existence of the imaginative faculty."[8] However, "it is impossible that we suppose there to be some body without at the same time thinking it to be endowed with its own dimensions or spaces. Therefore this space, which can be called real, is inherent in a body, as an accident in its subject, and would surely exist even if there were nothing that could imagine it."[9]

Imaginary space is therefore not an independent construct of the mind but rather an effect of the action of external bodies on the senses. To better understand this point, we should recall that Hobbes's materialistic ontology identifies body as the only substance. Because space cannot be identified with any particular body (or collection of bodies), it must be an "accident" of body. Hobbes defined an accident of body as "*the Manner by which any body is conceived*," or also as "*that faculty of any Body by which it works in us a Conception of it self*" (*DCo* 2.8.2). There are, in fact, two definitions of the term 'accident' at work here. The first proceeds "subjectively" by defining an accident in terms of how it is conceived; the second proceeds "objectively" by defining an accident as the means whereby a body generates a phantasm in a perceiving mind. This distinction allows Hobbes to define his two notions of the accident that is space: imaginary space is the subjective (or, in Hobbes's terms, "fictive") location of a body as it appears to a perceiver. Hobbes held, however, that this "feigned Extension" of a body in imaginary space is distinct from its *magnitude* or what he called its "true Extension" (*DCo* 2.8.5). The Hobbesian concept of real space thus characterizes it as the magnitude of a body, while imaginary space is the "phantasm" produced by external bodies acting upon the senses; or, as one commentator has put it, "the relation between *spatium reale* to *spatium imaginarium* exemplifies the cause-effect relation existing between our ideas and external things" (Leijenhorst 2002, p. 107).

Hobbes's account of time has important similarities to his theory of space, with one very important difference. Time is defined in *De Corpore* as "*the Phantasme of Before and After in Motion*" (*DCo* 2.7.3); this account is clearly analogous to his definition of imaginary space as "*the Phantasme of a Thing existing without the Mind simply*" (*DCo* 2.7.2). Time is therefore defined as a mental representation of motion rather than a self-subsistent entity that exists independently of any perceiving minds or moving bodies. However, where he had distinguished between imaginary and real space, Hobbes did not suppose a "real time" existing independently of the activity of minds that keep track of passage. He reasoned:

---

[8] "dicemus *spatium esse imaginem corporis, quatenùs corporis* . . . Manifestum hinc est existentiam spatii dependere non ab existentia corporis sed ab existentia imaginativae facultatis" (*CDM*, 3.1, p. 117).

[9] "Neque possible est ut corpus aliquot desse existimemus, quin simul putemus ipsum praeditum esse dimensionibus, sive spatiis suis. Hoc spatium igitur quod appelari potest reale inhaerens corpori, ut accidens in subiecto suo, existeret sane, et si nihil esset quod ipsum imaginari possit," (*CDM*, 3.2, p. 117).

> For seeing all men confess a Yeare to be Time, and yet do not think a Year to be the Accident of Affection of any Body, they must needs confesse it to be, not in the things without Us, but only in the Thought of the Mind . . . . What then can Dayes, Monthes and Years be, but the Names of such Computations made in our Mind? *Time* therefore is a Phantasme, but a Phantasme of Motion, for if we would know by what Moments Time passes away, we make use of some Motion or other, as of the Sun, of a Clock, or the sand in an Hourglasse. (*DCo* 2.7.3)

Notwithstanding his steadfast opposition to the great majority of Aristotelian doctrines, Hobbes took this theory of time to be consistent with Aristotle's definition. The Aristotelian definition of time (*Physics* IV, 11 220^{a} 24–25) characterizes it as "the number of movement in respect of the before and after." Hobbes said of this definition that because "that Numbering is an act of the mind; and therefore it is all one to say, *Time is the Number of Motion according to Former and Later*; and *Time is a Phantasme of Motion Numbered*" (*DCo* 2.7.3).

In Hobbes's account of space and time, it turns out that time is an accident of motion, where space is an accident of extension or magnitude. The accident that is time is mind-dependent, however, because it requires a perceiver to compute or reckon its passage. In the contrasting case of extension, the accident that is space has an observer-dependent aspect (namely, imaginary space) and an objective aspect (real space). Hobbes declared that "MOTION and MAGNITUDE … are the two most common Accidents of Bodies," (*DCo* 3.15.1) and he also announced that "The Extension of a Body is the same thing with the MAGNITUDE of it," (*DCo* 2.8.4). It follows that the two most fundamental properties of body are motion and extension. Space and time are consequences of these two basic accidents. Space derives from the accident of extension, with real space identified with a body's location, while imaginary space is the phantasm of its position in some perceiver's mind. Time, in contrast, is identified with the measure of a body's motion; and because all measurement presupposes a measuring mind, time is inevitably an *ens rationis*.

There are some oddities in this metaphysics of space and time that are worth considering. At the level of fundamental ontology, Hobbes took material bodies as absolutely basic. As we have seen, every body possesses two primary and inalienable accidents, namely magnitude and motion. From these, Hobbes then proceeded to define the concepts of space, place, and time. However, this all seems rather backward. One would normally expect space and time to be taken as primitives, with body and motion defined in the obvious way: body is that which occupies a region of space, and motion is translation of a body through a spatial distance in a given time. Hobbes did not take this route, evidently because he regarded the concepts of body and motion as so perspicuous as to be explanatorily prior to those of space and time. A second oddity in Hobbes's approach is that he seems to have fallen into the trap of trying to define everything: after he had defined space and time in terms of body and motion, he then defined body in terms of place or space (*DCo* 2.8.1) and motion in terms of space and time (*DCo* 2.8.10). The result of this typically Hobbesian definition-mania is a first philosophy that contains a set of seemingly circular definitions.

2 The Limits of Hobbes's Doctrine of Space and Time

Although Hobbes was not shy about claiming demonstrative certainty for a wide variety of physical, political, and philosophical principles, there are a number of fundamental questions concerning the nature of space (and, to a lesser extent, time) that he took to be either unanswerable, or (if answerable at all) only to be resolved empirically. Hobbes held that two a priori demonstrable principles governed all bodies in motion: first that a body could never initiate or extinguish its motion, and second that the only way the state a body could be altered is through contact with another body.[10] These comprise a very general framework within which natural-philosophical investigations can be undertaken.

Nevertheless Hobbes's epistemology also placed a number of traditional questions about the infinity and uniqueness of the world outside the scope of philosophical investigation. Hobbes held that the origin of all concepts "is that which we call SENSE; (For there is no conception in a mans mind, which hath not at first, totally, or by parts, been begotten upon the organs of Sense.) The rest are derived from that original" (*L* 1.1, p. 3). He further insisted that:

> Whatsoever we imagine, is *Finite*. Therefore there is no Idea, or conception of any thing we call *Infinite*. No man can have in his mind and Image of infinite magnitude; nor conceive infinite swiftness, infinite time, or infinite force, or infinite power. When we say any thing is infinite, we signifie onely, that we are not able to conceive the ends, and bounds of the thing named; having no Conception of the thing, but of our own inability. (*L* 1.3, p. 12)

A consequence of these epistemological restrictions is the outright rejection of traditional philosophical investigations into the question whether the world is finite or infinite, or whether there might be multiple worlds.

Natural philosophers had long been concerned to determine whether the world is finite or infinite, or whether there could be worlds other than ours, and discussions of these topics are a feature of much ancient and medieval philosophy.[11] Hobbes, however, regarded all such speculation as pointless. Considering disputes over the finitude or infinitude of the world, he insisted that "when we make question whether the World be Finite or Infinite, we have nothing in our Minde answering to the name *World*; for whatsoever we Imagine, is therefore Finite" (*DCo* 2.7.12). Although any given imagined space must be finite, we can nevertheless always imagine it to be further extended. Consequently, there can be no determinate limit to the extent of imaginary space. On Hobbes's principles, to say that the world is infinite is merely to say that we cannot conceive it as bounded; but this does not

---

[10] I term these two principles the "persistence principle" and the "action by contact" principle. Both are demonstrated (at least to Hobbes's satisfaction) in Part II, of *De Corpore*, Sections 8 and 9. The first is a version of the law of inertia that asserts "*Whatsoever is a Rest, will always be at Rest, unless there be some other Body besides it, which by endeavouring to get into its Place by motion, suffers it no longer to remain at Rest*." (*DCo* 2.8.19). The second claims "There can be no Cause of Motion, except in a Body Contiguous, and Moved." (*DCo* 2.9.7). I discuss them at greater length in (Jesseph 2006).

[11] SeeDuhem (1985) for a further account of such speculations.

exclude the possibility that the universe is of finite extent yet too vast to be grasped by our limited cognitive capabilities. However, we could never have sufficient evidence to justify the claim that imaginary space is literally infinite, because there can be no “phantasm” of the infinite. Likewise, we cannot exclude the possibility that there are other worlds besides this one, although we can have no empirical evidence of the existence of such worlds. Questions of this sort are therefore unanswerable, at least philosophically.

Hobbes’s principles thus lead to a more modest natural-philosophical program than many of his contemporaries or predecessors, at least where questions about the finitude or uniqueness of the world are concerned. In commenting upon these issues, Hobbes remarked

> And this is of it selfe so manifest, that I should not thinke it needed any explaining at all, but that I finde Space to be falsely defined by certaine Philosophers, who inferred from thence, One, that the world is Infinite; for taking *Space* to be the Extension of Bodies and thinking Extension may increase continually, he inferres that Bodies may be infinitely Extended; and Another from the same Definition concludes rashly, that it is impossible even to God himself to create more Worlds then one; for if another World were to be created, he says, that seeing there is nothing without this world, and therefore (according to his Definition) no Space, that new world must be placed in nothing, but in nothing nothing can be placed, which he affirms onely, without shewing any reason for the same; whereas the contrary is the truth: for more cannot be put into a Place already filled, so much is Empty Space fitter then that which is Full for the receiving of new Bodies. (*DCo* 2.7.2)

This passage is directed at the reasoning of Descartes and Thomas White. Descartes’ error (in the *Principles of Philosophy*, Part II, article 21) was to conclude that the universe must be infinite in extent because we can imagine any supposed limit to the world to be surpassed. As he put the matter “Wherever we imagine those limits [of the world] to be, we can always not only imagine some indefinitely extended spaces beyond them, but also perceive that these are truly imaginable, that is to say real; and thus indefinitely extended corporeal substance is also contained in them” (Descartes 1964–1976, 8: 52). White’s error (in the third part of the first dialogue of his *De Mundo Dialogi Tres*) was to assume the principle that “what is situated in nothingness has no place,” from which he concluded that the world as a whole could not be placed, since there is no location apart from the universe in which it could be placed.[12]

In contrast to these mistaken arguments, Hobbes held that the concept of space (whether imaginary space or real space) can yield no answer to these questions. All our knowledge of the world is derived from the “phantasms” of sense experience, but none of these is infinite, nor can we have experience of an alternate world beyond the spatio-temporal bounds of this one. Therefore, no sense experience can decide the issue of whether the universe is finite or infinite, or whether there are other worlds besides this one. Furthermore, our spatial concepts cannot resolve such

[12] White remarks “quid clarius esse potest quam positam in nihilo rem locum nullum habere?” (White 1643, p. 28).

questions a priori; consequently, questions about the infinity and uniqueness of the world can be dismissed as seventeenth-century pseudo-problems.

Hobbes returned to these themes in Part IV of *De Corpore,* again emphasizing that questions about the magnitude, duration, and uniqueness of the world must be "inscrutable" because no experience can decide the issue. As he framed the issue: "Whatsoever we know that are Men, we learn it from our Phantasmes, and of *Infinite* (whether Magnitude or Time) there is no Phantasme at all; so that it is impossible either for a man, or any other creature to have any conception of *Infinite*." Moreover, "whether we suppose the World to be Finite, or infinite, no absurdity will Follow. For the same things which now appear, might appear, whether the Creator had pleased it should be Finite or Infinite" (*DCo* 4.26.1). Neither experience nor a priori reasoning from the concepts can settle the issue, therefore such questions "are not to be determined by Philosophers, but by those that are lawfully authorized to order the Worship of God" (*DCo* 4.26.1).

Hobbes's dismissal of questions about the infinity or uniqueness of the world contrasts with his handling of debates over the existence of a vacuum. Aristotle had offered several a priori arguments against the possibility of a vacuum in the *Physics* (I 7), and the question of the void was a standard topic in medieval treatments of natural philosophy.[13] The issue remained the subject of philosophical dispute well into the seventeenth century, with Descartes declaring that it "is a contradiction to suppose there is such a thing as a vacuum, i.e. that in which there is nothing whatever," in the second part of his *Principles of Philosophy*. In elucidating this doctrine Descartes argued that, if we suppose a vessel to contain a literal vacuum, then "when there is nothing between two bodies, they must necessarily touch one another," so that the sides of the vessel containing the supposed vacuum must collapse upon each another (Descartes 1964–1976, 8: 48, 50). Thomas White accepted the cogency of this argument, offering that "If a there were a vacuum, there would be a place without body, that is, a concave body without anything to fill the cavity. So the sides of this concave body will close up because there is no *ens* between them. But if the sides are closed up, this shuts out the vacuum."[14]

Hobbes regarded such a priori arguments against the vacuum as comical misapplications of philosophical method. Both Descartes and White mistake the issue by assuming that a void must be a spatial region containing literally nothing, rather than nothing *other than space itself*. He remarked:

> And this is so easie to be understood, that I should wonder at some men, who being otherwise skillful enough in Philosophy, are of a different opinion, but that I finde that most of those that affect Metaphysical subtilties, wander from Truth, as if they were led out of the way by an *Ignis Fatuus*. For can any man that has his natural Senses, think that two Bodies must therefore necessarily Touch one another, because no other Body is between them? Or

[13] See Duhem (1985, part IV) for an account of medieval disputes regarding the vacuum.

[14] "Si enim vacuum est locus sine corpore, hoc est, corpus concauum, sine aliquot cauitatem implead, none vides concauitatem quanta esse sine tertiâ mediante. Vides ergo ex ipsa notione vacui conjuncta esse latera, etsi conjuncta sint, iam nullum reliquum esse vacuum" (White 1643, pp. 30–1).

> that there can be no *Vacuum*, because *Vacuum* is nothing, or as they call it, *Non Ens*? Which is as childish, as if one should reason thus; No man can Fast, because to Fast is to eat Nothing; but Nothing cannot be eaten. (*DCo* 2.7.9)

In Hobbes's view, the question of whether there is a vacuum is an empirical one, to be decided on the basis of experiment rather than a priori reasoning. Unlike the questions of the finitude or uniqueness of the world (where he held that experience cannot settle the issue) Hobbes argued that the experimental evidence convincingly refuted the hypothesis of a vacuum.

Hobbes proposed to "instance in onely one experiment, a common one, but (I think) unanswerable" to refute the hypothesis of the vacuum (*DCo* 4.26.2). This experiment involves a vessel "such as Gardiners use to water their Gardens withal," having small holes in the bottom and a larger opening at the top. When the vessel is filled with water and the top stopped shut, water does not flow through the holes; but when the top is opened, water does flow. Hobbes concluded that "the Water cannot by its natural endeavor to descend, drive down the air below," because there is no vacuum beneath it. The experiment therefore offers convincing evidence "that all Space is full; for without this, the natural motion of the water . . . downwards, would not be hindered" (*DCo* 4.26.21).

Those who accepted the reality of the vacuum based their opinion on what Hobbes termed "many specious arguments and experiments" that he took it upon himself to refute. He argued that all such evidence for the existence of a vacuum is either inconclusive or inconsistent with other a priori principles governing the motion of bodies. In order to clarify Hobbes's conception of space and his scientific methodology, it is worthwhile to consider two of these: first an a priori argument drawn from Lucretius concerning the necessity of vacuum to permit motion, and second an argument drawn from experiment of the "Torricellian tube."

The Lucretian argument that a vacuum is necessary for there to be motion reasons that "the office and property of Bodies is to withstand and hinder motion," so that in a world with no vacuum "motion would everywhere be hindered, so, as to have no beginning anywhere, & consequently there would be no motion at all" (*DCo* 4.26.3). Hobbes granted that "in whatsoever is full, and at rest in all its parts, it is not possible motion should have a beginning," (*DCo* 4.26.3), but this is because no body (or collection of bodies) can initiate its own motion. This principle Hobbes took to be demonstrably certain (and he credited himself with having demonstrated it in *DCo* 2.9.7). But, he argued, it is irrelevant to the case of the vacuum. At most, the Lucretian argument could show that "motion was either coeternal, or is of the same duration with that which is moved" (*DCo* 4.26.3), but it cannot rule out the possibility that the world is a plenum, albeit one containing many fluid bodies that offer no resistance to larger moving bodies. Thus, the argument yields no reason a priori in favor of the vacuum and against plenism.

The familiar Torricellian experiment involves inverting a tube filled with mercury and sealed at one end, placing the open end in a dish likewise filled with mercury and then inverting it. The mercury in the inverted tube descends, and the space at the top of the tube vacated by the mercury is characterized as a vacuum (at least by proponents of vacuuism). Hobbes claimed to "finde no necessity at all of a *Vacuum*" (*DCo* 4.26.4) in the experiment because nothing rules out the possibility that the whole apparatus may leak air. Hobbes imagined that the pressure of the mercury descending from the tube will act upon the ambient air at the surface of the vessel, so that "if the force with which the Quicksilver descends be great enough (which is greater or less, as it descends from a place of greater of less height) it will make the Aire penetrate the Quicksilver in the vessel, and go up into the Cylinder to fill the place which they thought was left empty" (*DCo* 4.26.4). Thus, at least in Hobbes's assessment, the Torricellian experiment gives no solid empirical reason in favor of the vacuum.

The reasoning Hobbes offered on the question of the vacuum is hardly free from difficulty, and I will not attempt to defend it. The reply to the Lucretian argument seems oddly question-begging, as Hobbes simply presupposes that motion is possible in a plenum, contrary to what the argument is intended to establish. An obvious problem with his reply to the Torricellian experiment is that very similar reasoning might undermine his favorite experimental evidence for plenism. If the hypothesis that fine particles of air might penetrate the column of mercury is legitimate, then surely a defender of the vacuum might make a similar appeal to unobservable pressures and suction forces to evade the plenist conclusions Hobbes derived from the experiment of the watering vessel. At best, then, Hobbes's reasoning does not make a decisive case for plenism. Still, it is important to recognize that Hobbes's procedure in reasoning about the vacuum is very much at odds with the kind of dogmatic, a priori adherence to plenism that some commentators have attributed to him.[15]

We have seen that Hobbes regarded some questions about the nature of space as unanswerable (namely those about the finitude or uniqueness of the world). He also held that some disputes about space (notably that over the vacuum) could be resolved only by empirical investigation. But he also thought that some doctrines concerning the nature of space and body were resolvable by a priori arguments. In particular, he held that the doctrine of rarefaction and condensation was an incoherent exercise in self-contradiction. The twin processes of rarefaction and condensation were invoked by Aristotelian natural philosophers who held that numerically the same body might gain or lose quantity, as when water rarefies to become vapor, or vapor condenses to become water.[16] In Hobbes's estimation this

[15] Shapin and Schaffer characterize Hobbes as a dogmatic adherent of plenism who had no interest in experiments. They assert that "What Hobbes was claiming . . . was that the systematic doing of experiments was not to be equated with philosophy: going on in the way Boyle recommended for experimentalists was not the same thing as philosophical practice . . . . This experimental way and the philosophical way were fundamentally different: they differed in their capacity to secure assent among intellectuals and peace in the polity" (1985, p. 129).

[16] The details of the doctrine are obscure enough to be left aside here. An oddity of the doctrine is that a rare body was taken (at least by some) to have *more* quantity in it than a dense one, contrary to our ordinary understanding of density and rarity. This derives from the fact that quantity is, in the words of Kenelm Digby, "nothing else but divisibility; and … a thing is bigge, by having a capacity to be divided" (1644, p. 9). Thus, a highly divisible body (such as a liter of water)

doctrine can be rejected on a priori grounds because it violates principles that follow from the concepts of space, body, and motion. According to Hobbes, it is a conceptual truth that “a Body keeps alwayes the same *Magnitude*, both when it is at Rest, and when it is Moved,” (*DCo* 2.8.5). Moreover, whenever a body appears to change magnitude (as when a plant grows, or water evaporates), it is a priori certain that such a change can be due only be due to the motion and impact of material bodies that interact with the changed body.

The consequence of such Hobbesian principles is that talk of rarefaction and condensation is literally incoherent. In *Leviathan*, Hobbes complained that the doctrine supposes “there could be Matter, that had not some determined Quantity; when Quantity is nothing else but the Determination of Matter,” or that there could be a body “made without any Quantity at all, and that afterwards more, or lesse were put into it, according as it is intended the body should be more or less Dense” (*L* 4.46, 375). In either case, the result is that there can be no content assigned to the concepts, or (as Hobbes put it in his *Six Lessons*), “nature abhorres even empty words, such as are . . . *Rarefying* and *Condensing*” (*SL* 2; *EW* 7: 225).

It should by now be clear that Hobbes’s conception of space is far removed from the dogmatic, anti-experimental, “rationalistic” enterprise that some commentators have found in his philosophy. As we have seen, his rejection of the vacuum was not based on ignorance of or indifference to experimental evidence, nor did it stem from an a priori argument against the possibility of a vacuum in the style of Descartes. The existence of a vacuum is not an issue settled by the foundational notions in Hobbes’s natural philosophy, i.e. his account of space, time, body, motion, and causation. Thus, if the issue can be resolved at all, it must be resolved on the basis of experiment, and Hobbes held that the experimental evidence was solidly against the hypothesis of a vacuum. Similarly, such phenomena as gravity, the freezing of water, or the propagation of sound cannot be accounted for by deducing them from the first principles of natural philosophy. Hobbes insisted that “where there is place for Demonstration, if the first Principles, that is to say the Definitions, do not contain the Generation of the Subject, there can be nothing demonstrated as it ought to be” (*SL* Epistle; *EW* 7: 184). However, when we lack access to the causes of phenomena (and this is typically the case in the investigation of nature), we must content ourselves with hypothetical causes. As Hobbes framed the issue, the fact that we must speculate hypothetically about the inner workings of nature means that “in natural causes, all you are to expect is but probability” (*SPP*, ch. 1; *EW* 7: 11).

(Footnote 16 continued)
contains more quantity than a less divisible body of the same volume, such as a liter of granite. For an account of the doctrine in the context of Hobbes’s dispute with Wallis, see Jesseph (1999, pp. 136–142).

3 Hobbes on Space, Body, and Geometry

Hobbes drew a sharp contrast between the inherently conjectural enterprise that he termed "physics" and the a priori demonstrations characteristic of true *scientia*. In Hobbes's words, "the Science of every Subject is derived from a praecognition of the Causes, Generation, and Construction of the same" (*SL* Epistle; *EW* 7: 183). As it happens, Hobbes held that there were only two fully demonstrative sciences: geometry and politics. In the case of geometry we can literally bring geometric objects into being by drawing lines and figures, while a science of the commonwealth can be founded on definitions that show how civil society is generated through human agreement. I will leave aside Hobbes's claims to have founded the one true science of politics, but would like to close by paying some attention to his program for the foundations of geometry and see how it connects with his treatment of space.[17]

Traditionally, geometry was taken to be an abstract inquiry into the properties of magnitudes that are not to be found in nature. Dimensionless points, breadthless lines, and depthless surfaces of Euclidean geometry were not traditionally taken to be the sort of thing one might encounter while walking down the street. Whether such items were characterized as Platonic objects inhabiting a separate realm of geometric forms, or as abstractions arising from experience, it was generally agreed that the objects of geometry and the space in which they are located could not be identified with material objects or the space of everyday experience. This approach raises a number of difficult philosophical problems concerning the relationship between geometric objects or space and their physical counterparts. For Hobbes, however, there are no such problems, because there is and can be no distinction between geometric space and physical space. Further, Hobbes's strict materialistic ontology requires that geometric objects be defined as bodies or as things produced by the motion of bodies.

Hobbes's geometric ontology can be summarized fairly readily by re-writing the traditional Euclidean definitions of the terms 'point', 'line', and 'surface'. A Euclidean point is defined as "that which has no parts" (*Elements*, 1, def. 1). Hobbes dismissed any such conception, remarking "That which is indivisible is not Quantity; and if a point be not Quantity, seeing it is neither substance nor Quality, it is nothing. And if *Euclide* had meant it so in his definition, . . . he might have defined it more briefly (but ridiculously) thus, *a Point is nothing*" (*SL* 1; *EW* 7: 201). The Euclidean definitions of line as "length without breadth" (*Elements* 1, def. 2) and surface as "that which has length and breadth only" (*Elements* 1, def. 5) fare no better, and for essentially the same reason. Instead of these definitions, Hobbes offered the following:

---

[17] More on Hobbes's program for geometry and his extensive dispute with John Wallis, Oxford's Savilian Professor of Geometry, can be found in Jesseph (1999).

> Though there be no Body which has not some Magnitude, yet if when any Body is moved, the Magnitude of it be not at all considered, the way it makes is called a LINE, or one single Dimension; & the Space through which it passeth, is called LENGTH; and the Body itself a POINT; in which sense the Earth is called a *Point*, and the Way of its yearly Revolution, the *Ecliptick Line*. (*DCo* 2.8.12)

Thus, a point is a body so small that its magnitude can be neglected in a demonstration, while a line is the path traced by a point in motion. A surface is then easily enough defined as the trace of a line in motion, and a solid is defined through the motion of a surface.

According to Hobbes, the benefit to be gained by these definitions is twofold. First, they purge geometry of a false ontology that supposes the existence of some sort of immaterial quasi-realm of non-physical objects distinct from physical space and its contents. Second, and perhaps more importantly, Hobbes's re-written definitions include the motions that generate lines, surfaces, and solids. This innovation, he thought, would found geometry on first principles that would give incontrovertible demonstrations and (he expected) would yield new theorems. As he explained

> In that part therefore of my Book where I treat of Geometry, I thought it necessary in my Definitions to express those Motions by which Lines, Superficies, Solids, and Figures were drawn and described; little expecting that any Professor of Geometry should find fault therewith; but on the contrary supposing I might thereby not only avoid the Cavils of the Scepticks, but also demonstrate divers Propositions which on other Principles are Indemonstrable. (*SL* Epistle; *EW* 7: 184–85)

Hobbes illustrated the supposed superiority of his approach by considering the Euclidean definition of a circle as "a plane figure contained by one line such that all the straight lines falling upon it from one point among those lying within the figure are equal to one another" (*Elements* 1, d. 15). Hobbes argued that, although this definition gives an adequate characterization of a circle, it does not show how to generate such a figure, since "if a man had never seen the generation of a Circle by the motion of a Compass or other aequivalent means, it would have been hard to perswade him, that there was any such Figure possible" (*SL* 1; *EW* 7: 205). Instead, a circle should be defined as the figure generated from the motion of a line about one fixed endpoint. With this definition in hand, Hobbes imagined that any question concerning the circle could then be answered. Having grasped the proper cause of the circle, the true geometer need do no more than deduce the properties of his own construction, and this (Hobbes thought) should make such problems as the quadrature of the circle readily solvable.

Employing definitions of geometric objects in terms of the motions that produce them, Hobbes concluded that he had developed a new geometric method, which he termed the "method of motions." The method exploited his kinematic conception of the genesis of magnitudes to solve geometric problems; by attending to the motions that generate magnitudes such as curves, Hobbes convinced himself that he had hit upon a method that would make short work of even the most difficult problem.

In fact, this method has some important similarities with techniques employed by Gilles Personne de Roberval, although it lacked the power that Hobbes attributed to it.[18]

4 Conclusion

Hobbes's program for the foundations of geometry was a thoroughly *sui generis* endeavor, and it found no favor with the philosophers and mathematicians of his day. The same, of course, might be said about his doctrine of space. These two are closely connected, of course, and what connects them is Hobbes's obsession with accounting for absolutely everything in terms of bodies in motion. Geometry is not, as some had taught, a science devoted to exploring the properties of abstract objects in a non-physical space, nor is physical space some kind of mysterious substance distinct from material bodies. In the end, Hobbes held, what is truly real is simply bodies moving and colliding; everything else is some kind of effect produced by that ultimate causal principle. Whether this sort of minimalist materialism has anything to recommend it aside from its parsimony is a question for another day, but we can at least thank Hobbes for trying to see how far one can go on such a basis.

Acknowledgments I would like to thank participants in the "Theories of Space" conference at the Max-Planck-Institut für Wissenschaftsgeschichte for comments on the presentation that led to this paper. My references to Hobbes's works use the following system of abbreviations: *English Works* (Hobbes 1839a–1845a) is abbreviated *EW* with references to volume and page number*; Opera Latina* (Hobbes 1839b–1945b) is abbreviated *OL* with references to volume and page number; *Leviathan* (Hobbes 1651) is abbreviated *L*, with references to part, chapter, and page number; *De Corpore* (Hobbes 1655) and its English translation *Of Body* (Hobbes 1656a) are abbreviated '*DCo*', with references to part, chapter, and section; *Six Lessons* (Hobbes 1656b) is abbreviated '*SL*', with a reference to the Lesson number and to *EW* after a semicolon; *Seven Philosophical Problems* is abbreviated *SPP* with a reference to dialogue number and *EW* after a semicolon; *Critique du de Mundo* (Hobbes 1973) is abbreviated *CDM* with references to chapter, section and page number; *The Correspondence* (Hobbes 1994) is abbreviated *CTH*. References to Euclid's *Elements* are to the Heath edition (Euclid 1925), by book number and proposition/definition number.

References

Brandt, F. (1928). *Thomas Hobbes' mechanical conception of nature*. London: Librairie Hachette.

Descartes, R. (1964–1976). *Oeuvres de Descartes*. In C. Adam & P. Tannery (Rev. ed). Paris: Vrin.

Digby, K. (1644). *Two treatises: In the one of which, the nature of bodies; in the other, the nature of man's soule is looked into: In way of discovery, of the immortality of reasonable souls*. Paris: Gilles Blaizot.

Duhem, P. (1985). *Medieval cosmology: Theories of infinity, place, time, void, and the plurality of worlds* (R. Ariew, Trans.). Chicago: University of Chicago Press.

[18] See Jesseph (1999, pp. 235–238) on Hobbes's method of motions; the connection with Roberval is explored in Malcolm (2002).

Engfer, H.-J. (1982). *Philosophie als analysis: Studien zur Entwicklung philosophischer analysiskonzeptionen unter dem einfluss mathematischer Methodenmodelle im 17. und frühen 18. Jahrhundert*. Stuttgart-Bad Cannstat: Fromann-Holzboog.
Euclid (1925). *The thirteen books of Euclid's "Elements" translated from the text of Heiberg* (T. L. Heath, Trans.), Vol. 3. Oxford: Oxford University Press.
Grant, E. (1969). Medieval and seventeenth-century conceptions of an infinite void space beyond the Cosmos. *Isis, 60*, 39–60.
Herbert, G. B. (1987). Hobbes's phenomenology of space. *Journal of the History of Ideas, 48*, 709–717.
Hintikka, J., & Remes, U. (1974). *The method of analysis: Its geometric origin and general significance*. Boston: D. Reidel.
Hanson, D. W. (1990). The meaning of 'demonstration', in Hobbes's science. *History of Political Thought, 9*, 587–626.
Hobbes, T. (1651). *Leviathan; or the matter, forme, & power of a common-wealth ecclesiasticall and civill*. London: A. Crooke.
Hobbes, T. (1655). *Elementorum philosophiae pars prima: De Corpore*. London: A. Crooke.
Hobbes, T. (1656a). *Elements of philosophy, the first section, concerning body*. London: R. & W. Leybourn for A. Crooke.
Hobbes, T. (1656b). *Six lessons to the professors of the mathematiques, one of geometry, the other of astronomy*. London: J.M. for Andrew Crook.
Hobbes, T. (1839–1845a). *The English works of Thomas Hobbes of Malmesbury* (Vol. 11). In Sir W. Molesworth (Ed.). London: J. Bohn.
Hobbes, T. (1839–1945b). *Thomae Hobbes Malmesburiensis Opera Philosophica* (Vol. 5). In Sir W. Molesworth (Ed.). London: J. Bohn.
Hobbes, T. (1973). *Critique du "De Mundo" de Thomas White*. In J. Jacquot & H. W. Jones (Eds.). Paris: Vrin.
Hobbes, T. (1994). *The correspondence* (Vol. 2). In N. Malcolm (Ed.). Oxford: Oxford University Press.
Jesseph, D. (1999). *Squaring the circle: The war between Hobbes and Wallis*. Chicago: University of Chicago Press.
Jesseph, D. (2006). Hobbesian Mechanics. In D. Garber & S. Nadler (Eds.), *Oxford studies in early modern philosophy* (pp. 118–152). Oxford: Oxford University Press.
Leijenhorst, C. (2002). *The mechanisation of aristotelianism: The late Aristotelian setting of Thomas Hobbes' natural philosophy*. Brill: Leiden, Boston, and Köln.
Malcolm, N. (2002). "Hobbes and Roberval," in Id., *Aspects of Hobbes*, pp. 156–199. Oxford: Oxford University Press.
Prins, J. (1990). Hobbes and the school of Padua: Two incompatible approaches to science. *Archiv für Geschichte der Philosophie, 72*, 26–46.
Shapin, S., & Schaffer, S. (1985). *Leviathan and the air-pump: Hobbes, Boyle, and the experimental life*. Princeton: Princeton University Press.
Talaska, R. A. (1988). Analytic and synthetic method according to Hobbes. *Journal of the History of Philosophy, 26*, 207–237.
Ward, S. (1656). *In Thomae Hobbii Philosophiam Exercitatio Epistolica*. Oxford: Davis.
Watkins, J. W. N. (1965). *Hobbes's system of ideas: A study in the political significance of scientific theories*. London: Hutchinson.
White, T. (1643). *De mundo dialogi tres*. Paris: D. Moreau.

Mathematics and Infinity in Descartes and Newton

Andrew Janiak

1 Introduction

The concept of the infinite has often been regarded as inherently problematic in mathematics and in philosophy. The idea that the universe itself might be infinite has been the subject of intense debate not only on mathematical and philosophical grounds, but for theological and political reasons as well. When Copernicus and his followers challenged the old Aristotelian and Ptolemaic conceptions of the world's finiteness, if not its boundedness, the idea of an infinite, if not merely unbounded, world seemed more attractive. Indeed, the infinity of space has been called the "fundamental principle of the new ontology" (Koyré 1957, p. 126). Influential scholarship in the first half of the twentieth century helped to solidify the idea that it was specifically in the seventeenth century that astronomers and natural philosophers fully embraced the infinity of the universe. As Kuhn writes in his *Copernican Revolution* (1957, p. 289): "From Bruno's death in 1600 to the publication of Descartes's *Principles of Philosophy* in 1644, no Copernican of any prominence appears to have espoused the infinite universe, at least in public. After Descartes, however, no Copernican seems to have opposed the conception." That same year saw the publication of Alexandre Koyré's sweeping volume about the scientific revolution, *From the Closed World to the Infinite Universe*. The decision to describe and conceive of the world as infinite might be seen as a crucial, if not decisive, aspect of the overthrow of Scholasticism. As Kuhn and Koyré knew, one finds a particularly invigorating expression of this historical-philosophical interpretation in an earlier article by Marjorie Nicholson (1929, p. 370):

A. Janiak (✉)
Department of Philosophy, Duke University, 209 West Duke Building,
Durham, NC 27708-0743, USA
e-mail: Janiak@duke.edu

V. De Risi (ed.), *Mathematizing Space*, Trends in the History of Science,
DOI 10.1007/978-3-319-12102-4_9

> Few, however, seem to have noticed the effect of the Cartesian idea of indefinite extension upon one of the most significant of all seventeenth-century conceptions: the idea of infinity, the problem of the possibility of an infinite number of worlds, or of a universe infinitely extended. In this conception lies the key to the characteristic form taken in England, at least, by the idea of progress, and to one of the most profound changes which occurred in seventeenth-century thought; for the real change in men's conceptions of themselves and of the universe came less through Copernicanism than through the expansion of the boundaries of thought through the idea of infinity.

Thus before Koyré and Kuhn's analyses, Nicholson had already forcefully argued that we should not focus all our historical attention on developments in astronomy proper, for it was actually Descartes's idea of the indefinite extension of the world that helped to engender the most important late seventeenth century developments in England, which centered around Isaac Newton's new conception of the universe.

Nicholson was careful not to attribute the idea of an infinite universe to Descartes himself; instead, in his main work in natural philosophy, the *Principia Philosophiae* (1644), the text Kuhn also cites in this connection, he deliberately avoided this idea, arguing instead for what he consistently called the "indefinite extension" of the material world. So if Nicholson is right in suggesting that by century's end philosophers in England like Newton were ready to embrace the world's infinity wholeheartedly—and she surely is right about that notion, as we will see—how did the dialectic from Cartesian indefinite extension to Newtonian infinity actually work? Was Descartes's rejection of the Scholastic bounded world the key maneuver, or was the English reaction to Descartes's unbounded but nonetheless finite world also a key aspect of the story? The suggestion from Koyré, Kuhn, and perhaps even Nicholson, is that it was the Cartesian worldview that was the key move toward the infinite universe, but as we will see, there are strong reasons to think that there was actually a three-stage dialectic in the seventeenth century, from the Scholastic finite bounded world, to the Cartesian unbounded but still finite world, and finally to the infinite world of Newton.

There is an even more remarkable aspect to this story. One might have thought that Newton's infinite world arises in the context of his thinking about the laws of motion, and especially the idea that the true motion of an object must be understood as its absolute motion, its motion with respect to absolute space. For that idea might reasonably be interpreted as entailing the conclusion that space itself, without relation to any objects or relations, must be infinite (given that Newton lacked inertial frames or the like). That may in fact be the case, or one might argue that a merely potentially infinite space is sufficient for absolute motion, since it can be arbitrarily large. Either way, there is another aspect to the dialectic. Newton explicitly adopts a critical attitude toward Cartesian finite but unbounded space on the grounds that it cannot accommodate the presence of the actually infinite creator of the world (cf. Harries 2001, p. 143). Newton's world must be actually infinite because God's infinity has a spatial character.

2 Descartes's Finite Unbounded World

In his original treatise in natural philosophy, *Le Monde*, which Descartes withdrew from publication after Galileo's house arrest in 1633, we find a strong endorsement of the Copernican view of the earth's motion, a view that Kuhn links to the idea of an infinite world. But since *Le Monde* did not see the light of day, Kuhn carefully notes that it was Descartes's *Principia Philosophie*, published in Amsterdam eleven years later, which played a key role in the shift toward reconceiving of the material world. As Nicholson notes, Descartes did not simply embrace the world's infinity in his text; instead, as with the question of the earth's motion, he articulated a complex view that has exercised his interpreters ever since. Like Galileo before him, he was "wary of infinites and infinitesimals" (Mahoney 1997, p. 745). His view centered on his idea of the world's indefinite extension, a concept introduced in part because Descartes explicitly contends that the concept of the infinite remains problematic. He makes this point near the very beginning of *Principia Philosophiae*, in section 26 of part one, noting that there are many authors who seek out paradoxes of one kind or another as soon as one contends that some item is infinite:

> 26. *We should never enter into disputes about the infinite; things in which we observe no limits, such as the world's extension, the division of the parts of matter, the number of stars,* etc.*, should be regarded as indefinite.*
>
> Thus we will never be involved in tiresome disputations about the infinite. Since we are finite, it would be absurd for us to determine anything concerning the infinite, for this would be to attempt to limit it and comprehend it. So we shall not respond to those who ask if half an infinite line would also be infinite, or whether an infinite number is odd or even, and the like; for unless one regards one's own mind as infinite, one should not judge such matters. And for us, with anything in which we are unable to discover a limit from point of view, we shall not affirm that it is infinite, but instead consider it indefinite. Thus, since there is no imaginable extension that is so great that we cannot understand the possibility of a greater one, we shall describe the magnitude of possible things as indefinite. And however many parts a body is divided into, each of the parts can still be understood to be divisible, so we shall think that quantity is indefinitely divisible. And no matter how great we imagine the number of stars to be, we still think that God could have created more, so we shall suppose their number is indefinite; and the same with the rest. (AT 8a: 14–15)

Descartes's motivation for introducing the concept of the indefinite into his metaphysics is clear: if one mentions infinite quantities, such as an infinite line, one is forced into "tiresome" debates and arguments. So one ought to avoid speaking of infinity altogether (except in the case of the divine, of course).[1] But there certainly are phenomena, processes and objects that would appear to be something other than ordinary finite items—these are items that at least appear to have no limits in the ordinary sense. For instance, if we divide a table in half, we recognize that we can divide each of those halves in half, and do the same again: at least as a conceptual

[1] The next section of *Principia Philosophiae* presents a clearer view: we should reserve the term 'infinite' for God, since we positively recognize that God has no limits, which apparently is not true in the cases quoted above.

matter, leaving aside any particular physical or nomological constraints on the process of division, there does not appear to be any limit to this process. So to recognize this fact, even while avoiding the idea that the table is infinitely divisible, Descartes declares that we should regard the table as indefinitely divisible.[2]

Some of Descartes's readers immediately regarded his distinction between the infinite and the indefinite as echoing Aristotle's distinction between actual and potential infinity, where the latter is explicitly introduced in order to avoid paradoxes involving the former. Aristotle's discussion of infinity (Bowin 2007) reflected the idea that there are various reasons for discussing infinite quantities: two of them are the possible use of such quantities in mathematics, and the suggestion, often attributed to Zeno, that ordinary phenomena such as a person running or a bird flying might somehow involve infinite quantities.[3] Of course, these two arenas need not *force* one to accept the idea of an infinite quantity. There was a strong tendency in some Greek authors, e.g., to insist that even in Euclidean geometry, we need not explicitly conceive of any object of our analysis as bearing an infinite property or feature: we can think of a line or of a space as extending an arbitrary distance, but need not think of it as extending infinitely (Moore 1990, p. 43), and Aristotle apparently endorsed that view (Bernadete 1964, p. 14). Nonetheless, there are many questions about quantities in mathematics, and the examples attributed to Zeno were often held to pose an especially difficult problem, one that Aristotle took seriously in the *Physica*. In that text, he outlines his general approach as follows, having just argued that there is no body that is actually infinite:

> But on the other hand to suppose that the infinite does not exist in way leads obviously to many impossible consequences: there will be a beginning and an end of time, a magnitude will not be divisible into magnitudes, number will not be infinite. If, then, in view of the above considerations, neither alternative seems possible, an arbiter must be called in; and clearly there is a sense in which the infinite exists and another in which it does not.
>
> We must keep in mind that the word 'is' means either what potentially is or what fully is. Further, a thing is infinite either by addition or by division.
>
> Now, as we have seen, magnitude is not actually infinite. But by division it is infinite. (There is no difficult in refuting the theory of indivisible lines.) The alternative then remains that the infinite has a potential existence. (*Physica* 206[a] 8–18; transl. Aristotle 1908-)

Aristotle then used his distinction between potential and actual infinity to solve various problems, including the problem of how to think about the paradoxes attributed to Zeno.[4] Whether this maneuver enables Aristotle to deal successfully

---

[2] Descartes seems to have thought that the idea of the world as indefinite was original with him, but there were Scholastic philosophers who held similar, if not identical, views; others, including some Jesuits, criticized the notion. See Ariew (1999, pp. 165–171).

[3] This is my way of approaching the material; in the *Physica*, Aristotle himself contends that "belief in the existence of the infinite" comes mainly from five considerations, including quantities dealt with in mathematics, the nature of time, the division of magnitudes, and so on (203[b] 15–24).

[4] In the *Physica*, he writes of Zeno (233[a] 22–31; transl. Aristotle 1908-, as modified by Barnes 1984): "Hence Zeno's argument makes a false assumption in asserting that it is impossible for a thing to pass over or severally to come in contact with infinite things in a finite time. For there are two ways in which length and time and generally anything continuous are called infinite: they are

with Zeno's paradoxes is an open question. But there is no doubt that his distinction enabled many later philosophers to think that they had avoided the fundamental problems associated with the idea of infinite quantities (Bernadete 1964, pp. 53–57).

Descartes employs his distinction between the infinite and the indefinite once again in a letter to Clerselier, the editor of his correspondence in the 1650s, aimed at explicating some of the ideas in the third meditation, which of course attempts to prove the existence of God, an actually infinite being. On 23 April 1649, amidst his correspondence with More among others, he writes:

> By infinite substance I understand a substance that has perfections that are true and real, actually infinite and immense. This is not an accident superadded to the notion of substance, but the very essence of substance taken absolutely and limited by no defects; such defects, in regard to substance, are accidents; but infinity or infinitude is not. It should be remarked that I never use the word *infinite* for signifying solely the lack of limits, which is negative and to which I have applied the word *indefinite*, but for signifying a real thing, which is incomparably greater than all those which are in some way limited. (AT 5: 355–56)

Descartes seems to indicate here that the word 'infinite' is reserved for referring to a being that has perfections within its very essence, rather than as features added to it. In contrast, the word 'indefinite' is reserved for referring to entities that lack limits of one kind or another. If it is merely the case that an entity lacks limits of some kind, then in the Cartesian system the word 'infinite' must not be applied to it.

Descartes holds at least two clear views in this area designed to express our positive knowledge in metaphysics: first, we know that we are finite; second, we know that God is infinite. The question is, has he expressed a third view using his concept of the indefinite? Many readers regard this idea as ambiguous, and in fact, as ambiguous in a twofold manner, with an intersection between them. For starters, it is ambiguous between these two claims: (1) all actual entities and processes are either finite or infinite (only God falls into the latter category), so "indefinite" items are *possible* entities and processes; and, (2) some actual entities and processes are indefinite. The reason to endorse (1) is clear: in the examples he presents, Descartes does not focus on the indefinite property of any *actual* entities or processes. Instead,

---

(Footnote 4 continued)
called so either in respect of divisibility or in respect of their extremities. So while a thing in a finite time cannot come in contact with things quantitatively infinite, it can come in contact with things infinite in respect of divisibility; for in this sense the time itself is also infinite: and so we find that the time occupied by the passage over the infinite is not a finite but an infinite time, and the contact with the infinites is made by means of moments not finite but infinite in number." Perhaps Aristotle is suggesting here that although a person walking a hundred meters across the college quad cannot traverse an infinite number of things in the sense of things that are quantitatively infinite—which we can read as an actual infinity—she can traverse an infinite number of things in the sense of things that are infinitely divisible—which we can read as a potential infinity. Just as the hundred meters that the person walks is infinitely divisible into smaller and smaller segments, from meters to centimeters to millimeters and so on, the time that it takes her to cross the quad is also infinitely divisible into smaller and smaller moments, from minutes to seconds to milliseconds and so on.

he focuses on the fact that *possible* objects—like a possible greater extension than the actual extension in the world—*possible* processes—like a possible process of dividing a material body into its constituents—and *counterfactual possibilities*—like the fact that God could have created more stars than currently exist—should be regarded as being indefinite, rather than infinite, in character. This has led some readers to infer that Descartes did not claim explicitly that any *actual* object, event or process is itself indefinite.

However, these considerations are not decisive, because Descartes also seems to contend that the material world itself—which is identical to space or extension—is indefinite. Since the world is actual, one has a reason to endorse (2) above. But we then find that (2) itself is ambiguous, between the following two claims: (2a) there are *actual* items that are indefinite, that is, which are neither finite nor infinite; and, (2b) as far as our *knowledge* reaches, we cannot say whether certain items are finite or not, so we *regard* them as indefinite. Not surprisingly, given these two potential disambiguations of the Cartesian concept of the actual indefinite, interpreters of *Principia Philosophiae* have articulated an epistemic construal and a metaphysical construal of Descartes.

On the epistemic construal (cf. Ariew 1987), the view reflects a fundamental limit to human knowledge: we cannot perceive any limits to the material world, indeed, we cannot even *conceive* of it as having any limits, but for all that, we do not have a *positive* conception of its infinity, as we do in the case of God.[5] So the appellation *indefinite* merely reflects our lack of knowledge. No actual item within our ontology is anything other than finite or infinite; nothing can be indefinite itself. This interpretation immediately raises a question: how can Descartes contend *both* that it is inconceivable that matter should have limits *and* that for all we know, the world might actually be finite and therefore have limits after all? Surely from the premise concerning inconceivability we can derive the conclusion that the world cannot be finite. But this inference is unacceptable for Descartes: in the Cartesian system, the fact that I cannot conceive, e.g., how "2 + 2 = 5" could possibly be true does *not* entail that it *cannot* be true—it simply means that God's ways are beyond my understanding (recall the doctrine of the eternal truths). This may also be bolstered by Margaret Wilson's famous argument: Descartes's claim that it is inconceivable that the world is limited is not identical to the claim that I have a clear and distinct perception that the world is unlimited (Wilson 1986, pp. 349–50). I simply cannot conceive that it is limited. Hence the *inconceivability* of matter's finitude is compatible with its *actual* finitude. Defenders of the epistemic construal might agree.

---

[5] In 1671, Leibniz argued that Descartes's distinction between the infinite and the indefinite is merely epistemological; he may have been the first to endorse an epistemic reading of Descartes's view. For his part, Leibniz certainly endorsed the idea that features of reality involve actual infinities, at least in his late work: in the monadology, it seems clear that Leibniz thinks that the world, or perhaps even an individual object like a chair, contains an actual infinity of monads. See Moore (1990, p. 79).

One benefit of this epistemic interpretation is that it seems to capture a significant aspect of Descartes's elaboration of his view: when discussing items or processes or events that are said to be indefinite, he lists various *possibilia* (e.g., the potential division of an object). This fits nicely with the view that all *actual* items, processes and events must be either finite or infinite. We would then speak of various possibilities as involving indefinite quantities (or something analogous) on the grounds that we lack the relevant knowledge of such items. Perhaps this idea would be supported by the thought that we do not obtain knowledge of such *possibilia* until they become actual, at which point they are either finite or infinite. Of course, the actuality of any such possible item does not entail that we have obtained, or can obtain, knowledge of it, but it does mean, perhaps, that it is knowable in some specifiable sense. Clearly, puzzles abound in this area.

On the metaphysical construal, we deny that Descartes is making any epistemic point, contending instead that there are actually three metaphysical categories into which everything fits: there are the finite things, such as me and the White House; there is the one infinite thing, viz., God; and then there are the indefinite things, such as space, i.e., the material world. What does this mean? Some have argued that it is best to view this idea through an Aristotelian lens: whereas Descartes thinks of God as *actually* infinite, as a completed infinity of which we have a positive conception, he thinks of the material world as merely *potentially* infinite, as an item that is limitless or unbounded, but actually finite. The clam that space or matter has no *limits* is compatible with the idea that it is potentially infinite, for the latter entails that there is no last part of space, or smallest piece of matter: for any given place, P, or any given bit of matter, M, we have P + n and M + n, their respective successors. However, the distinction between finite and indefinite items requires clarification, for any potentially infinite item is still actually finite, even if it has no limits.[6] So one possible reading is this: unlike the finite items, which have limits (whether we perceive or recognize them or not), the indefinite items are also finite, but lack any such limits. Whether this reading can be rendered rigorous is a remaining question that I will not tackle here.[7]

[6] Thanks to Henry Mendell for making this point.

[7] Perhaps the distinction between the metaphysical and the epistemic readings is not clear: either way, we still have only finite and infinite items in the ontology; according to one reading, the metaphysical, we can say that things like the material world are finite but potentially infinite, and therefore without limits in a certain sense—which would distinguish them from ordinary finite things like tables, which do have limits—and according to the epistemic reading, we would say that each item in our ontology is finite or infinite and by "indefinite" we would simply be signaling the fact that we do not know whether certain things, like the material world, are finite or infinite. But this might be compatible with the claim that the material world could be potentially infinite, i.e., finite but without limits. It could be. Of course, there would still be *one* distinction between the two readings: on the metaphysical reading, we *would know* that the material world is indefinite, by which we could mean, potentially infinite; and on the epistemic reading, we would *not know* whether the material world is finite—whether potentially infinite, and therefore without limits, or just plain finite, and therefore with limits—or infinite.

As Wilson showed some time ago, it is difficult to resolve the debate between the epistemic and the metaphysical interpretations of the Cartesian view because there are texts that count in favor of each of them, and there is no clear way of determining a priority among those texts (Wilson 1986). Happily, there is a meta-level view that captures the commonality of the epistemic and the metaphysical interpretations. The meta-level claim is this: Descartes argued that if we limit ourselves to expressing our knowledge, we must say that God alone is actually infinite—we do not *know* anything else that is actually infinite. Otherwise put, we lack a "positive" conception of any other actually infinite item. This is compatible with both the epistemic and the metaphysical interpretations because each can provide an analysis of what this claim means. On the epistemic construal, the claim means that although space *may be* infinite, we cannot know, or even conceive of the *possibility*, that it is. What the claim means on the metaphysical construal is that space is known *not* to be actually infinite; instead, it is potentially infinite. Thus it remains the case that the two interpretations are incompatible with one another, but each is compatible with the claim that we lack a positive conception of any actually infinite item (other than God) because each can construe that claim in a way that renders it compatible with the relevant interpretation.

Descartes clarifies our knowledge that God is actually infinite, along with the relation between that knowledge and our self-knowledge as finite but unlimited (in some sense) beings in a famous passage from the third meditation. It represents part of his attempt to reject an objection to his causal proof of God's existence. The objection is roughly this: in reply to Descartes's argument that only an actually infinite substance (or being) could be the cause of my idea of such a substance, on the grounds that there must be at least as much formal reality in the cause of my idea of an actually infinite substance as there is objective reality in that idea, the objector suggests that perhaps a finite being like me could in fact be the cause if I could simply begin with my limited knowledge and my limited positive features and then increase them "more and more to infinity." In that case, I could possibly be the cause of my idea of an actually infinite—and perfect—being after all, which would block Descartes's argument. One aspect of the objection is this: the idea of an actually infinite being could possibly be caused by a finite being that exhibits characteristics, such as an endless growth in its knowledge, which highlight the fact that it bears a potential infinity.[8] This aspect of the objection seems reasonable: leaving aside various obvious temporal limits and physiological factors, it does seem unobjectionable to assert that there is no *inherent* limit to the amount of knowledge a person can achieve, hence each person bears what we might call potentially infinite knowledge. That is, our knowledge is always finite, but perhaps we can regard it as unbounded.

[8] As Broughton highlights (2002, pp. 151–53), one might also find the third meditation proof unpersuasive because it seems to rely on an obscure, or at least not fully clarified, conception of what the representation of an infinite being involves. That representation is connected with a cluster of ideas, including the notion that there is more "reality" in an infinite being than in a finite one, that require clarification beyond what Descartes provides.

Descartes's reply to this objection in the third meditation is illuminating because it reflects his understanding and employment of the distinction between potential and actual infinity:

> But none of this is possible. For first, although it is true that there is a gradual increase in my knowledge, and many things are potential and not yet actual, none of this pertains to the idea of God, in which there is absolutely nothing potential; indeed, this gradual increase in knowledge is the surest sign of imperfection. Furthermore, even if my knowledge always increases more and more, I understand that it will never be actually infinite, since it will never reach the point where it is incapable of another increase. In contrast, I take God to be actually infinite [*actu infinitum*], such that nothing can be added to his perfection. (AT 7: 46–47)

Hence the distinction between actual and merely potential infinity, as Descartes understands it, is crucial because it enables him to argue that there is a clear epistemic difference—a clear difference in our ideas—between a finite being with potentially infinite knowledge and an actually infinite being with actually infinite knowledge. This seems to connect with a difference in the sense in which a finite being like me and the actually infinite being are each unbounded. The former type of being is unbounded in just the sense that she could always add to her knowledge; the latter is unbounded in the sense that it already encompasses an actual epistemic infinity. In that way, the notion that God already knows everything that there is to know is expressed here through the concept of actual infinity.

3 The Cartesian Origins of Newton's Infinite World

As with the fundamental idea of inertia and the laws of motion, the rejection of the Aristotelian distinction between the sublunary and the superlunary, and even the scope of natural philosophy, Newton begins where Descartes left off, or more precisely, he begins where the debate between Descartes and More in 1648–1649 left off (Lewis 1953). Whereas Descartes expended considerable energy rejecting Aristotelian ideas within natural philosophy, Newton seems to have believed that such a project would be a waste of time: unlike Descartes and other influential predecessors such as Galileo, Newton does not bother to ridicule such ideas as the Aristotelian definition of motion (much lampooned throughout the century, of course). Instead, he focused his principal critical energies on the Cartesian system. He also follows Henry More's detailed critical reaction to Cartesianism, which is most clearly in evidence in the unjustly ignored correspondence between More and Descartes, just before Descartes's death. It was More who first criticized and rejected the Cartesian distinction between the infinite and the indefinite in just the way that Newton does in his now famous, unpublished, anti-Cartesian tract *De Gravitatione* (since it was untitled, it is known after its first line).

In his first letter to Descartes, written on the 11th of December 1648, from Christ's College, Cambridge, Henry More heaps praise upon his correspondent and then humbly suggests that there are a few arguments and concepts in *Principia Philosophiae* that puzzle him. One of the most significant is the Cartesian view that space and body are numerically identical; another is the notion that God is not extended, despite the divine omnipresence throughout nature; and a third is the connected view that space or the material world is indefinite rather than infinite (I tackle the question of God's omnipresence below). About this third issue, More writes:

> Fourthly, I do not comprehend your indefinite extension of the world. For this indefinite extension is either infinite *simpliciter* or infinite only to us. If you mean infinite extension *simpliciter*, why do you hide your meaning with excessively modest words? If you mean infinite only for us, the extension will in reality be finite, for our mind is not the measure of things or of the truth. (AT 5: 242)[9]

More favors an epistemic construal of the Cartesian doctrine: Descartes must believe that all ontological items are finite or infinite, so by calling something *indefinite*, he can mean only that the item is actually finite or infinite; it would be "infinite for us" only in the sense of having no perceived limits. It would be an expression of our lack of knowledge.

More's criticisms of Descartes set the stage for Newton's systematic deconstruction of Cartesian natural philosophy in *De Gravitatione*. In that text, Newton discusses three Cartesian doctrines regarding infinity. First, he agrees with the view found in the third meditation, for example, that it is an error to say that "we do not understand what an infinite being is, save by negating the limitations of a finite being" (AT 7: 45). Newton says that when we conceive of a limited being—say, a wooden table in front of us—part of our conception of that being is bound up with our conception of its limits. The table of course is bounded by its wooden surface. But when we think of an infinite being, we are conceiving of something that is "maximally positive," for we are thinking of it as having no features that involve limitations. This idea must obviously be clarified. But what is significant for our purposes is that even in agreeing with Descartes here, Newton may also be diverging from him by holding his view for a reason that Descartes cannot accept. For in this very same paragraph, Newton emphasizes that those who believe that we cannot understand an infinite being should consider the fact that geometers have no difficulty in understanding the infinite: they "accurately" know "positive and finite quantities of many surfaces infinite in length" (*De Gravitatione*, p. 24). This means that Newton differs from some ancient geometers. It also hints that Newton will diverge from the Cartesian view that there is an important sense in which we cannot fully grasp the infinite being: we can in fact use geometry to assist us in understanding infinite beings, such as infinite objects in geometry. I tackle the dialectic between Descartes and Newton on this issue below.

The second Cartesian doctrine concerns the distinction between the infinite and the indefinite, which Newton decidedly rejects:

---

[9] In a letter of 5 May 1651, More agrees with Anne Conway's claim that there isn't any clear distinction in Descartes between the infinite and the indefinite (this is connected with an interpretation of section 21 of part two of *Principia Philosophiae*). Gabbey (1977, pp. 589–90).

> If Descartes should now say that extension is not infinite but rather indefinite, he should be corrected by the grammarians. For the word 'indefinite' ought never to be applied to that which actually is, but always looks to a future possibility, signifying only something which is not yet determined and definite. Thus before God had decreed anything about the creation of the world (if ever he was not decreeing), the quantity of matter, the number of the stars, and all other things were indefinite; once the world was created, they were defined. Thus matter is indefinitely divisible, but is always divided either finitely or infinitely (Part I, article 26; Part II, article 34). Thus an indefinite line is one whose future length is still undetermined. And so an indefinite space is one whose future magnitude is not yet determined; for indeed that which actually is, is not to be defined, but either does or does not have boundaries and so is either finite or infinite. (*De Gravitatione*, p. 24)

Here we see that unlike More, who apparently favors an epistemic construal of Descartes's view, chiding him for not clearly articulating it, Newton indicates that the concept of the indefinite should be applied only to *possibilia*, and not to actual items or processes. This certainly captures an aspect of the discussion from part one, section 26 in *Principia Philosophiae* quoted above, for Descartes's examples often involve possible properties and processes, such as the possible—for Descartes, indefinite—division of some quantity. Had Descartes restricted himself to thinking of mere *possibilia* as indefinite, this objection would lack any bite. But of course, Descartes also claims that the material world itself is indefinite. For Newton, this is an error: we can say that my future granddaughter is of indefinite height, because she does not exist, but we cannot say of my eight-year-old son that his height is indefinite: he is actual, and therefore his height is determined.

The third doctrine discussed by Newton in *De Gravitatione* is this: we should avoid considering space to be infinite because "it would perhaps become God because of the perfection of infinity." Here we find a classic pre-modern metaphysical issue. On Newton's reading, Descartes sides with the old way of thinking by taking infinity to be a perfection per se; Newton rejects this view, siding with the moderns, which would eventually include Leibniz. From Newton's point of view, infinity is not a perfection per se; it is, as it were, value neutral. This is an essential component in the shift from what has been called a metaphysical conception of the infinite, with deep roots in medieval philosophy, to a mathematical conception, which was bound up with new mathematical techniques in the second half of the seventeenth century (Moore 1990). Those techniques will become relevant below.

What then is the outcome of Newton's reaction to these three Cartesian doctrines? First, we find that Newton takes us to understand infinite beings; second, we find that he wishes to regard space as infinite rather than as merely indefinite; and third, he denies that his second view raises problems for his view of God.

What is most remarkable about this discussion of Descartes, however, is that this set of passages in *De Gravitatione* sets the stage for Newton's introduction of his most significant, and complex, doctrine concerning both the ontology of space and its exact relation to the divine. More precisely: after criticizing Cartesian metaphysics and natural philosophy in depth in the first eight pages of *De Gravitatione*, Newton shifts toward a discussion of his own understanding of the ontology of space. He famously begins by denying that space is either a substance or an accident—a view found in other thinkers in this period, such as Charleton and

Gassendi, and to some extent, Isaac Barrow—and then presents a series of numbered paragraphs concerning space. The first paragraph details the mathematical figures that exist within space (see McGuire 2007); the second contends that space is "extended infinitely in all directions," which leads to his long digression concerning the Cartesian doctrine of the *indefinite*; the third paragraph indicates that the parts of space are motionless, which is an aspect of his anti-Cartesian distinction between space and body. And this is the fourth paragraph:

> 4. Space is an affection of a being just as a being. No being exists or can exist which is not related to space in some way. God is everywhere, created minds are somewhere, and body is in the space that it occupies; and it follows that space is an emanative effect of the first existing being, for if any being whatsoever is posited, space is posited. (*De Gravitatione*, p. 25)

We now have two key Newtonian views before us: (1) space is extended infinitely in all directions; and (2) God is everywhere within space. Clearly, (1) contradicts the Cartesian view that space or the material world is merely indefinite, for on either the epistemic or the metaphysical reading of that view, we cannot assert positively that space is infinite (we cannot say that it is actually infinite). If Descartes is understood as holding that the material world is finite—reserving the use of 'indefinite' for possibilities—then Newton clearly denies that view as well.

Regarding (2), the claim that God is everywhere in space is precisely one of the ideas that More presses Descartes to accept, without success. So once again, the correspondence sets the stage for *De Gravitatione*. In More's December 1648 letter to Descartes—already quoted from above—we find the following argument:

> And, indeed, I judge that the fact that God is extended in his own way follows from the fact that he is omnipresent and intimately occupies the universal machine of the world and each of its parts. For how could he have impressed motion on matter, which he did once and which you think he does even now, unless he, as it were, immediately touches the matter of the universe, or least did so once? This never could have happened unless he were everywhere and occupied every single place. Therefore, God is extended in his own way and spread out; and so God is an extended thing [*res extensa*]. (AT 5: 238–39)

More is arguing as follows: Descartes must agree that God's power to act is omnipresent, for that view is entailed by his occasionalism and is also endorsed on more general grounds; yet how could God act on any body to which God was not present? More takes the impossibility of this notion to entail that God must be extended. That is, God must in fact be substantially omnipresent.

One might infer that in *De Gravitatione*, Newton is explicitly endorsing the view outlined by More in his correspondence with Descartes, for Newton says that "God is everywhere." But the correspondence actually serves to highlight the fact that Newton's claim is ambiguous between: (1) God's power to act, or God's action, is everywhere; and, (2) God is actually everywhere, or substantially present everywhere. It is not clear from the text of *De Gravitatione* that Newton's readers have the resources to resolve this ambiguity in his view. It is therefore remarkable that Newton tackles precisely this same issue in the General Scholium, which was added to the second (1713) edition of *Principia mathematica* under the editorship of

Roger Cotes. Indeed, the ambiguity of the idea expressed in *De Gravitatione* highlights the central importance of the General Scholium for determining Newton's considered view on this topic. In the General Scholium, Newton writes:

> Every sentient soul, at different times and in different organs of sense and motions, is the same individual person. There are parts that are successive in duration and coexisting in space, but neither of these exist in the person of man or in his thinking principle, and much less in the thinking substance of God. Every man, insofar as he is a thing that has senses, is one and the same man throughout his lifetime in each and every organ of his senses. God is one and the same God always and everywhere. He is omnipresent not only in power, but in substance: for power cannot subsist without substance [Omnipraesens est non per *virtutem* solam, sed etiam per *substantiam*: nam virus sine substantia non potest]. (Newton 1972, vol. 2, p. 762)[10]

This text clearly resolves the potential ambiguity in *De Gravitatione*: Newton now argues explicitly that God is substantially present everywhere. This is clearly a strong endorsement of More's view against Descartes's contrary opinion.

This discussion in the General Scholium is important for another reason. This canonical formulation of the Newtonian conception of the divine being indicates that Newton regarded space itself as actually infinite. Consider this argument:

1. God is actually infinite.
2. Claim (1) should be read to mean that God's substance, and not just God's power, is actually infinite.
3. Claim (2) means that God substantially occupies all of space.
4. If there were finitely many spatial points, God's substance would be finite and bounded.
5. If there were a potential infinity of spatial points, God would be potentially infinite, that is, finite but unbounded.
6. Not 4 and not 5, by 1, 2, 3.
7. Therefore, space is actually infinite.

This would appear to show that Newton is fully committed to the idea that space itself is actually infinite, and not merely potentially infinite, as Descartes may have believed. He therefore rushed in where Descartes feared to tread.

This interpretation has two components worth mentioning. First, Newton's distinction between virtual and substantial omnipresence, itself a reflection of More's debate with Descartes, maps onto the distinction between potential and actual infinity.

---

[10] It is possible that for More, who was a more or less standard Anglican, the notion of God's substantial omnipresence had no special connection with the Trinity; but for Newton, who was obviously a heretical Anglican, one who rejected the Anglican view of the Trinity, it may also be possible that the doctrine of divine omnipresence was in fact connected with his view of the Trinity. See Snobelen (2005) and (2006). For his part, Clarke defended the view that God is substantially omnipresent in the twelfth section of his third letter to Leibniz, quoting from the General Scholium passage reproduced above; see Koyré (1957, p. 248).

If God is virtually omnipresent, we might construe this to mean that if God chooses to act within a certain spatial area—say, to make a rainbow after 40 days of rain, or to set a bush alight without it being consumed—then God becomes actually present at that location at that time. But this is not the case for Newton's God: in his view, God is already actually present everywhere in space, even when God is not acting or choosing to act in a given location. Because Newton follows More in thinking that God is actually, substantially omnipresent, he must conceive of space as actually infinite.

Second, the interpretation also indicates why we must think of space from an absolute or mathematical perspective, rather than from a relative or common perspective, to capture God's infinity. Consider this famous passage from the Scholium to the Definitions in *Principia mathematica*:

> Absolute space, by its own nature without relation to anything external, is always homogeneous and immobile. Relative space is any movable measure or dimension of this absolute space, which is determined by our senses from the situation of the space with respect to bodies and is popularly used for immobile space, as where the dimension of space under the earth, in the air, or in the heavens, is determined by its situation relative to the earth. (Newton 1972, vol. 1, p. 46)

We can also profitably map the distinction between absolute and relative space onto the distinction between actual and potential infinity. The conception of relative spaces, relative places, and relative times involves at most the notion of a potential infinity. For we can never say that we have determined the *largest possible* relative space—we can always take the three objects that are definitive of the space in three dimensions and conceive of them as being a greater distance apart (we might think of this as mapping onto the natural numbers). We also can never contend that we have found the smallest possible relative place, or the shortest possible relative time. What these ideas about space, time and place give us *just is* the idea of potentially infinite measures of the absolute quantities of space and time. No measure of space or of time will ever be the smallest or largest, shortest or longest—for any arbitrary measure M, there will be what we can call a successor to M, M + n, and this is true along the one temporal dimension and along all three spatial dimensions. This would mean, in turn, that the quantities themselves—absolute or mathematical space and time—would be actually infinite. And it makes good sense, finally, to think of a measure of some actually infinite quantity like Euclidean space itself as involving a merely potential infinity, for the measure can be as large as one likes—it is unlimited in that crucial but restricted sense. (As we know, Newton would not call it *indefinite*.) When we restrict ourselves to the objects of sense perception, we develop the idea of arbitrarily large—but still finite—measures of the quantities space and time. In order to conceive of space and time themselves, however, we must conceive of two infinite quantities, and in order to do that, in turn, we require the representational capacities of geometry. Sense perception will not do.

Newton is clearly more sanguine than Descartes that philosophers need not fear infinite quantities because geometers and other mathematicians are capable of generating clear ideas of such quantities when pondering lines, planes and other geometric objects, not to mention Euclidean space itself. Descartes's avoidance of the infinite was not merely an expression, or a reflection, of his view that we must avoid the Morean idea of God's extendedness; it was also a reflection of his conviction that various paradoxes and irrelevant questions will attend any discussion of the infinite, at least within philosophy. (Whether Descartes regarded such paradoxes as attending mathematical discussions of the infinite is another, and intriguing, question). If one entertains the thought that a line is infinite, one will be forced into "tiresome" questions about whether half that line is also infinite. Descartes and Newton obviously held many distinct, if not opposed, views in the philosophy of mathematics (Guicciardini 2009), just as they held distinct, if not opposed, views in philosophy more generally. So the question is: did Newton hold some specific view that he regarded as saving him from "tiresome" arguments concerning the infinite?

4 Mathematics, Wallis, and Newton's Avoidance of Paradoxes of the Infinite

A key document for understanding Newton's view of the infinite falls in between *De Gravitatione* and the General Scholium.[11] It is part of Newton's correspondence with Richard Bentley, who was preparing to publish his Boyle lectures in 1693, having recently delivered them—the first in what became a famous series—in London. The entire correspondence is framed by Newton's first letter to Bentley, which indicates that he wrote *Principia mathematica* with an eye toward inclining people to believe in a deity. Questions about infinite space arise immediately in the second paragraph of the first letter, although we do not here see any distinction between actually infinite and potentially infinite space. Bentley is clearly struggling with understanding various kinds of arguments about space and the distribution of matter within it that make use of the notion of the infinite. He apparently does not see how to bolster his criticisms of those who employ concepts of infinite matter or infinite space to oppose him. Intriguingly, Newton chooses to instruct Bentley in thinking about the infinite rather generally by referring him to some basic ideas from John Wallis's famous 1656 text, *Arithmetica Infinitorum*. Newton suggests that if Bentley comes to understand Wallis's basic approach to the infinite, he will see that his opponents are committing a fallacy in their arguments against him.

[11] Presumably, on any reasonable view of the dating of the former, it must have been written before the first edition of *Principia mathematica*, which was ready in 1686—see Ruffner (2012).

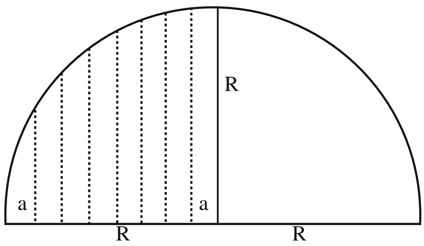


Fig. 1 Wallis tackles the quadrant of the circle

In *Arithmetica Infinitorum*, which Newton read and commented on as a young man at Trinity College,[12] Wallis sought to create a new "arithmetic of indivisibles" to parallel Cavalieri's famous "geometry of indivisibles".[13] The goal was to make progress on an old problem tackled by Cavalieri among many others, namely the problem of the "quadrature" of curvilinear figures, which included both the calculation of the area under some curve, and the calculation of a volume enclosed by some curvilinear figure. Wallis argued in particular that any plane surface can be conceived of as comprising an infinite number of parallelograms—this obviously required him to think of the parallelograms as infinitesimals or as indivisibles, since they must somehow sum to a finite quantity expressing a feature of the plane surface. Wallis proposed to use this general method as a means of "squaring the circle." He tackled the quadrant of the circle as follows (Fig. 1).

The parallelograms comprising the circle would have equal bases a, such that:

$$\mathrm{a} = \mathrm{R}/\infty.$$

[12] In his Trinity notebook, Newton may have made use of some of Wallis's techniques from *Arithmetica infinitorum* (Newton 1983, pp. 106–107, including footnote 168), and he made two pages of annotations from Wallis in that text (Newton 1967–1981, vol. 1, pp. 89–90). As Whiteside indicates, in another pocket book from 1664–1665, Newton made detailed entries concerning Wallis: Newton (1967–1981, vol. 1, pp. 91–141). In the *Que[a]estiones quedam Philosoph[i]cae*, Newton noted: "one infinite extension may be greater than another," a key point from Wallis that Newton would describe to Bentley nearly 30 years later (Newton 1967–1981, vol. 1, p. 89). Whiteside notes, intriguingly, that Newton may have read Hobbes's attack on Wallis first, and then proceeded to read Wallis for himself (Newton 1967–1981, vol. 1, p. 89 note 1). Newton also retained a copy of Wallis's *Opera mathematica* in his personal library (Harrison 1978). David Rabouin points out that by roughly 1680, Newton had decided that he no longer needed the techniques of Wallis. This is an important point, but it's compatible with the fact that Newton still regarded Wallis as indicating how we can avoid various kinds of paradoxes when thinking about infinitesimals, infinite divisibility, and infinity more generally, as his correspondence with Bentley a decade later indicates.

[13] For discussions of Wallis's work, see Guicciardini (2009, pp. 140–47), which places it in the context of understanding Newton's work in mathematics; and, Stedall (2010), which places it within the history of mathematics more broadly.

Wallis argued that we could use a summation of an infinite number of indivisible or infinitesimal quantities in order to make the analysis of some finite quantity—such as the area under a curve—tractable. The key to Wallis's summation techniques, which help to transform geometric problems through the use of arithmetic sequences, is to remember that his infinite number of constituents of any finite quantity, such as a plane surface, retain a definite ratio to that original quantity, such that they sum to the original quantity.

Wallis's technique obviously raises the question: how precisely are we to think about these infinitesimals or indivisibles? For instance, is an infinitesimal parallelogram, which we are meant to conceive of as a constituent of some finite quantity, distinct from a line? That is, does it lack width altogether? Wallis apparently thought that the infinitesimal parallelogram differed from a line because the former's width is not zero; instead, its width is smaller than any assignable finite width. One question, of course, is whether such a notion can be made rigorous and clear. For his part, Wallis did not seem especially concerned with this issue (Stedall 2010, p. xxix). He dealt very freely with infinite products and infinitesimals (Guicciardini 2009, p. 146), even as others, most prominently Hobbes, raised objections against his practice.

Wallis's general approach to thinking about infinity is important for understanding Newton's letter to Bentley of 17 January 1693. The letter bears quoting at length:

> But you argue in the next paragraph of your letter that every particle of matter in an infinite space has an infinite quantity of matter on all sides & by consequence an infinite attraction every way & therefore must be *in equilibrio* because all infinites are equal. Yet you suspect a paralogism in this argument, & I conceive the parallogism lies in the position that all infinites are equal. The generality of mankind consider infinites no other ways than definitely, & in this sense they say all infinites are equal, though they would speak more truly if they should say they are neither equal nor unequal nor have any certain difference or proportion one to another. In this sense therefore no conclusions can be drawn from them about the equality, proportions or differences of things, & they that attempt to do it, usually fall into paralogism. So when men argue against the infinite divisibility of magnitude, by saying that if an inch may be divided into an infinite number of parts, the sum of those parts will be an inch, & if a foot may be divided into an infinite number of parts, the sum of those parts must be a foot, & and therefore since all infinites are equal those sums must be equal, that is, an inch equal to a foot. The falseness of the conclusion shows an error in the premises, & the error lies in the position that all infinites are equal. There is therefore another way of considering infinites used by mathematicians, & that is under certain definite restrictions & limitations whereby infinites are determined to have certain differences or proportions to one another. Thus Dr Wallis considers them in his *Arithmetica Infinitorum*, where by the various proportions of infinite sums he gathers the various proportions of infinite magnitudes: which way of arguing is generally allowed by mathematicians & and yet would not be good were all infinites equal. According to the same way of considering infinites, a mathematician would tell you that though there be an infinite number of infinitely little parts in an inch, yet there is twelve times that number of such parts in a foot; that is, the infinite number of those parts in a foot is not equal to, but twelve times bigger than, the infinite number of them in an inch. And so a mathematician will tell you that if a body stood *in equilibrio* between any two equal and contract attracting infinite forces, & if to either of those forces you add any new finite attracting force: that new force how little so ever will destroy the equilibrium & put the body into the same motion into

> which it would put it were those two contrary equal forces but finite or even none at all: so that in this case two equal infinites by the addition of a finite to either of them become unequal in our ways of reckoning. And after these ways we must reckon if from the consideration of infinites we would always draw true conclusions. (Newton 1959, vol. 3, p. 239)

We might read Newton here as explaining to Bentley that mathematicians such as Wallis, and presumably, Newton himself, have ways of thinking about infinite quantities, or infinite processes—such as infinite divisibility—that avoid the false assumption guiding philosophical discussions, viz. that all infinites are equal. In particular, mathematicians following Wallis contend that the proportion between two finite quantities is preserved when one considers each of those quantities to be infinitely divisible. Hence an infinitely divisible foot remains twelve times the size of an infinitely divisible inch. Since there can be preserved proportions between items that have an infinite feature, such as being infinitely divisible, it follows, says Newton, that there can be different sized infinities. Newton had already grasped this exact point in 1664–1665 as a student at Trinity College reading Wallis (Newton 1967–1981, vol. 1, p. 89). So we have a specific mathematical view that enables us to reject the faulty philosophical presumption guiding reasoning about the infinite.

This conception of infinity in Wallis would also, *mutatis mutandis*, enable Newton—and Bentley, if he follows Newton in this respect—to evade exactly the kind of "tiresome" questions that Descartes mentions in *Principia Philosophiae*. If we reject the presumption that all infinites are equal, then we have a straightforward answer to the question Descartes mentions in section 26 of part one of *Principia Philosophiae*: would an infinite line divided in half result in two infinite lines? The answer is that half an infinite line would remain infinite, and indeed, it would retain its proportion (1/2) to the original line, for as Newton wrote in his undergraduate book, "one infinite extension may be greater than another."

Wallis's approach to thinking about infinity, which guided Newton already in his very earliest days, and which Newton cited 30 years later when instructing Bentley on how to present arguments concerning nature that will incline his readers toward believing in the deity, enabled Newton to accomplish a task that Descartes eschewed explicitly in part one of *Principia Philosophiae*. Newton explicitly sought to discuss infinity in his philosophizing about the deity, nature and motion because he thought he could avoid exactly the kinds of "tiresome" arguments, and classical paradoxes, that had hampered discussions of the infinite since antiquity. So Wallis's bold translation of old geometrical problems into problems involving the summation of arithmetic sequences, where he is willing to sum over an infinite number of parts of finite figures, enables Newton to conclude that one can conceive of infinity in a way that avoids paradoxes. Once liberated from paradox, the philosopher is free to embrace the concept of infinity, and to think of various quantities as infinitely large or as infinitely small. Once that move has been made, in turn, it is not a stretch for the philosopher to contend that we ought to regard space as actually infinite. And that is precisely what Newton contends.

5 Conclusion

Two conclusions seem apt. First of all, it is tempting to interpret Cartesian metaphysics as exhibiting a kind of tension. On the one hand, whatever we make of the distinction between the infinite and the indefinite, it is essential to Descartes's view that we regard ourselves as having a positive conception of God's actual infinity. Indeed, if we lack such a conception, then the argument for the existence of God in the third meditation cannot be valid, for that argument applies the causal principle to the distinction between the objective and the formal reality of my ideas by indicating that a finite substance (like me) cannot be the cause of my idea of an actually infinite substance, for the formal reality of the cause of an idea must be at least equal to the objective reality of the idea itself.[14] So we must have a positive conception of the actually infinite substance or being. On the other hand, it is not clear that Descartes explicates what the *content* of our positive conception of the one and only infinite substance really is. There are at least two reasons to think that this explication is hampered—I do not say, rendered impossible—by other Cartesian doctrines. First, we know from sections 26–27 of part one of *Principia Philosophiae* that although we positively know that God has no limits, and is therefore infinite, we are also warned against discussing infinity in any other context because it leads to endless debates and paradoxes. Second, we know from other texts that Descartes exhibits a strong reluctance to discuss God. When discussing More's view that God is everywhere, which I have discussed above, Descartes replies (on 15 April 1649) as follows:

> This "everywhere" I cannot admit. You seem here to make God's infinity consist in his existing everywhere, an opinion I cannot accept. I think on the contrary that by reason of his power, God is everywhere; but by reason of his essence, God has no relation to place at all. But since in God power and essence are not distinguished, I think it is better to argue in such cases about our mind or angels, which are more on the scale of our perception, rather than to dispute about God. (AT 5: 343; cf. Lewis 1953, pp. 160–161)[15]

More and Newton could certainly be forgiven for concluding that Descartes could not, or did not wish, to articulate what his positive conception of divine infinity is.

Perhaps Descartes can resolve this tension (it may merely be a surface tension). It remains important philosophically, however, because it enables us to see the significance of the fact that we can understand Newton's system more thoroughly if we recognize that it lacks any such tension, even on the surface. For in the Newtonian metaphysical system, not only do we have a positive conception of God's actual infinity as a matter of doctrine, if one can put it that way, but in fact one can give a clear content to that positive conception through the clear mathematical

---

[14] This is uncontroversial, although it leaves open Wilson's intriguing question (1986, pp. 354–355): can a finite substance like me be the cause of my idea of the world, which is merely indefinite?

[15] See also Descartes's letter to Mersenne of May 1630 (AT 1: 152), although that is obviously from a much earlier period in his career.

concept of the actual infinity of space. The mathematician has a perfectly clear conception of actually infinite Euclidean space extending in all directions. He knows that this conception is clear because his reasoning about the infinite lies in the Wallisian tradition of thinking about infinity, where various paradoxes and problems are avoided. This idea connects, in turn, to the crucial view that Newton articulates following More's rejection of the Cartesian conception of the divine: God is not merely virtually or potentially present everywhere, but substantially or actually present everywhere. So actually infinite space gives us an entrée into achieving a clear conception of God.

This claim leads to the second moral of my story, which is more speculative. We sometimes read of a distinction between mathematical conceptions of the infinite and metaphysical conceptions of it. Moore (1990) makes this distinction the centerpiece of his recent book on the subject. The intriguing thing about Descartes is this: as we learn from the end of part two of *Principia Philosophiae*, Descartes claims that in "Physica," he admits and requires only the principles of "Geometria" and "Mathesis abstracta."[16] But in "metaphysica," or in "Philosophia prima," Descartes does *not* employ a mathematical conception of the infinite: he does not employ geometrical, arithmetic *or* algebraic notions to conceive of infinity, an infinite being, infinite properties, etc. Instead, Descartes focused on a *metaphysical* conception of the infinite. And as a result, he expresses a strong reluctance to discuss the infinite, and maintains the old view that the infinite is perfect per se. Newton, in contrast, focused primarily on a mathematical conception of the infinite, which is something that had occupied him since he took extensive notes on Wallis's work in his early days at Trinity College. As we have seen, this work remained important to him long after the publication of *Principia mathematica*. Newton also jettisoned Descartes's reticence to speak of the infinite outside of metaphysical contexts. Newton argued, in turn, that we should employ the mathematical conception of the infinite in order to grasp the most important object within traditional metaphysics, the divine being. For Newton, it is precisely the infinity of Euclidean space that *enables* us to conceive clearly of the infinity of the divine.[17] And now for the speculation: does Newton's attempt to give a mathematical conception of infinity *logical priority* in metaphysical contexts reflect his overarching attitude toward the Cartesian system, namely that it fails to employ mathematical principles in a systematic and substantive way? Descartes proclaimed that he employed and required only the principles of geometry and mathematics, but for Newton, Cartesian physics failed to live up to this billing. Hence Newton famously replied

---

[16] Descartes writes: "The only principles that I admit—or require—in physics are those of geometry and abstract mathematics; they explain all natural phenomena, and enable us to provide quite certain demonstrations concerning them" (*Principia Philosophiae*, Part two, § 64; AT 8: 78–79).

[17] As Ted McGuire writes in a recent paper: "If Newton's theology of divine existence grounds the *actuality* of infinite space, geometry underwrites his claim to understand its infinite nature. Clearly, the depth of Newton's dialogue with Descartes must be appreciated if we are adequately to understand his path to this conception" (McGuire 2007, p. 125). I agree.

by proclaiming again the need for *mathematical principles of natural philosophy*. Perhaps we can add to this proclamation an intriguing addendum, viz. the need for the *mathematical principles of metaphysics*.

Acknowledgments For very helpful conversations that substantially altered my argument in this paper, I would like to thank Vincenzo De Risi and David Sanford. Many thanks to the audience at the Max-Planck-Institut für Wissenschaftsgeschichte in Berlin for their help on an earlier version of this paper—the audience included Lorraine Daston, Dan Garber, Graciela De Pierris, Michael Friedman, Jeremy Gray, Gary Hatfield, Doug Jesseph, Clarissa Lee, Brandon Look, Henry Mendell, David Rabouin. All translations are my own unless otherwise noted.

References

Ariew, R. (1987). The infinite in Descartes' conversation with Burman. *Archiv für Geschichte der Philosophie, 69*, 140–163.
Ariew, R. (1999). *Descartes and the last scholastics*. Ithaca: Cornell University Press.
Aristotle, (1908). *The works of Aristotle. Translated into English*. Oxford: Clarendon Press, later Oxford University Press.
Barnes, J. (1984). *The complete works of Aristotle: The revised Oxford translation*. Princeton, N. J.: Princeton University Press.
Bentley, R. (1842). In C. Wordsworth (Ed.), *The correspondence of Richard Bentley*, London: John Murray.
Bernadete, J. (1964). *Infinity: An essay in metaphysics*. Oxford: Oxford University Press.
Bowin, J. (2007). Aristotelian infinity. *Oxford Studies in Ancient Philosophy, 32*, 233–250.
Broughton, J. (2002). *Descartes's method of doubt*. Princeton: Princeton University Press.
Descartes, R. (1969–1975). In C. Adam & P. Tannery (Ed.), *Oeuvres de Descartes* (new ed., Vol. 11). Paris: Vrin.
Descartes, R. (1991). *Principles of philosophy* (V. Rodger, Ed. & R. P. Miller, Trans.). Dordrecht: Kluwer.
Domski, M. (2003). The constructible and the intelligible in Newton's philosophy of geometry. *Philosophy of Science, 70*, 1114–1124.
Gabbey, A. (1977). Anne Conway et Henry More: Lettres sur Descartes (1650–1651). *Archives de Philosophie, 40*, 379–388.
Guicciardini, N. (1999). *Reading the principia: The debate on Newton's mathematical methods for natural philosophy from 1687 to 1736*. Cambridge: Cambridge University Press.
Guicciardini, N. (2009). *Isaac Newton on mathematical certainty and method*. Cambridge, MA: MIT Press.
Harries, K. (2001). *Infinity and perspective*. Cambridge, MA: MIT Press.
Harrison, J. (1978). *The library of Isaac Newton*. Cambridge: Cambridge University Press.
Henry, J. (2013). The reception of Cartesianism. In P. Anstey (Ed.), *Oxford handbook of British philosophy in the seventeenth century*. Oxford: Oxford University Press.
Koyré, A. (1957). *From the closed world to the infinite universe*. Baltimore: Johns Hopkins University Press.
Kuhn, T. (1957). *The Copernican Revolution*. Cambridge, MA: Harvard University Press.
Lewis, G. (Ed.). (1953). *Descartes: Correspondance avec Arnauld et Morus, texte Latin et traduction*. Paris: Librairie philosophique Vrin.
McGuire, J. E. (2007). A dialogue with Descartes: Newton's ontology of true and immutable natures. *Journal of the History of Philosophy, 45*, 103–125.
Mahoney, M. S. (1997). The mathematical realm of nature. In D. Garber & M. Ayers (Eds.), *Cambridge history of seventeenth century philosophy*. Cambridge: Cambridge University Press.

Moore, A. W. (1990). *The infinite*. London: Routledge.
More, H. (1655). *An Antidote Against Atheism*. London: J. Flesher.
More, H. (1659). *Immortality of the Soul, so farre forth as it is demonstrable from the knowledge of nature and the light of reason*. London: J. Flesher.
More, H. (1662). *A collection of several philosophical writings*. London: J. Flesher.
Newton, I. [*De Gravitatione*] (Unknown date). *De Gravitatione et Aequipondio Fluidorum*, MSS Add. 4003. Cambridge: University Library.
Newton, I. (1959–1977). In H. W. Turnbull, et al (Eds.), *The correspondence of Isaac Newton*. Cambridge: Cambridge University Press.
Newton, I. (1967–1981). In D. T. Whiteside, & M. A. Hoskin (Eds.), *The mathematical papers of Isaac Newton*. Cambridge: Cambridge University Press.
Newton, I. (1972). In A. Koyré, I. B. Cohen, & A. Whitman (Eds.), *Philosophiae Naturalis Principia Mathematica* (the 3rd edition with variant readings). Cambridge, MA: Harvard University Press.
Newton, I. (1983). In J. E. McGuire & M. Tamny (Eds.), *Certain philosophical questions: Newton's trinity notebook*. Cambridge: Cambridge University Press.
Newton, I. (1999). *The Principia: Mathematical principles of natural philosophy* (I. B. Cohen & A. Whitman Trans.). Berkeley: University of California Press.
Newton, I. (2014). In A. Janiak (Ed.), *Newton: Philosophical writings* (2nd ed.). Cambridge: Cambridge University Press.
Nicholson, M. (1929). The early stage of Cartesianism in England. *Studies in Philology, 25*, 356–374.
Ruffner, J. (2012). Newton's *De gravitatione*: A review and reassessement. *Archive for the History of the Exact Sciences, 66*, 241–264.
Shapiro, L. (Ed.). (2007). *The correspondence between Princess Elisabeth of Bohemia and René Descartes*. Chicago: University of Chicago Press.
Snobelen, S. (2005). 'The true frame of nature': Isaac Newton, Heresy, and the reformation of natural philosophy. In J. Brooke & I. Maclean (Eds.), *Heterodoxy in early modern science and religion*. Oxford: Oxford University Press.
Snobelen, S. (2006). 'To us there is but one God, the Father': Antitrinitarian textual criticism in seventeenth and early-eighteenth century England. In A. Hessayon & N. Keene (Eds.), *Scripture and scholarship in early modern England* (pp. 116–136). Aldershot, Hampshire: Ashgate.
Stein, H. (1970). Newtonian space-time. In R. Palter (Ed.), *The Annus Mirabilis of Sir Isaac Newton, 1666–1966*. Cambridge, MA: MIT Press.
Stein, H. (2002). Newton's metaphysics. In I. B. Cohen & G. E. Smith (Eds.), *The Cambridge companion to Newton*. Cambridge: Cambridge University Press.
Wallis, J. (1656). *The arithmetic of infinitesimals* (J. Stedall, Ed. and Trans.). New York: Springer. (2010).
Wilson, M. (1986). Can I be the cause of my idea of the world? (Descartes on the infinite and indefinite). In A. Oskenberg Rorty & G. E. Smith (Eds.), *Essays on Descartes' meditations*. Berkeley: University of California Press.

Leibniz's Transcendental Aesthetic

Daniel Garber

Though my title may suggest otherwise, I don't want to argue that Leibniz actually had a transcendental philosophy, or that his account of space and time is in any way identical to that of Kant's. My title is meant to be somewhat provocative. While I will not discuss Kant in any explicit way, the correspondences should be obvious. I hope that I got your attention, but, at the same time, I hope that I won't strain your credulity.

I want to show how for Leibniz, as for Kant, things in themselves, bodies, are in an important sense not geometrical. Instead, I will argue, geometrical extension, which Leibniz characterizes as "ideal" and radically distinct from concrete reality, is something external to the concrete world of bodies which we apply to them. I will begin by looking at some published and unpublished texts of Leibniz's that deal with body, and its relation to force, motion, extension, and individuality, setting aside Leibniz's views on the ultimate metaphysical foundations. Only after that will I then examine Leibniz's views on the ultimate make-up of the world, be it corporeal substances or monads, and discuss the relation that that metaphysic has on the question of the relation between extension and body.

I should mention at the very beginning that I am not going to talk about Leibniz's theory of physical space, and his celebrated disputes with Newton and Clarke. My subject is related but different. What interests me most centrally is the relation between body and geometrical extension.

Let me begin with a passage from an unpublished essay thought to be from the mid-1680s, "De modo distinguendi phaenomena realia ab imaginariis":

D. Garber (✉)
Department of Philosophy, Princeton University,
1879 Hall, Princeton NJ 08544-1006, USA
e-mail: dgarber@Princeton.edu

V. De Risi (ed.), *Mathematizing Space*, Trends in the History of Science,
DOI 10.1007/978-3-319-12102-4_10

> Concerning bodies I can demonstrate that not merely light, heat, color and similar qualities are apparent but also motion, figure, and extension. And that if anything is real, it is solely the force of acting and suffering, and hence that the substance of a body consists in this (as if in matter and form).[1]

Leibniz's point here is that what is real in bodies is not their geometrical properties, but their forces, active and passive, and that active and passive forces constitute, in some sense, the form and matter that join together to form a body.

In order to understand Leibniz's account of body here, we must turn to his doctrine of force. The locus classicus of his views is in the *Specimen dynamicum* (1695), where he presents a detailed account of this conception of body and force. Leibniz writes:

> *Active force* (which might not inappropriately be called *power* [*virtus*], as some do) is twofold, that is, either *primitive*, which is inherent in every corporeal substance per se … or *derivative*, which, resulting from a limitation of primitive force through the collision of bodies with one another, for example, is found in different degrees. Indeed, primitive force (which is nothing but the first entelechy) corresponds to the *soul or substantial form* …. Similarly, passive force is also twofold, either primitive or derivative. And indeed, the *primitive force of being acted upon* [*vis primitiva patiendi*] or of *resisting* constitutes that which is called *primary matter* in the schools, if correctly interpreted. This force is that by virtue of which it happens that a body cannot be penetrated by another body, but presents an obstacle to it, and at the same time is endowed with a certain laziness, so to speak, that is, an opposition to motion, nor, further, does it allow itself to be put into motion without somewhat diminishing the force of the body acting on it. As a result, the *derivative force of being acted upon* later shows itself to different degrees in *secondary matter*.[2]

There are two distinctions here, between active and passive force, and between primitive and derivative force, giving rise to four kinds of force, primitive and derivative active force, and primitive and derivative passive force.

Derivative forces, are the forces most familiar to us, the variable physical magnitudes connected with motion and the resistance to motion. Leibniz writes in the *Specimen dynamicum*:

> Therefore, by derivative force, namely, that by which bodies actually act on one another or are acted upon by one another, I understand … only that which is connected to motion (local motion, of course), and which, in turn, tends further to produce local motion. For we acknowledge that all other material phenomena can be explained by local motion.[3]

Derivative forces are the kinds of forces that appear in the laws of physics. Leibniz writes:

---

[1] A VI, 4, 1504 (L 365). Cf. *Discours de métaphysique* § 12, A VI, 4, 1545 (AG 44). (English translations, when available, are given in parentheses following the original language citations). There are many other passages in which Leibniz claims that our ideas of extension contain something imaginary. See, e.g., A VI, 4, 1622 (RA 315); A VI, 4, 1465; A VI, 4, 1612–13; etc.

[2] GM VI 236–37 (AG 119–20). Cf. the account given in the earlier draft, Leibniz (1982, p. 66). A very similar account is given in an unpublished essay dated May 1702; see GP IV 395 (AG 252).

[3] GM VI 237 (AG 120).

> It is to these notions [i.e., the derivative forces] that the laws of action apply, laws which are understood not only through reason, but are also corroborated by sense itself through the phenomena.[4]

These variable derivative forces, Leibniz holds, are grounded in something in body, the primitive active and passive forces. These primitive forces are constituents of body, something found in body, whose modifications are the derivative forces found in the physical world. Writing to Johann Bernoulli in 1698, Leibniz explains:

> If we conceive of soul or form as the primary activity from whose modification secondary [i.e. derivative] forces arise as shapes arise from the modification of extension, then, I think, we take sufficient account of the intellect. Indeed there can be no active modifications of that which is merely passive in its essence, because modifications limit rather than increase or add.[5]

In an exposition of his metaphysics of body, written in May 1702, Leibniz writes similarly that "active force is twofold, primitive and derivative, that is, either substantial or accidental."[6] In the passage from the *Specimen dynamicum* quoted above Leibniz characterizes the primitive active force as corresponding to "the soul or substantial form"; the primitive passive force, on the other hand, is characterized as constituting "that which is called *primary matter* in the schools, if correctly interpreted." Form and matter join together to constitute a (corporeal) substance. In the May 1702 exposition of his account of body, Leibniz writes:

> Primitive active force, which Aristotle calls first entelechy and one commonly calls the form of a substance, is another natural principle which, together with matter or [primitive] passive force, completes a corporeal substance.[7]

The primitive passive force of a body, what Leibniz calls its matter in the *Specimen dynamicum*, is that from which arises impenetrability and the resistance to motion. And it is from this that arises extension in bodies. In an important passage appended to a letter to Arnauld, Leibniz argues that insofar as matter is just passive force, "…in this sense matter would not be extended or divisible, although it would be the principle of divisibility or of that which amounts to it in the substance."[8] Similarly, in another passage dated at 1685 (?) Leibniz writes that "Matter is the force of being acted upon or of resisting in any body whatsoever, from which follows a certain extension in body, unless the Author of things desires otherwise."[9] (This passage is primarily concerned with the problem of the Eucharist. In this context, the fact that God can create passive force without thereby creating something extended is

---

4 GM VI 237 (AG 120).

5 Leibniz to Johann Bernoulli, 18/28 November, 1698, A III, 7, 944 (AG 169).

6 GP IV 395 (AG 252).

7 GP IV 395 (AG 252).

8 Leibniz to Arnauld, 9 October 1687, A II, 2, 251. This is generally thought to be a later addition to the letter.

9 "Specimen demonstrationum catholicae seu apologia fidei ex ratione," A VI, 4, 2326.

something that can only happen by miracle).[10] Leibniz's larger point is simply that extension is not basic to body, but derives from the passive force in body, in particular, from impenetrability by virtue of which one body excludes other bodies from occupying the same place.

Now that we have a better conception of Leibniz's conception of body in relation to force, we can return to the passage quoted earlier:

> Concerning bodies I can demonstrate that not merely light, heat, color and similar qualities are apparent but also motion, figure, and extension. And that if anything is real, it is solely the force of acting and suffering, and hence that the substance of a body consists in this (as if in matter and form).[11]

Leibniz's foil here is obviously Descartes. For Descartes, the essence of body is extension. This is understood in a very strong sense: when Descartes claims that the essence of body is extension, he means to say that strictly speaking, the only properties that bodies really have are geometrical. In this sense one might say that Cartesian bodies are the objects of geometry made real.[12] The contrast is with sensible qualities, such as heat, color, taste, etc. These qualities are grounded in the geometrical qualities of bodies in the sense that heat, color, taste, etc. are caused by the broadly geometrical configuration of smaller parts that make up larger bodies, their corpuscular substructure. But strictly speaking, these sensible qualities aren't really in bodies themselves. What Leibniz is claiming in these passages is that despite Descartes' claims, the geometrical properties have roughly the same status that the sensible qualities have: like the sensible qualities, they are appearances, the effects of a causal process that starts with something physically and metaphysically more basic. In the strictest sense, the geometrical properties of body aren't really in body any more than the sensible properties are. But what is basic in body? Force, Leibniz claims, "the force of acting and suffering."

This argument aims to distinguish geometrical extension from body. But there are others. Leibniz writes in the 1692 "Animadversiones in partem generalem Principiorum Cartesianorum":

> … the notion of extension is not a primitive one but is resolvable. For an extended being implies the idea of a continuous whole in which there is a plurality of things existing simultaneously. To speak of this more fully, there is required in extension, the notion of which is relative, a something which is extended or continued as whiteness is in milk, and that very thing in a body which constitutes its essence; the repetition of this, whatever it may be, is extension.[13]

---

[10] Adams (1994, p. 349ff.), discusses passages like these as part of an argument that Leibniz was an idealist in the period under discussion. However, the fact that extension can be separated from force only supernaturally makes these passages problematic for his case.

[11] A VI, 4, 1504 (L 365).

[12] On Descartes' conception of body, see Garber (1992, Chap. 3).

[13] GP IV 364 (L 390).

The argument is given at greater length in a letter to de Volder in 1699:

> I do not think that there is a substance constituted from extension alone, since the concept of extension is incomplete. Nor do I think that extension is conceived of through itself but that it is a resolvable and relative notion. For it is resolved into plurality, continuity and coexistence, i.e. number and continuity, but also to time and motion, while coexistence is only added in that which is extended. But from this it appears that something must always be assumed that is continued or diffused, such as whiteness is in milk, color ductility, and weight are in gold, and resistance is in matter. For in itself continuity (for extension is nothing but simultaneous continuity) no more completes a substance than multitude or number, where there must be something numbered, repeated and continued. And I believe that our thinking is completed and terminated more in the notion of dynamism than in that of extension, and no other notion of power or force should be sought than that it is an attribute from which change follows whose subject is substance itself.[14]

Again, the target here is Descartes and the idea that bodies are just geometrical extension made real. A comparison between number and extension is helpful here to see Leibniz's point. It is obvious that with numbers, while we can talk about three loaves of bread in the physical world, or three persons, it doesn't make sense to talk about an object in the physical world that is *just* three, and nothing else. Leibniz's point is that extension works the same way. While we can talk about some quality—whiteness, or ductility, or weight—being extended in the world, it doesn't make sense to talk about an object in the world that is just extended and nothing else. In this way Leibniz carefully distinguishes bodies as they are in themselves—forces, active and passive—from their extension.

The contrast between the realm of mathematical extension and the realm of real bodies is emphasized in yet another way in Leibniz's solution to the labyrinth of the composition of the continuum.

In order to appreciate Leibniz's solution, though, I must say something about the problem of the continuum, as Leibniz understood it. The problem goes back to Aristotle and *Physics* Chap. 6:

> Now if the terms 'continuous', 'in contact', and 'in succession' are understood as defined above—things being continuous if their extremities are one, in contact if their extremities are together, and in succession if there is nothing of their own kind intermediate between them—nothing that is continuous can be composed of indivisibles: e.g. a line cannot be composed of points, the line being continuous and the point indivisible. For the extremities of two points can neither be *one* (since of an indivisible there can be no extremity as distinct from some other part) nor *together* (since that which has no parts can have no extremity, the extremity and the thing of which it is the extremity being distinct).[15]

That which is continuous for Aristotle is that whose parts are such that "their extremities are one." That is, a continuum is that which is made up of parts that share a boundary and are thus not genuinely distinct from one another. It is this

---

[14] LDV 72–73. Leibniz repeats the same basic argument a number of times in his correspondence with de Volder: see LDV 224–25, 304–305, 322–23. It is also found in later writings, e.g. GP IV 393–4 (AG 251) and GP VI 584 (AG 261).

[15] *Physica* VI 231a 18–28 (Aristotle (1984, vol. 1 p. 390)).

passage that Leibniz seems to have in mind when he writes in the 1676 dialogue, “Pacidius Philalethi”:

> I remember that Aristotle, too, distinguishes the contiguous from the continuous in such a way that things are *continuous* whose extrema are one, and *contiguous* whose extrema are together.[16]

Or, as he often puts it, “a continuum that whose parts are indefinite.”[17] A *dis*continuous magnitude is thus one whose parts don’t share boundaries, and thus have a kind of natural division. And so Leibniz sometimes characterizes the discontinuous as that “which has already been actually cut into parts by nature.”[18] The problem, then is this. It is obvious that a continuous magnitude (a line, a surface, a solid) cannot be made up out of points. But if not of points, then of what can it be composed?

In response to Simon Foucher’s critique of the “Système nouveau” of 1695, Leibniz wrote one of the clearest accounts of his solution to the problem. In the reply to Foucher, Leibniz draws a clear distinction between the ideal world of mathematical entities (lines, surfaces, numbers), and the world of concrete things. About geometrical objects he writes:

> Extension or space and the surfaces, lines, and points one can conceive in it are only relations of order or orders of coexistence, both for the actually existing thing and for the possible thing one can put in its place. Thus they have no bases of composition, any more than does number. A number divided, 1/2 for example, can be further divided into two fourths or four eighths, etc. to infinity, without our being able to arrive at any smallest fractions or to conceive of the number as a whole that is formed by the coming together of ultimate elements.[19]

The problem of the composition of the continuum is concerned with the parts from which continua can be constructed. Leibniz’s point is that the ideal geometrical continuum does not have such parts, strictly speaking, nor does it need them: insofar as a geometrical continuum has parts, they come from the division of the line, and these parts are not properly elements of that line. However, in real concreta, matters are different:

> But, in actual substantial things, the whole is a result or coming together of simple substances, or rather of a multitude of real unities.[20]

When dealing with real concrete things, the whole is indeed composed of parts. However, real concrete bodies are not *continua* but *contigua*, things which have real parts with genuine boundaries. The problem of the composition of the continuum is thus solved:

[16] A VI, 3, 537 (RA 149).

[17] A VI, 4, 668; cf. A VI, 4, 393, 565, 637, 1002, etc.

[18] A VI, 3, 563 (RA 205).

[19] GP IV 491 (AG 146).

[20] Ibid.

> It is the confusion of the ideal with the actual which has muddled everything and caused the labyrinth of *the composition of the continuum.* Those who make up a line from points have looked for the first elements in ideal things or relations, something completely contrary to what they should have done; and those who found that relations like number or space (which contain the order or relation of possible coexistent things) cannot be formed by the coming together of points were wrong, for the most part, to deny that substantial realities have first elements, as if the substantial realities had no primitive unities, or as if there were no simple substances.[21]

The objects of geometry, which exist in the realm of the ideal, are continuous, but not composed of parts; the real objects that exist in the physical world are composed of parts, but they are not continuous. In this way the problem of the composition of the continuum is resolved.[22]

But what does Leibniz mean when he says that continuous extension is *ideal*? Let me begin with some quotations. In a 1702 reply to Bayle, "Réponse aux réflexions continues dans la seconde Edition du Dictionaire Critique de M. Bayle," Leibniz writes:

> I acknowledge that time, extension, motion, and the continuum in general, as we understand them in mathematics, are only ideal things—that is, they express possibilities, just as do numbers. Even Hobbes has defined space as a phantasm of the existent.[23]

Here there are two elements worth noting. First of all, the continuum "expresses" possibilities. But the reference there to Hobbes introduces something else interesting. The reference is to Chap. 7 of *De corpore* (1655), Hobbes's physics. Hobbes begins the physics by imagining the entire world annihilated, except for one subject's mind. He then asks himself how this mind would conceive of various things in the world. The first thing that Hobbes considers is space, which he conceives of as the phantasm (imagination) left in the mind of the magnitude of a body, as he puts it.[24] Or, more formally:

> Space is the Phantasme of a Thing existing without the Mind simply [*Spatium est phantasma rei existentis quatenus existentis*]; that is to say, that Phantasme, in which we consider no other Accident, but onely that it appears without us.[25]

Which is to say, space as defined by Hobbes is a phantasm: an object that exists in the mind, indeed, in the imagination. It seems here as if Leibniz means to be endorsing this view, as applied to continuous magnitude.

These two elements, that the continuum relates to possibilities and that it is phenomenal, (only) in the mind, come up again and again in passages where Leibniz talks about the continuum as something ideal. The connection between

21 Ibid.

22 There is also an echo of this discussion in Leibniz to de Volder, 19 January 1706, LDV 332–33.

23 GP IV 568 (L 583).

24 Hobbes (1655, Chap. 7 §2).

25 Hobbes (1655, Chap. 7 § 2), translation from Hobbes (1656).

mathematics and the mind, and between continuous magnitude and the imagination are frequent, and not surprising. In a famous passage in the preface to the *Nouveaux Essais*, Leibniz argues that the principles of mathematics (along with logic, metaphysics and morals) are innate, that is, in the mind:

> …it appears that necessary truths, such as we find in pure mathematics and particularly in arithmetic and geometry, must have principles whose proof does not depend on instances nor, consequently, on the testimony of the senses, although without the senses it would never occur to us to think of them. This is a distinction that should be noted carefully, and it is one Euclid understood so well that he proves by reason things that are sufficiently evident through experience and sensible images. Logic, together with metaphysics and morals, of which the one shapes natural theology and the other natural jurisprudence, are full of such truths, and consequently, their proof can only arise from internal principles, which are called innate. It is true that we must not imagine that we can read these eternal laws of reason in the soul from an open book, as the edict of the praetor can be read from his tablet without effort and scrutiny. But it is enough that they can be discovered in us by dint of attention…[26]

In a letter to de Volder, Leibniz asserts that the objects of mathematics (I suspect that he has geometry specifically in mind) are in the mind:

> From the fact that mathematical body cannot be resolved into primary constituents, it may be inferred that it is certainly not real, but something mental, designating nothing other than the possibility of parts, not something actual.[27]

Later in the same paragraph of the letter he refers to (geometrical?) space, where the parts are “indefinite” as “a mental thing.”

It should not be surprising that the ideal is located in the mind. This, indeed, is directly connected with the central meaning of “ideal.” The Académie Française dictionary of 1694 defines “idéal” as “Qui n’est qu’en idée.” The 1762 edition is a bit more explicit: “Qui existe dans l’idée, qui n’existe que dans l’entendement.”

It should be noted here, though, that even though continuous geometrical extension is in the mind, something mental, Leibniz is not committed to a completely subjective account. Though in the human mind, it is not *only* in the human mind. In a letter to the Electress Sophie, 31 October 1705, Leibniz writes:

> Continuity uniformly ordered, though it is only a matter of assumption and abstraction, grounds eternal truths and necessary sciences. It is the object of the divine understanding, as are all truths, and its rays extend to ours.[28]

In this way it seems that the continuous geometrical extension that exists in the human mind is just a reflection of what there is in the divine mind.

Geometrical extension is ideal in the sense that it exists (only) in mind, in our mind or in God’s. But, Leibniz also holds, it is closely connected with possibility. Writing to de Volder in January 1705 (?) he notes:

[26] A VI, 6, 50 (AG 292–93).

[27] LDV 302–3.

[28] GP VII 564.

> …in mathematical extension, by which possibles are understood, there is no actual division and there are no parts except those that we make by thinking.[29]

In his last letter to de Volder, 19 January 1706 he writes:

> …continuous quantity is something ideal that pertains to possible things and to actual things insofar as they are possible things … Indeed, when we—confusing ideal things with real substances—seek actual parts in the order of possible things and indeterminate parts in an aggregate of actual things, we entangle ourselves in the labyrinth of the continuum and in inexplicable contradictions.[30]

In an earlier draft of that same letter, Leibniz writes:

> In real things there is nothing but discrete quantity, i.e., a multitude resulting from true unities. Continuous quantity, which is not apparent but exact, pertains to ideal things and possibilities since it involves something indefinite or indeterminate, which is not allowed by the actual nature of things.[31]

Continuous extension is ideal, and thus relates to "possibilities."

But this last quotation suggests a puzzle. The ideal relates to the possible, rather than to the actual. But, Leibniz also suggests, the "actual nature of things" does not allow for the real existence of continuous extension. Why is continuous extension impossible in nature? And if so, in what sense is continuous extension related to the possible?

To understand why continuous extension is impossible *in rerum natura*, I want to turn to the Correspondence with Arnauld, where this is a major theme. Basic to the view that Leibniz articulates in the exchange is the principle that properly speaking, existence must be grounded in genuine individuals. Leibniz writes in the letter of 30 April 1687:

> …there is no multiplicity without true unities. To be brief, I hold as axiomatic the identical proposition which varies only in emphasis: that what is not truly *one* entity is not truly one *entity* either. It has always been thought that 'one' and 'entity' are interchangeable. Entity is one thing, entities another; but the plural presupposes the singular, and where there is no entity, still less will there be many entities. What clearer statement can be made?[32]

Leibniz seems to have liked this formulation. He used a very similar one more than 15 years later in a letter to de Volder (20 June 1703): "…if there is nothing that is *truly one*, then every *true thing* will be eliminated."[33] There can, of course, be entities composed of collections of such true unities, aggregates. But their existence presupposes the existence of true unities. Leibniz writes, again in the letter of 30 April 1687:

---

[29] LDV 320–21.

[30] LDV 332–33.

[31] LDV 338–39.

[32] A II, 2, 185–86 (M 121).

[33] LDV 262–63

> I believe that *where there are only entities through aggregation, there will not even be real entities*: for every entity through aggregation presupposes entities endowed with a true unity, because it obtains its reality from nowhere but that of its constituents, so that it will have no reality at all if each constituent entity is still an entity through aggregation; or one must yet seek another basis to its reality, which in this way, if one must constantly go on searching, can never be found.[34]

Extended things, insofar as they are divisible, must be made up of parts, and those parts must ultimately be genuine individuals. Leibniz wrote in a draft for the letter of 8 December 1686:

> Now, each extended mass can be considered as composed of two or a thousand others; there exists only an extension achieved through contiguity. Thus one will never find a body of which it may be said that it is truly one substance. It will always be an aggregate of many. Or rather, it will not be a real entity, since the parts making it up are subject to the same difficulty, and since none never arrives at any real entity, because entities made up by aggregation have only as much reality as exists in their constituent parts.[35]

In order to be real, bodies must either be genuine individuals, or composed of genuine individuals.

What, though, constitutes a genuine individual for Leibniz? In the letter of 8 December 1686 he writes:

> There is as much difference between a substance and such an entity [i.e. a pool of fish or a flock of sheep] as there is between a man and a community, such as a people, army, society or college, which are moral entities, where something imaginary exists, dependent upon the fabrication of our minds. Substantial unity requires a complete, indivisible and naturally indestructible entity … which cannot be found in shape or in motion … but in a soul or substantial form after the example of what one calls self.[36]

What exactly are these genuine individuals that Leibniz has in mind here? In the Correspondence with Arnauld, I would argue, it is corporeal substances, extended bodies unified by non-extended souls that are at issue. In the same letter Leibniz wrote:

> …every part of matter is in fact divided into other parts…and since it continues endlessly in this way, one will never arrive at a thing of which it may be said: 'Here really is an entity', except when one finds animate machines whose soul or substantial form creates substantial unity independent of the external union of contiguity. And if there are none, it follows that apart from man, there is apparently nothing substantial in the visible world.[37]

Later on, for example in the correspondence with de Volder, the basic unities seem to be monads. Whether the unities are corporeal substances or monads is not important for the argument at hand: what is important is that the reality of extended things requires that there be such genuine unities.

It should be noted that in the view of the Correspondence with Arnauld, we needn't stop with the first layer of corporeal substances we come upon. Leibniz's

---

[34] A II, 2, 184 (M 120).

[35] A II, 2, 114–15 (M 88).

[36] A II, 2, 121 (M 94).

[37] A II, 2, 122 (M 95).

basic building blocks themselves contain further corporeal substances, and so on ad infinitum.[38] So, Leibniz writes about human beings in the letter to Arnauld of 9 October 1687:

> … man … is an entity endowed with a genuine unity conferred on him by his soul, notwithstanding the fact that the mass of his body [*la masse de son corps*] is divided into organs, vessels, humours, spirits, and that the parts are undoubtedly full of an infinite number of other corporeal substances endowed with their own forms.[39]

And what is true about human beings is true about all corporeal substances, for Leibniz. As a consequence, Leibniz holds that “everything is full of animate bodies.”[40]

Because of this, it follows that concrete extended bodies are *contigua*, and *not continua*. Indeed, this is metaphysically *necessary* for Leibniz: if extended things did not contain real unities as genuine parts, they could not be real. In a letter from 4 March 1687 Arnauld tried to press Leibniz on this. At one point, he asked the following:

> I see no drawback to believing that in the whole of corporeal nature there are only ‘machines’ and ‘aggregates’ of substances, because of none of these parts can one say, accurately speaking, that it is a single substance. That indicates only what is very proper to note, as did St. Augustine, that thinking or spiritual substance is in that respect much more excellent than extended or corporeal substance, that only the spiritual has a true unity and a true self, which the corporeal does not have … [I]t may be of the essence of matter not to have true unity, as you admit of all those bodies which are not joined to a soul or substantial form.[41]

Leibniz answered in his letter of 20 April 1687:

> I grant you, Sir, that in the whole of corporeal nature there are only machines (which are often animated), but I do not grant that ‘there are only aggregates of substances’, and if there are aggregates of substances there must also be genuine substances from which all the aggregates result. One must therefore necessarily arrive either at mathematical points from which certain authors make up extension, or at Epicurus’s and M. Cordemoy’s atoms (which you, like me dismiss), or else one must acknowledge that no reality can be found in bodies, or finally one must recognize certain substances in them that possess a true unity.[42]

To Arnauld’s last suggestion, that unity may simply not be essential to matter, Leibniz replied:

> You object, Sir, that it may be of the essence of body to be devoid of true unity; but it will then be of the essence of body to be a phenomenon, lacking all reality as would a coherent dream, for phenomena themselves like the rainbow or a heap of stones would be wholly imaginary if they were not composed of entities possessing true unity.[43]

[38] This view is developed in detail in Garber (2009, Chap. 2).

[39] A II, 2, 251 (M 154).

[40] A II, 2, 249 (M 151).

[41] A II, 2, 154 (M 108).

[42] A II, 2, 184–85 (M 120–21).

[43] A II, 2, 186 (M 122).

It follows from this line of argument that it is metaphysically impossible, on Leibniz's view, for there to be genuine continuously extended bodies in the real world. If *continua* lack natural parts, then a continuous extended body would lack constituent parts. And if it lacked constituent parts, it would lack unities. And if it lacked unities, it could not exist in reality. An indeterminate extended magnitude, which is what a *continuum* is, cannot really exist. To repeat a quotation from the letters with de Volder I gave earlier,

> In real things there is nothing but discrete quantity, i.e., a multitude resulting from true unities. Continuous quantity, which is not apparent but exact, pertains to ideal things and possibilities since it involves something indefinite or indeterminate, which is not allowed by the actual nature of things.[44]

And so, for Leibniz it is *literally impossible* that there should be continuously extended bodies in the world.

(As an aside here, I should note that this argument from the Correspondence with Arnauld applies only to continuous extension in body: it does not apply to time, motion, or empty space. There are numerous passages, some of which I have quoted earlier, that suggest that (continuous) time, space, and motion are also ideal, like continuous extension, but these require independent arguments. In these cases, though, I would strongly suspect that the ideality of the notions in question—time, space and motion—are not derived from the fact that they are continuous, but from other features of their natures. But this is not a question I can take up at this time.)

A real continuously extended body cannot exist in nature. What, then, does Leibniz mean when he says that continuous magnitude, as something ideal, "expresses possibilities" or is something "by which possibles are understood" or "pertains to possible things and to actual things insofar as they are possible things" and the like?

In the letter to de Volder of January 1705 Leibniz wrote:

> Hence number, hour, line, motion, i.e. degree of speed, and other ideal quantities of this kind, i.e., mathematical entities, are not in fact aggregated from parts, since the way in which someone may wish to assign parts in them is completely undetermined. Actually, it is necessary that they be understood in this way, since they signify nothing other than the mere possibility of assigning parts in any way whatever.[45]

This is one way in which we can think of continuous mathematical extension as dealing with possibles: while it is not itself actually divided, it is something that *can be* divided, that contains parts in a potential way. In a passage from the letter to de Volder of 30 June 1704, Leibniz writes:

> From the fact that mathematical body cannot be resolved into primary constituents, it may be inferred that it is certainly not real, but something mental, designating nothing other than the possibility of parts, not something actual. Indeed, a mathematical line is like an arithmetical unity, and in both cases the parts are only possible and absolutely indefinite. And a line is no more an aggregate of the lines into which it can be cut up, than a unity is an

[44] LDV 338–39.

[45] LDV 320–23.

> aggregate of the fractions into which it can be broken up. … But in real things, namely bodies, the parts are not indefinite (as they are in space, a mental thing), but are actually assigned in a certain way, in accordance with the divisions and subdivisions that nature actually institutes according to different motions. And, although these divisions proceed to infinity, nonetheless, they all result from certain primary constituents, i.e., from real unities, though infinite in number.[46]

In ideal continuous extension the parts are possible, while in real things, they are actual. In this way we can use the ideal continuous extension to represent—in the mind, of course—bodies outside of us by choosing to divide it up in one way rather than another.

This is relatively straightforward: while real concrete bodies are actually divided in one particular way (though they are divided to infinity), geometrical continuous extension contains all possible divisions, in a sense. That is, it is potentially divisible any way in which we would like. And this gives us another way to understand why real bodies cannot be continuously extended. Continuous extension contains *all possible* divisions of a body: but in real bodies, there must be *some one* division, even if the division goes to infinity.

But this captures only part of what Leibniz thinks is interesting and important about geometrical extension. Though geometrical extension contains in this way all possible ways of dividing actual extended things, Leibniz claims that it also puts certain constraints on the world of real extended bodies. In the reply to Foucher from which I quoted earlier, Leibniz wrote:

> However, number and line are not *chimerical* things, even though there is no such composition [i.e., even though they are not composed of parts, as real material things are], for they are relations that contain eternal truths, by which the phenomena of nature are ruled.[47]

This view is echoed in later writings as well. In the 1702 "Reponse aux reflexions," he wrote:

> Although mathematical thinking is ideal, this does not thereby diminish its utility, because actual things cannot escape its rules.[48]

And in the letter to the Electress Sophie from 31 October 1705, from which I quoted earlier:

> [Geometrical space] is the ground of the relation of the order of things … insofar as we conceive of them existing simultaneously.[49]

In this way geometrical extension not only contains everything possible in the concrete world, but also *constrains* it in a way, by imposing certain laws or eternal truths on reality.

---

[46] LDV 302–3.

[47] GP IV 491–92 (AG 146–7).

[48] GP IV 569 (L583).

[49] GP VII 564.

But how? What I think Leibniz has in mind is this. When we take the indeterminate geometrical continuum, and break it into parts, those parts have different magnitudes that are related to one another, and have relations of contiguity or non-contiguity, as well as distance, shape, congruity, and so on, all of the relations that they have by virtue of being geometrical objects, in the case of the actual division of the geometrical space. (One could say the same about arithmetic properties, in case of the division of the unit into fractions.) These “eternal truths” of geometry are, as it were, built into the structure of this mathematical continuum, for Leibniz: the mathematical continuum brings with it the eternal truths of geometry insofar as they are connected with the relations among the different possible way of dividing it into parts.[50] In this way one can say that geometrical continuous extension embodies relations of magnitude and order among the various possible parts that consist in the different ways of breaking it up.

How does the ideal impose itself and its laws on the concrete? Leibniz is a bit unclear about this. But I imagine that the story goes something like this. As I noted earlier, the realm of the ideal is connected with the realm of ideas in our minds, but not only in our minds: as the letter to the Electress Sophie suggests, the ideas are not just in our minds but in God’s as well. The ideal notions of geometry are reflections of ideas in the divine intellect. And as such, they are the models for possible bodies in the possible worlds among which God chooses. For that reason the laws that pertain to things in the ideal world must be observed in the real world as well.

Actually, the story has to be a bit more complicated than that. Because of the infinite division of bodies, bodies in the real world don’t *exactly* instantiate geometrical figures, and thus are not *exactly* governed by geometrical laws, but only to as close an approximation as one would like. Leibniz writes in the important *Specimen inventorum* of 1688 (?):

> Indeed, even though this may seem paradoxical, it must be realized that the notion of extension is not as transparent as is commonly believed. For from the fact that no body is so very small that it is not actually divided into parts excited by different motions, it follows that no determinate shape can be assigned to any body, nor is a precisely straight line, or circle, or any other assignable shape of any body found in the nature of things, although certain rules are observed by nature even in its deviation from an infinite series. Thus shape involves something imaginary, and no other sword can sever the knots we tie for ourselves by misunderstanding the composition of the continuum.[51]

There is a similar passage in the ‘Primae veritates’ of 1689 (?):

> There is no determinate shape in actual things, for none can be appropriate for an infinite number of impressions. And so neither a circle, nor an ellipse, nor any other line we can define exists except in the intellect, nor do lines exist before they are drawn, nor parts before they are separated off.[52]

---

[50] Here I am agreeing with Vincenzo De Risi’s claim that for Leibniz geometrical space is a genuine mathematical structure. See De Risi (forthcoming).

[51] A VI, 4, 1622 (RA 315).

[52] A VI, 4, 1648 (AG 34).

Leibniz's view here is closely connected with his metaphysics of body, and the claim that any finite body is made up of corporeal substances, conceived of as smaller animals, bodies and souls, which, in turn, contain corporeal substances, tiny animals ad infinitum. As a consequence, whenever we look at any boundary of any body with sufficient magnification, it will appear irregular to us, the result of tinier bodies that come together in irregular ways. Of course, any such surface at a given degree of magnification can be fit to some geometrical shape, complicated though it might be. But if we were to look at that same surface at a higher magnification, then new irregularities would appear. And since bodies are divided to infinity in this way, the irregularities will go to infinity as well. Take a bronze sphere. Its mathematical properties can be approximated by those of a geometrical sphere. Or, if that's not good enough, we can examine the sphere under a microscope, and fit it to a more complex irregular geometrical solid. Or if that's not good enough, we can examine it with a more powerful microscope, and fit it to a geometrical solid more complicated still. In this way, we can use a geometrical solid from the ideal geometrical world to approximate the real world as closely as we would like. And so, geometry as it was known in Leibniz's day would be incapable of capturing the true shape of any body in nature. And thus, he argues, the imposition of geometrical shapes, cubes, spheres, whatever, bounded in straight and curved lines, in plane and curved surfaces, is just an imposition of ideal geometrical notions onto a much more complex world. And so he wrote to the Electress Sophie in 1705:

> It is our imperfection and the defects of our senses which makes us conceive of physical things as mathematical entities …. And one can demonstrate that there is no line or shape in nature that has the properties of a straight or circular line or of any other thing whose definition a finite mind can comprehend, or that retains it uniformly for the least time or space. … However, the eternal truths grounded on limited mathematical ideas don't fail to be of use to us in practice, to the extent to which it is permissible to abstract from inequalities too small to be able to cause errors that are large in relation to the end at hand ….[53]

In this way one may say that the geometrical extensionality of bodies is, in a way, the result of our imperfect senses which impose geometrical concepts onto bodies which are, in their real nature, quite something different and which don't fit them exactly.[54] And so, the eternal truths that govern geometrical objects in the

[53] Leibniz to Sophie, 1 October 1705, GP VII 563–4. There is a good discussion of this passage in Hartz and Cover (1988, p. 501). Although I would claim that Leibniz's metaphysics of body and the ultimate make-up of substance is somewhat different when he wrote this letter than it was earlier in the 1680s and early 1690s, the view expressed in the passage quoted is very much continuous with the earlier period.

[54] Some commentators have been tempted to read the no-exact-shape argument as an attempt to establish the claim that the world is made up of non-extended simple substances, and that the extension of bodies is an illusion in a strong sense. See, e.g., Adams (1994, pp. 229–32) and Sleigh (1990, pp. 112–14). But I think that it is more plausible to see Leibniz's intention here to point out the difference between what Sellars has called the manifest view of the world, the world as it appears to us, bodies with real geometrical shapes, and the scientific image of the world, bodies of infinite complexity, beyond our power to grasp in sense. See the excellent discussion of their views in Levey (2005, pp. 84–92).

ideal realm apply to real bodies in nature only approximately, though to any degree of accuracy that one would like.

Let me try to pull these discussions together. For Leibniz, then, there are a number of senses in which the concrete world of bodies is not geometrical. In its nature, the world of bodies is constituted by force, active and passive, and not by geometrical extension, as it is for the Cartesians. In this way Leibniz rejects the Cartesian metaphysic of body on which bodies are the objects of geometry made real. But even more surprisingly, *Leibniz seems to deny that bodies share a mathematical structure with geometrical objects.* For Leibniz, the ideal world of geometry and the concrete world of bodies are fundamentally distinct. The world of geometry is continuous, objects whose parts are not defined and indistinct. In the world of bodies, on the other hand, things are made up of well-defined smaller parts, parts that go to infinity, in fact: as a matter of metaphysical necessity, concrete bodies in the real world are not *continua* but *contigua*. And because of their infinite complexity, he argues, in the strict sense, concrete bodies don't have geometrical shapes. When we look more closely at what appears to be a sphere or a cone or a pyramid in nature, we find that its boundaries are not straight lines, but surfaces of infinite complexity of the sort that go beyond the bounds of geometry, at least as it was understood by Leibniz and his contemporaries. Geometry is thus *fundamentally distinct* from the concrete world of bodies.

So, does geometry apply to Leibnizian bodies? Yes and no. Bodies are not, in their nature, mathematical: they are composed of matter and form, primitive passive and active force. But even so, it seems as if geometrical extension applies to bodies in a relatively straightforward way: bodies are extended insofar as their passive primitive force (impenetrability) gives rise to structures that instantiate the relations treated in pure geometry, at least approximately. Bodies are extended *not* insofar as extension is in them, in any real sense, but insofar as geometrical relations are (approximately) true of them.

So far I have bracketed Leibniz's fundamental metaphysics and emphasized his reflections on the relation between body and extension. But how does the question of the extension of body relate to his fundamental metaphysics?

I have discussed at some length what Leibniz thinks about the notion of extension and the way in which it is connected with a certain account of body and its make-up. For Leibniz, at least in the texts written in the 1680s and 1690s, body is grounded in corporeal substance, and corporeal substance is, in turn, grounded in the notions of form and matter, which are, in turn, understood in terms of active and passive force. And so, it is fair to say, in these texts, at least, the whole edifice of the physical world rests on the notions of force, active and passive.[55]

[55] The claim that Leibniz's metaphysics in the so-called "middle years," the 1680s and most of the 1690s was grounded in corporeal substance, and not in monads is discussed and argued at great length in Garber (2009). Among many problems with the view is the relation between corporeal substance regarded as the union of active and passive force and corporeal substance regarded as living individuals with souls (form) and bodies (matter). These two models for the corporeal substance are not obviously compatible with one another. Even so, it seems clear that Leibniz held

Now, this conception of body first enters into Leibniz's thought in 1678 or 1679 or so, and marks the beginning of what we might call Leibniz's mature thought. There is every reason to believe that in its essentials, this conception of body and the physical world remains pretty constant from then to the end of his life. And at that time, at least from the late 1670s to sometime in the mid-1690s or so, there is every reason to think that at least as regards the physical world, there is no deeper metaphysics of substance at work. That is, there is every reason to believe that he thought of active and passive forces not only as the ground-level *physical* realities, but as the ultimate *metaphysical* realities that ground the created world. A correspondent, Jacques L'Enfant wrote Leibniz with the following remark on 7 November 1693:

> The whole question is thus to know if the force to act in bodies is in matter something distinct and independent of everything else that one conceives there. Without that, this force cannot be its essence, and will remain the result of some primitive quality or another.[56]

Leibniz replies:

> And since everything that one conceives in substances reduces to their actions and passion and to the dispositions that they have for this effect, I don't see how one can find there anything more primitive than the principle of all of this, that is to say, than force.[57]

And, as he wrote even more clearly to Bossuet from 2 July 1694, "I find nothing so intelligible as force."[58] If the active and passive forces that ground body in the physical world also ground Leibniz's fundamental metaphysics of substance, then it would seem as if the reality of extended body is assured, at least if we understand extended body in the way Leibniz does, the way we have discussed it above.

But this fundamental metaphysics will change sometime in the mid- or late 1690s, when Leibniz introduces monads as the new metaphysical foundation of everything.[59] For our purposes here it doesn't really matter exactly when things change. But at a certain point Leibniz seems to add a new sub-basement to his metaphysical edifice, and grounds everything in monads, non-extended and mind-like. As Leibniz explained his view to de Volder in a classic passage:

> Indeed, considering the matter carefully, it should be said that there is nothing in things except simple substances and in them perception and appetite. Moreover, matter and motion are not so much substances or things as the phenomena of perceivers, the reality of

(Footnote 55 continued)
both conceptions during the middle years, and that he thought that they were alternative developments of the same basic metaphysics. Again, see Garber (2009) for a fuller discussion of these issues.

[56] A II, 2, 751.

[57] Leibniz to L'Enfant, 25 November/5 December 1693, A II, 2, 753. It is interesting to observe that even though L'Enfant had written Leibniz about *body*, his reply is quite clearly about *substance*.

[58] Leibniz to Bossuet, 2/12 July 1694, A I, 10, 143–4 (Leibniz (1997, p. 30)).

[59] On the introduction of monads into Leibniz's thought, see Garber (2009, Chap. 8) and Garber (forthcoming).

> which is located in the harmony of perceivers with themselves (at different times) and with other perceivers.[60]

But what sense can we make of the extension of bodies in this new context?

One might think that with the introduction of monads, the relation between body and geometry might change in a fundamental way. In a recent book, *Geometry and Monadology: Leibniz's Analysis Situs and Philosophy of Space*, Vincenzo De Risi has made a very ingenious suggestion about how Leibniz might ground a world of extended bodies directly in the metaphysical level of monads.

Basic to De Risi's account is Leibniz's program for *analysis situs*. Leibniz wanted to reform geometry, and ground it on the foundation of *situs* (situation). As part of that project, as De Risi understands it, Leibniz wants to ground the notion of extension in geometry on *situs*. Here is his summary:

> Leibniz understands that extension is not created by any amount of unextended elements (because an aggregate of unextended elements is unextended as well), nor does it consist of (minimal) atoms of extension, or infinitesimals having an indeterminate size.… Extension is generated only through the mutual relations between the elements that make it up. … Space turns out for him to be a set of relations between unextended (but situated) elements.[61]

In his last years Leibniz successfully defined space as the locus of all points, where points were understood as "absolute situation." In this way, space is *constituted* by points, that is to say, *generated* by points and their relations among one another, though it isn't *composed* of points.[62]

De Risi then tries to use this mathematical project to understand how Leibniz might be thought to have grounded the extension of bodies in the relational properties of non-extended monads. His strategy is this. On De Risi's view, phenomena in monads express other monads.[63] But "all properties distinguishing one monad from the other, i.e. the individuating properties, are relational properties."[64] It is the representation of these non-geometrical relations *as* situation *within* the phenomena of the monad that gives rise to extension.[65] Insofar as the *analysis situs* teaches us that situation gives rise to geometrical space, he concludes that "a simple substance can be represented as an extended body in space."[66] De Risi draws a remarkable conclusion from this: "Every phenomenon expressing a noumenon expresses (non-situated) monadic relations as situational relations. Thus, the extension originating from this set of situational relations … is essential to each phenomenon—each phenomenon is situated and extended…."[67] In essence, on De

[60] Leibniz to de Volder, 30 June 1704, LDV 306–7.

[61] De Risi (2007, p. 174).

[62] De Risi (2007, pp. 173–4).

[63] De Risi (2007, p. 320).

[64] De Risi (2007, p. 321).

[65] De Risi (2007, p. 323).

[66] De Risi (2007, p. 325; cf. pp. 324, 383–4).

[67] De Risi (2007, p. 341).

Risi's view, if I understand it, the relations among monads can give them *situs*, situation, from which extension can be generated by way of the *analysis situs*. There is, of course, a lot of detail to be filled in, but in outline, this seems to be what De Risi has in mind.

De Risi's view is very interesting, but ultimately I'm somewhat skeptical. The main difficulty is textual: the case is based largely on philosophical argument and, to a certain extent, on reading Kantian ideas back into Leibniz's texts. There is no place where Leibniz actually says what De Risi wants him to say. I should point out that De Risi is completely explicit about his mode of argument. He admits a number of times that Leibniz's own statements are often sketchy, hesitant, and confused, indeed to the point that we must "abandon textual analysis … and try instead to answer the question from a theoretical point of view…."[68] Indeed, he often appeals to Leibniz's later readers, particularly Wolff and Kant to make his points.

De Risi's proposal is, in a way, the most rigorous attempt to work out what many commentators have tried to do, show how extended bodies can be grounded directly in the world of non-extended monads. But in Leibniz's own texts, one can find a rather different strategy.

It is quite striking how, even after the appearance of monads in Leibniz's metaphysics, the catalogue of forces, active and passive, primitive and derivative, persist. Perhaps unsurprisingly, in different texts we can find different conceptions of how we are to understand these forces in the new world of monads. In the correspondence with de Volder, in the letter of 20 June 1703 for example Leibniz writes:

> Properly and rigorously speaking, perhaps one will not say that the primitive entelechy impels the mass of its body, but only that it is joined with the primitive passive power that it completes, i.e. with which it constitutes a monad.[69]

Here the primitive forces, both active and passive seem to be in the monad. Later in the same letter he writes:

> I therefore distinguish: (1) the primitive entelechy, i.e. the soul; (2) matter, namely, primary matter, i.e., primitive passive power; (3) the monad completed by these two things; (4) the mass, i.e. the secondary matter, i.e. the organic machine, for which innumerable subordinate monads come together…[70]

In this way, bodies seem to be considered as aggregates of monads. And it is in those aggregates that Leibniz locates derivative forces:

> But in the phenomena, i.e. in the resulting aggregate, everything is indeed explained mechanically, and masses are understood to impel one another. And in these phenomena,

---

[68] De Risi (2007, p. 399; cf. pp. 396, 401).

[69] LDV 260–61. I have altered Paul Lodge's translation here slightly, substituting "*the* passive power" for "*a* passive power".

[70] LDV 264–65.

> nothing is needed except the consideration of derivative forces, once it is agreed where they result from, namely the phenomena of aggregates from the reality of monads.[71]

As he summarizes the view later in the same letter:

> I regard the substance itself, endowed with primitive active and passive power, like the *I* or something similar, as the indivisible, i.e. perfect, monad, not those derivative forces that are continually found to be one way then another. … The forces that arise from mass and speed are derivative and belong to aggregates, i.e. phenomena.[72]

Extension is grounded in passive forces. But insofar as the derivative forces are phenomenal and, for Leibniz, thus less than fully real, so will extension. Leibniz wrote in a draft of a letter to de Volder in 1706:

> I hold that the primitive or derivative force that is conceived of in extension and bulk is not a thing outside perceivers but a phenomenon, just like extension itself, bulk, and motion, which are no more things than an image in a mirror or a rainbow in a cloud.[73]

In this way, one can say that just as primitive and derivative active and passive forces continue in the new context of monads, so does the earlier view of extension as grounded in passive force suitably altered to fit the new context.

But the view continues to evolve. In the exchange with de Volder, primitive forces seem to be located in the monads, while derivative forces are in the phenomena, in aggregates of monads. However, in the correspondence with Des Bosses Leibniz tries out a different idea. In a letter to Des Bosses from 15 February 1712, Leibniz turns to the question of corporeal substance. This is the letter where Leibniz introduces to Des Bosses his famous *vinculum substantiale*, the substantial bond by which genuine corporeal substances are to be formed from collections of monads. And almost immediately, once again, the question of force arises. He writes:

> … from the union of the passive powers of monads there in fact arises primary matter, which is to say, that which is required for extension and antitypy, or for diffusion and resistance. From the union of monadic entelechies, on the other hand, there arises substantial form; but that which can be generated in this way, can also be destroyed and will be destroyed with the cessation of the union, unless it is miraculously preserved by God. However, such a form then will not be a soul, which is a simple and indivisible substance.[74]

---

[71] LDV 260–61. Here bodies (and their derivative forces) are taken to be phenomenal in the sense that their unity as aggregates is contributed by the mind: individual monads are substances, but it is the mind that makes an aggregate of monads a thing by considering them together. In general, for Leibniz to say that some being is phenomenal is to say that it is in some way or another constituted by a perceiver. For the different senses of the phenomenal and the different ways in which Leibniz understands bodies and their forces to be phenomenal, see Garber (2009, Chap. 7).

[72] LDV 262–63.

[73] LDV 336–37. In this context, though extension and the derivative forces are clearly dependent on perceivers, it is not clear that Leibniz means to say that extension and the derivative forces are phenomenal in exactly the same sense that he meant them to be in the passages above from the letter of 20 June 1703. On this see again Garber (2009, Chap. 7).

[74] LDB 224–25.

As in the earlier passages from the letters with de Volder, it seems as if monads are composed of primitive active and passive force. But here there seems to be a primitive active and passive force in the corporeal substance as well, forces that arise from the primitive forces in the monads. And from the primitive passive force in the corporeal substance, Leibniz tells us, there arise, in turn, antitypy, that is, impenetrability, and extension.

But this view evolves as well. In his letter to Des Bosses from 29 May 1716, not long before his death, Leibniz wrote:

> I do not say that there is a mediating bond between form and matter but, rather, that the substantial form itself of the composite and primary matter taken in the Scholastic sense, that is, primitive active and passive power, belongs to that bond, as the essence of the composite.[75]

Here the view is that the primitive forces belong to the bond itself: the substantial bond is now the seat of primitive active and passive force. But Leibniz then continues in a remarkable way:

> However, this substantial bond is naturally, and not essentially, a bond. For it requires monads but does not involve them essentially, since it can exist without monads and monads without it.[76]

Though Leibniz says here that the bond "requires monads," it also has a kind of independence from them insofar as it can exist without them. And, in a passage later in the same letter he seems to indicate that the substance is similarly independent of the monads:

> Composite substance does not formally consist in monads and their subordination, for then it would be a mere aggregate or a being *per accidens.* Rather, it consists in primitive active and passive force, from which arise the qualities and the actions and passions of the composite which are discovered by the senses, if they are assumed to be more than phenomena.[77]

Here Leibniz seems to suggest that the monads play no role at all in the corporeal substance. And we are back to the view we saw earlier in the *Specimen dynamicum*.

In addition to the grounding of body in force, in Leibniz's later writings we can also find the earlier view that body is actually divided to infinity, corporeal substances in corporeal substances without end, the view that underlies Leibniz's idea that bodies are discontinuous and of infinite complexity. And so Leibniz writes in a letter to de Volder on 30 June 1704:

> …[A] mathematical line is like an arithmetical unit, and in both cases the parts are only possible and absolutely indefinite…But in real things, namely bodies, the parts are not indefinite … but are actually assigned in a certain way, in accordance with the divisions and subdivisions that nature actually institutes according to different motions. And although these divisions proceed to infinity, nonetheless they all result from certain primary

[75] LDB 366–67.

[76] Ibid.

[77] LDB 370–71. I have slightly altered the translation by Look and Rutherford.

> constituents, i.e. from real unities, though infinite in number… [P]henomena can always be divided into lesser phenomena that could appear to other more subtle animals, and the smallest phenomena will never be reached.[78]

Similarly, Leibniz writes to Des Bosses on 11 March 1706:

> When I say there is no part of matter that does not contain monads, I illustrate this with the example of the human body or that of some other animal, any of whose solid and fluid parts contain in themselves in turn other animals and plants. And this, I think, must be said again of any part of these living things, and so on to infinity.[79]

This view can also be found in the so-called "Monadology" of 1714, a summary outline of Leibniz's monadological metaphysics:

> Each portion of matter can be conceived as a garden full of plants, and as a pond full of fish. But each branch of a plant, each limb of an animal, each drop of its humors, is still another such garden or pond….Thus there is nothing fallow, sterile, or dead in the universe, no chaos and no confusion except in appearance, almost as it looks in a pond at a distance, where we might see the confused and, so to speak, teeming motion of the fish in the pond, without discerning the fish themselves. Thus we see that each living body has a dominant entelechy, which in the animal is the soul; but the limbs of this living body are full of other living beings, plants, animals, each of which also has its entelechy, or its dominant soul.[80]

In this way the later monadological metaphysics seems continuous with another aspect of Leibniz's earlier view of body.

There is a sense, I think, in which Leibniz's reflections about physical body and its extension have a kind of philosophical autonomy from his views on ultimate reality, ultimately, the realm of monads. There is a sense in which Leibniz's philosophy is modular. It can be regarded as a series of positions in a variety of domains—physics, metaphysics, logic, theology, etc.—that cohere with one another to form a unified and harmonious whole. At the same time, each domain can be regarded independently, understood and accepted or rejected without having to accept or reject the entire package. This is true, I think, of Leibniz's account of body and the physics of body, which can be considered in independence of his fundamental metaphysics. Furthermore, I think it has a kind of stability that is somewhat independent of the fundamental metaphysics: as the latter changes with the introduction of monads at the bottom level, the theory of body can and does remain largely unchanged, in a way.[81]

In these later texts that we have been surveying, Leibniz doesn't talk as much about the relation between geometry and extended body as he did in the earlier texts. Perhaps with the new monadological metaphysics his interest had shifted

---

[78] LDV 302–303. I have slightly altered Lodge's translation here, substituting "unit" for "unity" in the first sentence quoted.

[79] LDB 34–35. I presume here that the monads that he finds in every part of matter are the souls of the plants and animals (corporeal substances) contained there.

[80] Monadology §§ 67, 69–70 (AG 222).

[81] For a fuller development of the idea of Leibniz's philosophy as modular, albeit in a rather different context, see Garber (2014).

away from differentiating his own position from that of the Cartesians and toward working out the details of his new metaphysical picture. But even so, it is interesting how the essentials of Leibniz's earlier account of body and the grounds of its extension persists throughout his later years. Bodies continue to be characterized in terms of force, active and passive, primitive and derivative, even after monads are introduced, and along with that conception of body the account of how extension arises from the passive forces in bodies. To be sure, there are differences as well: the various claims Leibniz makes for the greater reality of the level of monads and the phenomenality in some sense of the level of force, at least derivative and perhaps even primitive force have consequences for the reality of extension. There is much work that needs to be done to understand what becomes of Leibniz's catalogue of forces in his later years. But even so, it seems to me reasonable to suppose that at least the main lines of Leibniz's earlier account of body and its relation to the ideal world of geometry persists into his later years.

References

Adams, R. M. (1994). *Leibniz: Determinist, theist, idealist*. New York: Oxford University Press.

Aristotle (1984). In J. Barnes (Ed), *The complete works of Aristotle* (Vol. 2.). Princeton: Princeton University Press.

De Risi, V. (2007). *Geometry and monadology: Leibniz's analysis situs and philosophy of space*. Basel, Boston and Berlin: Birkhäuser.

De Risi, V. (forthcoming). *Leibniz on the parallel postulate and the foundations of geometry*.

Garber, D. (1992). *Descartes' metaphysical physics*. Chicago: University of Chicago Press.

Garber, D. (2009). *Leibniz: Body, substance, monad*. Oxford: Oxford University Press.

Garber, D. (2014). Monads and theodicy: Reading Leibniz. In L. Jorgensen & S. Newlands (Eds.), *New essays in Leibniz's theodicy* (pp. 218–232). Oxford: Oxford University Press.

Garber, D. (forthcoming). Monads on my mind. In A. Nita (Ed.), *Leibniz's metaphysics and the adoption of substantial forms*. Dordrecht: Springer.

Hartz, G. A., & Cover, J. A. (1988). Space and time in the Leibnizian metaphysic. *Noûs, 22*, 493–519.

Hobbes, T. (1655). *Elementorum philosophiae secio prima de corpore*. London: Andrew Crooke.

Hobbes, T. (1656). *Elements of philosophy the first section concerning body*. London: Andrew Crooke.

Leibniz, G. W. (1849–1863). In C. I. Gerhardt (Ed.), *Leibnizens mathematische Schriften* (Vol. 7). Berlin: A. Asher (References include volume and page. So 'GM VII 80' is volume 7 p. 80. Abbreviated 'GM').

Leibniz, G. W. (1875–1890). In C. I. Gerhardt (Ed.), *Die philosophischen Schriften* (Vol. 7). Berlin: Weidmann (Abbreviated 'GP.' References include volume and page. So 'GP VII 80' is volume 7 p. 80).

Leibniz, G. W. (1923-). *Sämtliche Schriften und Briefe*. In Deutsche Akademie der Wissenschaften zu Berlin (Ed.), Berlin: Akademie-Verlag (Abbreviated 'A'. References include series, volume, and page. So 'A VI, 4, 1394' is series 6, volume 4, p. 1394).

Leibniz, G. W. (1967). *The Leibniz-Arnauld correspondence* (H. T. Mason, Ed. & Trans.). Manchester: Manchester University Press (Abbreviated 'M').

Leibniz, G. W. (1976). *Philosophical papers and letters* (2nd ed) (L. E. Loemker, Ed. & Trans.). Dordrecht: Reidel. (Abbreviated 'L').

Leibniz, G. W. (1982). *Specimen dynamicum* (H. G. Dosch, G. W. Most, E. Rudolph, & J. Aichelin, Eds. & Trans.). Hamburg: Meiner.

Leibniz, G. W. (1989). *Philosophical essays* (R. Ariew & D. Garber, Eds. & Trans.). Indianapolis: Hackett (Abbreviated 'AG').

Leibniz, G. W. (1997). *Leibniz's 'New system' and associated contemporary texts* (R. S. Woolhouse & R. Francks, Eds. & Trans.). Oxford: Oxford University Press.

Leibniz, G. W. (2001). *The labyrinth of the continuum: Writings on the continuum problem, 1672–1686* (R. Arthur, Ed. & Trans.). New Haven: Yale University Press (Abbreviated 'RA').

Leibniz, G. W. & B. Des Bosses (2007). *The Leibniz-Des Bosses correspondence* (B. Look & D. Rutherford, Eds. & Trans.). New Haven: Yale University Press (Abbreviated 'LDB').

Leibniz, G. W. & B. de Volder (2013). *The Leibniz-de Volder correspondence* (P. Lodge, Ed. & Trans.). New Haven: Yale University Press (Abbreviated 'LDV').

Levey, S. (2005). Leibniz on precise shapes and the corporeal world. In D. Rutherford & J. A. Cover (Eds.), *Leibniz: Nature and freedom* (pp. 69–94). Oxford: Oxford University Press.

Sleigh, R. C. (1990). *Leibniz and Arnauld: A commentary on their correspondence*. New Haven: Yale University Press.

Hume's Skepticism and Inductivism Concerning Space and Geometry

Graciela De Pierris

1 Introduction: The Inexactitude of Geometry

I have argued elsewhere that Hume's views on space and geometry in Book One, Part Two of *A Treatise of Human Nature* are rooted in his radically empiricist model of apprehension and ultimate evidence.[1] This epistemological model ultimately relies on phenomenologically given particular sensory images. Diagrams in geometry are always regarded as themselves particular sensory images, thus they cannot be taken as representatives of ideal geometrical objects. Guided by this conception, Hume offers a radically skeptical view of geometry (the study of continuous quantity) according to which it cannot be a fully exact and certain science. Only arithmetic and algebra (the sciences of discrete quantity) attain complete exactitude and certainty.

Hume's discussion of space and geometry is centrally concerned with arguing against the possibility of infinite divisibility—showing that our ideas of space and time (and space and time themselves) are not phenomenologically infinitely divisible. He offers arguments based on principles about the infinite and argues that to suppose that space and time are infinitely divisible is absurd, impossible, and

G. De Pierris (✉)
Department of Philosophy, Stanford University, Building 90, 450 Serra mall,
Stanford, CA 94305-2155, USA
e-mail: graciela@stanford.edu

[1] See De Pierris (2012, 2013).

V. De Risi (ed.), *Mathematizing Space*, Trends in the History of Science,
DOI 10.1007/978-3-319-12102-4_11

contradictory.[2] The most notorious such argument, presented at T 1.2.2.2 (SBN 29), traces back to Zeno's metrical paradox of extension. If any finite interval of space (or time) is infinitely divisible, then it must consist of an infinite number of ultimate parts. But these parts, when added together, must then result in an extension that is infinitely great, contrary to the supposition that it was a finite interval with which we started. From the point of view of pure mathematics, there is an obvious objection to this argument: for, if we divide a finite interval into an infinite sequence of *decreasing* finite parts, their sum when added together (as a limit) can still be finite (as $1/2 + 1/4 + 1/8 + \cdots$ approaches 1 as its limit). I have argued elsewhere more extensively than I shall have the occasion to do here that, when we place Hume's argument within his epistemological model, we can appreciate that obvious and correct criticisms issuing from a purely mathematical conception of geometry and the infinite are not relevant to Hume's discussion.[3]

In order to appreciate how Hume can avoid these criticisms, we should pay attention to two interconnected tenets of Hume's view, which are a direct result of his empiricist phenomenological model of geometry and the infinite. The first is his acknowledgement that our perception of a homogeneous space can give us a phenomenological appearance of continuity. The second is that our complex impression of space can be perceptually presented to us only as finitely divisible. Hume invokes, on the one hand, the "confounding" of indivisible minima in the complex phenomenological appearance of continuity and, on the other, the fact that in undertaking perceptual divisions or diminutions of space we always reach a visual or tactual threshold beyond which our sensory experience reduces to nothing. What we experience at the last stage of divisions or diminutions—at the threshold before annihilation—is a simple, perceptually indivisible, minimum or unextended "point."

Hume devotes a large portion of T 1.2.4 (SBN 39–53) to answering the objection that demonstrations in geometry prove the infinite divisibility of space. The outcome of this discussion is that geometry, unlike arithmetic, is not a perfectly exact science, because demonstrations in geometry are not perfectly exact (T 1.2.4.17/SBN 44–45): "[N]one of these demonstrations can have sufficient weight to establish such a principle, as this of infinite divisibility; and that because with

---

[2] Hume claims that his arguments against infinite divisibility (including his "experiments" discussed below) are demonstrative (T 1.2.2.6/SBN 31): "I doubt not but it will readily be allow'd by the most obstinate defender of the doctrine of infinite divisibility, that these arguments are difficulties … But here we may observe, that nothing can be more absurd, than this custom of calling a *difficulty* what pretends to be a *demonstration*, and endeavouring by that means to elude its force and evidence. 'Tis not in demonstrations as in probabilities, that difficulties can take place, and one argument counter-ballance another, and diminish its authority. A demonstration, if just, admits of no opposite difficulty; and if not just, 'tis a mere sophism, and consequently can never be a difficulty. 'Tis either irresistible, or has no manner of force." All citations of *A Treatise of Human Nature* (abbreviated T), are from Hume (2000), and thus include the book, part, section, and paragraph numbers; this is followed by a reference to the corresponding page number in Hume (1978), abbreviated SBN. All citations of *An Enquiry concerning Human Understanding* (abbreviated EHU) are from Hume (1999), which includes section and paragraph numbers; this is followed by a reference to the page number in Hume (1975), abbreviated SBN.

[3] See again note 1 above.

regard to such minute objects, they are not properly demonstrations, being built on ideas, which are not exact, and maxims, which are not precisely true. When geometry decides any thing concerning the proportions of quantity, we ought not to look for the utmost *precision* and exactness. None of its proofs extend so far. It takes the dimensions and proportions of figures justly, but roughly, and with some liberty. Its errors are never considerable, nor wou'd it err at all, did it not aspire to such an absolute perfection." Geometers cannot establish the exact number of minima in the figures they compare (T 1.2.4.19/SBN 45): "For as the points, which enter into the composition of any line or surface, whether perceiv'd by the sight or touch, are so minute and so confounded with each other, that 'tis utterly impossible for the mind to compute their number, such a computation will never afford us a standard, by which we may judge of proportions."

Geometers could attain the absolute perfection to which they aspire—an ideal exactness concerning proofs of dimensions and proportions of figures—only on the basis of determining an exact (finite) number of minima in each of figure. However, this is impossible since we cannot reach all the minima simultaneously. Thus, the irremediable inexactitude of geometry issues from the "confounding" of the finite indivisible minima in the complex phenomenological appearance of continuity together with Hume's implicit assumption that in undertaking perceptual divisions of space we can reach only one minimum at a time.

Concerning the impossibility of establishing exact equality, Hume writes (T 1.2.4.24/SBN 48): "We are sensible, that the addition or removal of one of these minute parts, is not discernible either in the appearance or measuring; and as we imagine, that two figures, which were equal before, cannot be equal after this removal or addition, we therefore suppose some imaginary standard of equality, by which the appearances and measuring are exactly corrected, and the figures reduc'd entirely to that proportion. This standard is plainly imaginary. For as the very idea of equality is that of such a particular appearance corrected by juxta-position or a common measure, the notion of any correction beyond what we have instruments and art to make, is a mere fiction of the mind, and useless as well as incomprehensible." Nonetheless, in common life and science, we surrender to this natural fiction and imagine that we have exact measures of equality and proportions for continuous extended magnitudes (ibid.): "But tho' this standard be only imaginary, the fiction however is very natural; nor is any thing more usual, than for the mind to proceed after this manner with any action, even after the reason has ceas'd, which first determin'd it to begin." Proofs of equality based on congruence fail for the same reason (see, e.g., T 1.2.4.21/SBN 46). When placing one figure upon the other, the supposition that we can determine whether *all* their parts correspond to and touch one another is fictitious.

Hume's phenomenological model of apprehension and ultimate evidence leads him also to conclude that the allegedly ideal geometrical notions of a right line (or a curve) and of a plain surface (or a curved one) are inexact. In particular, there is no precise boundary between curves and straight lines: we cannot ascertain in the sensory appearance the order of the minimum points in a line (T 1.2.4.25/SBN 49):

"Nothing is more apparent to the senses, than the distinction betwixt a curve and a right line … But however easily we may form these ideas, 'tis impossible to produce any definition of them, which will fix the precise boundaries betwixt them. When we draw lines upon paper or any continu'd surface, there is a certain order, by which the lines run along from one point to another, that they may produce the entire impression of a curve or right line; but this order is perfectly unknown, and nothing is observ'd but the united appearance. Thus even upon the system of indivisible points, we can only form a distant notion of some unknown standard to these objects. Upon that of infinite divisibility we cannot go even this length; but are reduc'd merely to the general appearance, as the rule by which we determine lines to be either curve or right ones." As a result, we have no perfectly exact notion of straight line at all (T 1.2.4.30/SBN 52): "The original standard of a right line is in reality nothing but a certain general appearance; and 'tis evident right lines may be made to concur with each other, and yet correspond to this standard, tho' corrected by all the means either practicable or imaginable."

In sum, precisely because the minima appear confounded in any phenomenologically given extension, the number of minima cannot be determined. And since geometry is based on nothing but such phenomenological appearances, the geometer can never provide perfectly precise definitions or demonstrate exact equality or proportion of extensions. We could achieve the complete certainty of exact demonstrations only if, *per impossibile*, we could establish an exact number of minima for each of the extensions compared (T 1.2.4.19/SBN 45): "[L]ines or surfaces are equal, when the numbers of points in each are equal; and that as the proportion of the numbers varies, the proportion of the lines and surfaces is also vary'd. But tho' this answer be *just*, as well as obvious; yet I may affirm, that this standard of equality is entirely *useless*, and that it never is from such a comparison we determine objects to be equal or unequal with respect to each other … No one will ever be able to determine by an exact numeration, that an inch has fewer points than a foot, or a foot fewer than an ell or any greater measure; for which reason we seldom or never consider this as the standard of equality or inequality [of extension]." The only exact standard of equality or proportion—a one-to-one correspondence between the simple units composing each of the magnitudes compared—is "useless" when we are concerned with extension, and this shows that geometry cannot be a perfectly exact science but only work with rough approximations. The science of arithmetic is the discipline that provides the "just" ideal standard (a one-to-one correspondence) of exact enumeration of simple parts or discrete units. Indeed, it is precisely because in arithmetic there is no confounded appearance of discrete units that arithmetic, unlike geometry, can successfully apply the standard of equality it provides (see T 1.3.1.5/SBN 71). However, despite his skepticism regarding the exactitude of geometry, Hume still takes the approximate results of geometry to amount to a demonstrative science, that is, to be based on comparisons of intrinsic properties of ideas, not on their external relations. I return to this point in the last section below.

In the present essay I aim to clarify the relationship between our phenomenologically given complex idea of space as a homogeneous continuous image and the simple minima out of which it is composed by examining in more detail than in my previous work some subtleties and difficulties arising from Hume's examples. This examination will illuminate the other issues with which I am concerned here. I shall discuss the character of the *inductive* advance in the study of nature that leads to "new sensible minima": the inductive discovery of new minima with the aid of instruments such as the microscope and the telescope, starting from impressions to which we have access simply with our unaided senses of sight and touch. I shall argue, in particular, that such inductive advancement does not commit Hume to realist assumptions about minima existing independently of our impressions and ideas. I shall also highlight Hume's inductivist conception of the application of geometry to nature in common life and science (outside his radical skeptical standpoint), and discuss the relationship of this conception to his own radical skeptical arguments concerning geometry's inexactitude and lack of complete certainty. I shall conclude that Hume's inductivist conception of the application of geometry to nature—together with his unmitigated skeptical view that geometry as a demonstrative a priori science is intrinsically inexact and never fully certain—show that Hume's commitment to inductivism in common life and science goes well beyond Newton's commitment to inductivism in the *Principia*.[4] For neither Newton's focus on "manifest Effects" or "Phænomena" nor Newton's own inductivism generates doubts concerning the perfect exactitude of the science of geometry or its exact application to all regions of space, whether very small or very large.

2 Hume's "Experiments" Concerning Divisibility

Hume appeals to "experiments" in order to undermine the geometricians' claim that space is infinitely divisible. Immediately after he introduces the notion of a simple minimum, he illustrates this notion with an experiment with images of the imagination (T 1.2.1.3/SBN 27): "'Tis therefore certain, that the imagination reaches a *minimum*, and may raise up to itself an idea, of which it cannot conceive any subdivision, and which cannot be diminish'd without a total annihilation. When you tell me of the thousandth and ten thousandth part of a grain of sand, I have a distinct idea of these numbers and of their different proportions; but the images, which I form in my mind to represent the things themselves, are nothing different from each other, nor inferior to that image, by which I represent the grain of sand itself, which is suppos'd so vastly to exceed them. What consists of parts is distinguishable into them, and what is distinguishable is separable. But whatever we may imagine of the

[4] For the background to my brief discussion in the present essay of Hume's relationship to Newton's scientific methodology, see De Pierris (2006).

thing, the idea of a grain of sand is not distinguishable, nor separable into twenty, much less into a thousand, ten thousand, or an infinite number of different ideas."[5]

The next paragraph offers a further experiment in order again to illustrate the notion of a minimum, but this time concerning impressions of sensation (T 1.2.1.4/ SBN 27–28): "'Tis the same case with the impressions of the senses as with the ideas of the imagination. Put a spot of ink upon paper, fix your eye upon that spot, and retire to such a distance, that at last you lose sight of it; 'tis plain, that the moment before it vanish'd the image or impression was perfectly indivisible. 'Tis not for want of rays of light striking on our eyes, that the minute parts of distant bodies convey not any sensible impression; but because they are remov'd beyond that distance, at which their impressions were reduc'd to a *minimum*, and were incapable of any further diminution. A microscope or telescope, which renders them visible, produces not any new rays of light, but only spreads those, which always flow'd from them; and by that means both gives parts to impressions, which to the naked eye appear simple and uncompounded, and advances to a *minimum*, what was formerly imperceptible."

Hume's explanation of the modification of one's visual field by the use of a microscope or telescope might be read as making a realist point about the existence of minima independently of our apprehension—which are waiting, so to speak, to be discovered. Yet Hume is not interested in the question whether there are minima independently of what an observer can perceive; he is rather concerned with what, at a given time and under specific circumstances, a perceiver apprehends after a series of diminutions or divisions, just before the impression or image is annihilated. In the last quoted text, for example, Hume says that a telescope or microscope "gives parts to *impressions*" (emphasis added)—thus revealing that he focuses on sensory impressions and their modes of presentation rather than independently existing physical referents. Thus, in the spot of ink example, the minima or simples are sensory impressions, which are the ultimate perceptible ingredients into which the complex sensory impression *cannot* be further subdivided without causing the

[5] Notice that this and subsequent experiments of reaching a simple indivisible minimum shed light on Hume's earlier distinction between simple and complex impressions and ideas (introduced at T 1.1.1.2/SBN 2) and on the closely related separability principle, which he renders here as "what consists of parts is distinguishable into them, and what is distinguishable is separable". Hume first introduces the separability principle at T 1.1.3.4 (SBN 10): "Where-ever the imagination perceives a difference among ideas, it can easily produce a separation." At T 1.1.7.3 (SBN 18) he adds the condition of distinguishability: "[W]hatever objects are different are distinguishable, and whatever objects are distinguishable are separable [and vice versa] by the thought and imagination." Separability is also said to be a necessary property of "distinct" ideas (T 1.3.3.3/SBN 79): "[A]ll distinct ideas are separable from each other." The above quoted text from T 1.1.3.4 (SBN 10) occurs in the context of pointing out that the imagination (unlike memory) has perfect "*liberty ... to transpose and change its ideas*," Here Hume makes explicit the close relationship between the simple/complex distinction and the separability principle (T 1.1.3.4/SBN 10): "Nor will this liberty of the fancy appear strange, when we consider, that all our ideas are copy'd from our impressions, and that there are not any two impressions which are perfectly inseparable. Not to mention that this is an evident consequence of the division of ideas into simple and complex. Where-ever the imagination perceives a difference among ideas, it can easily produce a separation."

perception in question simply to vanish. Similarly, in the example of the grain of sand, the minima or simples are ideas, which are the indivisible ingredients of a complex idea of a sandy aggregate.

It is undeniable that Hume uses realist language in both of these texts. In the first he hypothetically refers to the "things themselves." In the second, he writes about removing oneself to such a distance from a spot of ink on paper until one finally loses sight of it, of rays of light flowing from the minute parts of distant bodies, of optical instruments modifying these rays, and so on. When, in realistic language, I gradually remove myself to the threshold distance just before the spot vanishes, what I actually perceive is a series of ever smaller closely resembling impressions until I finally reach one that cannot be further diminished without annihilation. Nonetheless, Hume is not committed to postulating a relation of numerical identity between the impression of an ink spot at one particular time and the slightly smaller (or larger) impression arising at a slightly different time—nor such a relation of identity between their supposed physical referents. On the contrary, all that we are actually presented with, on Hume's account, are two different impressions related to one another by *resemblance*.

The relationship between the original phenomenological appearance of an extended ink spot and the appearance of a simple minimum out of which it is composed is a relationship between two different visual fields presented at two different times. The minimum appears as the last member of a temporal sequence of closely resembling visual appearances of ever-smaller ink spots, ending at the threshold immediately before the appearance vanishes. In this sense, therefore, the minimum is not literally perceived as a simple component in the initial complex (extended) appearance with which we began—nor in any of the succeeding complex (extended) appearances that follow. The relationship between the complex (extended) appearance and the simple (unextended) minimum is not the geometrical relationship between the (extended) whole appearance and the (extended) sub-wholes that are simultaneously perceived in and with it.

There are therefore two senses of "perceptible part" in a perceptible extended whole given at a single time. Any such extended whole consists of *both* the extended (complex) sub-wholes that are perceivable as such at the given time (the left-side and right-side of the initially given ink spot, for example) *and* the simple (unextended) minima whose "confounding" results in a homogeneous appearance of extension at that same time (the darkly colored "points" out of which the ink spot is composed). The latter, however, are not separately perceived *as minima* at this time, for they constitute the appearance of continuous extension only by being "confounded"—and thus by not being individually perceived, and by their *manner* of appearing as co-existing. They only become individually perceived (as minima) in the context of a (finite) temporal sequence of appearances beginning with some initial complex whole (the original spot of ink).[6]

The complex ideas of space and time emerge from the perceived manner or order of arrangement of simple indivisible minima. Whereas the manner of appearance of space is a co-existence of homogeneous simple parts (minima), the manner of appearance of time is a succession of changing (heterogeneous) perceptions which are themselves unchanging (T 1.2.3.7/SBN 35): "As 'tis from the disposition of visible and tangible objects we receive the idea of space, so from the succession of ideas and impressions we form the idea of time, nor is it possible for time alone ever to make its appearance, or be taken notice of by the mind." Hume summarizes his characterization of the visual apprehension of space as follows (T 1.2.3.4/SBN 34): "[W]e may conclude with certainty, that the idea of extension is nothing but a copy of these colour'd points, and of the manner of their appearance." The difference between the relevant relations (coexistence as opposed to succession) marks the difference between space and time. The fact that the phenomenological apprehension of space and time always involves the apprehension of such relations explains Hume's central assumption that the impressions and ideas of space and time are, respectively, always complex and thus divisible. The copy principle here works in tandem with the separability principle. Neither space nor time reduces to any of the simple units of which it is composed. Nor does either space or time reduce to an additional simple impression on a par with the simple components.[7]

---

[6] Given these two notions of perceptible part, there are two different senses in which space can appear to have parts of equal or different "sizes." On the one hand, there are the extended sub-wholes of extension that are perceivable as such in a single given whole appearance. Hume accepts a rough and approximate *geometrical* notion of "size" for such sub-wholes, in so far as they may phenomenologically appear (roughly) as either of the same magnitude or of different magnitudes. (These are only rough appearances, for Hume, since the exact magnitude—the exact number of minima in each of these sub-wholes—cannot be determined). On the other hand, with respect to the simple and indivisible minimum parts whose "confounding" results in all the appearances of extension in question, the situation is completely different. For, despite the fact that a minimum can be determined to be such only after a (finite) temporal process of diminutions or divisions, there is nonetheless a sense in which Hume takes the minima, as I argued in the first article cited in note 1 above, to be all of the same *arithmetical* "size": they are ultimate simple units, which sensibly appear as discrete at the phenomenological threshold immediately before the appearance vanishes.

[7] Hume puts this point very clearly for the case of time (T 1.2.3.10/SBN 36–37): "In order to know whether any objects, which are join'd in impression, be separable in idea, we need only consider, if they be different from each other; in which case, 'tis plain they may be conceiv'd apart. Every thing, that is different, is distinguishable; and every thing, that is distinguishable, may be separated, according to the maxims above-explain'd. If on the contrary they be not different, they are not distinguishable; and if they be not distinguishable, they cannot be separated. But this is precisely the case with respect to time, compar'd with our successive perceptions. The idea of time is not deriv'd from a particular impression mix'd up with others, and plainly distinguishable from them, but arises altogether from the manner, in which impressions appear to the mind, without making one of the number. Five notes play'd on a flute give us the impression and idea of time; tho' time be not a sixth impression, which presents itself to the hearing or any other of the senses. Nor is it a sixth impression, which the mind by reflection finds in itself. These five sounds making their appearance in this particular manner, excite no emotion in the mind, nor produce an

Because the ideas of space and time copy impressions of relations among simples (these relations are the manner of arrangement of simples), I characterize the impressions of relations and their corresponding ideas as second-order impressions and ideas. The complex idea of space is an image perceptibly "filled" with colored or tangible but extensionless (not further divisible) simples, arranged in the manner of co-existence. Whereas the perceptible sub-wholes of a given whole of extension are perceived as such simultaneously with the whole, the indivisible simple minima are not themselves perceived *as minima* simultaneously with the whole. Indeed, the characteristic "confounding" of spatial minima in a given whole of extension consists in precisely the fact that they are presented in the manner of co-existence (as simultaneous) without being so presented as minima. Perceptible sub-wholes of a given whole of extension are also related to one another in the manner of co-existence, of course, but this is a first order relation of co-existence between parts that are presented as parts simultaneously with the whole. Nonetheless, they are ultimately constituted, just like the whole itself, by the second-order manner of coexistence of unextended minima.

3 The Inductive Advance to "New Sensible Minima"

Let us return, with these points in mind, to the experiment in which one obtains the presentation of an indivisible (unextended) minimum as the last member of a temporal sequence of closely resembling visual appearances, beginning with an initial appearance of a (homogeneous) extended ink spot and followed by ever smaller such appearances until the minimum is reached. Here, as I have argued, the relationship between a complex (extended) appearance and the simple (unextended) minimum reached at the threshold is not immediately presented in any single phenomenological field; rather, it is a relation between any complex (extended) appearance in the sequence and the final appearance of an indivisible (simple) minimum at the very end of the sequence. All members of the sequence are connected by the relationship of resemblance, but they are also connected by the actual temporal order in which the sequence takes place (large spot followed by slightly smaller spot, followed by slightly smaller spot, and so on). This temporal order

(Footnote 7 continued)
affection of any kind, which being observ'd by it can give rise to a new idea. For *that* is necessary to produce a new idea of reflection, nor can the mind, by revolving over a thousand times all its ideas of sensation, ever extract from them any new original idea, unless nature has so fram'd its faculties, that it feels some new original impression arise from such a contemplation. But here it only takes notice of the *manner*, in which the different sounds make their appearance; and that it may afterwards consider without considering these particular sounds, but may conjoin it with any other objects. The ideas of some objects it certainly must have, nor is it possible for it without these ideas ever to arrive at any conception of time; which since it appears not as any primary distinct impression, can plainly be nothing but different ideas, or impressions, or objects dispos'd in a certain manner, that is, succeeding each other." Note the importance of the explicit appeal to the separability principle at the very beginning of this passage.

involves a sequence of perceived changes—both in the perceived visual fields (the ink spot and its background) and in the concurrent bodily sensations (of moving backwards and forwards in relation to these visual fields)—which perceived sequence, in turn, involves the past observation of regularities or constant conjunctions among these various perceptions. The use of induction is therefore presupposed here.

Hume's discussion of the microscope and the telescope immediately following the example of the ink spot at T 1.2.1.4 (SBN 28) indicates that further empirical inductive advance in the study of nature might always give us "new" minima. We might always increase our capacity to have presentations before the mind that "give parts" to impressions—that is, we obtain *new* impressions closely resembling those which, at an earlier time, were indivisible minima not separable into "parts." In examining the distant stars with a telescope, for example, those which formerly appeared as simple minima (at the threshold of perception) may now appear as complex and extended; and, from a visual field in which no star at all appeared, we might now, as Hume suggests, "advance" to a perceptible "*minimum*, what was formerly imperceptible" (T 1.2.1.4/SBN 28). In both cases, we are comparing perceptible visual fields, which in a particular temporal sequence closely resemble one another. In the case of "advancing" to something that was "formerly imperceptible," in particular, we are adding a new visible point to a background of (closely resembling) visible items (the appearance of surrounding stars) present in both fields.

It is now possible to make sense of the paragraph immediately following Hume's mention of the microscope (T 1.2.1.5/SBN 28): "We may hence discover the error of the common opinion, that the capacity of the mind is limited on both sides, and that 'tis impossible for the imagination to form an adequate idea, of what goes beyond a certain degree of minuteness as well as greatness. Nothing can be more minute, than some ideas, which we form in the fancy; and images, which appear to the senses; since these are ideas and images perfectly simple and indivisible. The only defect of our senses is that they give us disproportion'd images of things, and represent as minute and uncompounded what is really great and compos'd of a vast number of parts. This mistake we are not sensible of; but taking the impressions of those minute parts, which appear to the senses, to be equal or nearly equal to the objects, and finding by reason, that there are other objects vastly more minute, we too hastily conclude, that these are inferior to any idea of our imagination or impression of our senses. This is however certain, that we can form ideas, which shall be no greater than the smallest atom of the animal spirits of an insect a thousand times less than a mite. And we ought rather to conclude, that the difficulty lies in enlarging our conceptions so much as to form a just notion of a mite, or even of an insect a thousand times less than a mite. For in order to form a just notion of these animals, we must have a distinct idea representing every part of them; which, according to the system of infinite divisibility, is utterly impossible, and according to that of indivisible parts or atoms, is extremely difficult, by reason of the vast number and multiplicity of these parts." This passage has (understandably) puzzled many commentators.

First, Hume begins by appearing to contradict his initial claim at T 1.2.1.2 (SBN 26): "'Tis universally allow'd, that the capacity of the mind is limited, and can never attain a full and adequate conception of infinity: And tho' it were not allow'd, 't wou'd be sufficiently evident from the plainest observation and experience." This, in fact, is one of Hume's central principles about infinity, to which he often appeals throughout (for example, at T 1.2.4.1/SBN 39). How, then, can "the common opinion, that the capacity of the mind is limited on both sides," be in "error" (T 1.2.1.5/SBN 28)? Hume is certainly not denying in this paragraph that we must always reach indivisible minima in our attempts to diminish or divide a given whole of extension. Indeed (ibid.): "Nothing can be more minute, than some ideas, which we form in the fancy; and images, which appear to the senses; since these are ideas and images perfectly simple and indivisible." This is precisely because there can be nothing "smaller" than a minimum: for all minima, as I have argued elsewhere, must have the same "size," namely, that of an indivisible arithmetical unit.[8] Nevertheless, something that appears as a minimum at one stage in a sequence of closely resembling phenomenologically given fields (as seen with the naked eye, for example) may not appear as a minimum at a later stage (as seen by a microscope which "gives parts to impressions"). Hume is not claiming, of course, that any such sequence can literally be extended to infinity (for this would indeed contradict his fundamental principle concerning the limited capacity of the human mind), but there is no reason to prescribe, at any given *finite* stage, that we cannot advance one step further.

Second, Hume is not making a realist point when he complains (again at T 1.2.1.5/SBN 28) that our senses "give us disproportion'd images of things, and represent as minute and uncompounded what is really great and compos'd of a vast number of parts." In particular, he is not suggesting that there are two distinct entities—impressions and mind-independent referents of these impressions—when he goes on to say that we (ibid.) take "the impressions of those minute parts, which appear to the senses, to be equal or nearly equal to the objects, and finding by reason, that there are other objects vastly more minute, we too hastily conclude, that these are inferior to any idea of our imagination or impression of our senses." What "we too hastily conclude," Hume says, is that there are some "objects" that are more minute than any of our (minimal) impressions or ideas. Indeed, it is "certain," Hume continues, that we can form ideas of *extremely* small "objects"—even those that are far beyond the threshold of ordinary perception with the naked eye.

Hume's example of the mite indicates the importance of his appeal to the microscope in the previous paragraph, for Robert Hooke's *Micrographia* (1665) contained a famous drawing of a greatly enlarged mite seen through the microscope.[9] Using this image, we can then find indivisible minima representing something vastly smaller than a mite—but these "new" minima now have the same "size" as the mite itself might have presented when seen with the naked eye,

[8] See note 6 above and the first article cited in note 1 above.

[9] See the Editors' Annotations to T 1.2.1.5 in Hume (2000, p. 435).

namely, as I have argued, the size of an ultimate arithmetical unit. Hume is not setting up a relationship between our ideas or impressions and absolutely imperceptible “objects” existing independently of our ideas. He is rather (once again) assuming a temporal sequence of closely resembling impressions or ideas obtained in the advance of natural science: for example, the image of a mite seen first with the naked eye, then with a magnifying glass, then with a low-power microscope, then with a more powerful microscope, and so on. The image in Hooke’s *Micrographia* does not immediately resemble the image seen with the naked eye, but the two are nonetheless connected by a temporal sequence of impressions in which any two immediately adjacent images do closely resemble one another.[10]

Third, whereas (T 1.2.1.5/SBN 28) “we too hastily conclude” that there are very tiny objects “inferior to any idea of our imagination or impression of our senses,” Hume says that “we ought rather to conclude, that the difficulty lies in enlarging our conceptions so much as to form a just notion of a mite, or even of an insect a thousand times less than a mite.” Nevertheless, that we can “enlarge” our conception (idea or image) of both the mite and its even more minute parts is shown by the image in Hooke’s *Micrographia*. So what is the “difficulty” in question? Hume continues as follows (ibid.): “For in order to form a just notion of these animals, we must have a distinct idea representing every part of them; which, according to the system of infinite divisibility, is utterly impossible, and according to that of indivisible parts or atoms, is extremely difficult, by reason of the vast number and multiplicity of these parts.” The “system of infinite divisibility” is the traditional conception of continuous extension that Hume is most concerned to reject, while the system of “indivisible parts or atoms,” in this context, is his own conception of sensible minima.[11] Why does Hume say that forming a “distinct idea representing every part” of a given (enlarged) image is “utterly impossible” in the first system, but only “extremely difficult” in the second?

Hume’s first principle concerning infinity explains why forming a “distinct idea representing every part” of a given whole of extension is “utterly impossible” in the system of infinite divisibility. To form such a distinct idea, on this system, would require an infinite number of steps (diminutions or divisions), and would therefore contradict the limited (finite) capacity of the human mind. In Hume’s own system of indivisible minima, by contrast, each of the parts composing a given whole of extension (including the minima themselves) can be reached in a finite number of steps (diminutions or divisions); yet, precisely because we go through diminutions or division that can only end up with one minimum at a time, we cannot obtain all of the minima simultaneously. Thus, although it is “utterly impossible” to reach all the minima in one *single* appearance (because of their confounding), it is also

[10] Compare the discussion in the section “Of skepticism with regard to the senses” of how the “fiction” of numerical identity of external objects arises from resemblance (e.g., at T 1.4.2.34–35).

[11] Hume cannot here be referring to *physical* atoms, for he has just introduced *sensible* minima as “perfectly simple and indivisible” parts of extension at T 1.2.1.2–4 (SBN 27–28)—he is explicitly considering “impressions of the senses” and “ideas of the imagination” (T 1.2.1.4/SBN 27–28). (The word “atom” simply means *indivisible*).

"extremely difficult" to reach them one at a time in any temporal *sequence* of appearances: their number, although finite, is still exceedingly "vast."

Hume's contrast here between the system of infinite divisibility and his own conception does not contradict his arguments concerning the inexactitude of the science of geometry. It is possible (albeit "extremely difficult") to arrive at an image of *each* of the minima originally confounded in a given whole of extension by separate temporal sequences of diminutions or divisions. However, it is impossible to determine, by a given fixed number of such separately constructed sequences, that we have thereby arrived at *all* of them. For example, we could have chosen to start the divisions with a sub-whole in the original whole different from the sub-whole with which we actually started. Because of the confounding of the minima, in choosing sub-wholes with which to start the process of divisions we might leave out minima to which we do not return when we proceed with separately constructed sequences of divisions of the sub-wholes we have left behind. Thus, it is indeterminate whether any particular number of separately constructed sequences has reached *all* of the minima for precisely the same reason (confounding) that their number in the original given whole is indeterminate. Since the finite number of minima is indeterminate due to their "confounding," we cannot establish a one-to-one correspondence between them and any finite (arithmetical) aggregate of discrete items.

In general, Hume's conception of the advance of our study of nature into the microstructure of bodies does not postulate hidden mechanisms independently of what we can actually observe. Hume follows Newton in rejecting the hypothetical method of the mechanical philosophy in favor of the desideratum of purely inductive inferences based only on uniform observed phenomena. In the discussion of tiny parts of bodies vastly smaller than a mite, Hume is not postulating (or hypothesizing) a relationship between our perceptions and some hidden microstructure existing independently of the appearances (in opposition to the realist assumptions of both the rationalist mechanical philosophers and Locke). Hume's process of advancing to sensible minima via smaller and smaller perceptible images is a purely inductive advance involving a continually increasing knowledge of causal laws of nature. And this is true whether the laws in question concern everyday causal relationships between visual experiences and the motions of our bodies or sophisticated optical instruments such as the microscope or the telescope—where the primary causal laws concern the reflection and refraction of light rays.[12] Although Hume rejects the realistic postulation (in advance of our inductive inquiries) of hidden structures below the threshold of perception (such as Locke's fixed microstructure of primary

[12] Hume points out at T 1.2.1.4 (SBN 28) that "[a] microscope or telescope, which renders [previously unseen objects] visible, produces not any new rays of light, but only spreads those, which always flow'd from them." The reference to "spreading" rays of light suggests Newton's famous experiments on the refraction of light rays by a prism. In my article on Hume, Locke and Newton cited in note 4 above, I develop Hume's understanding of Newton's prism experiments, and I argue, in addition, that the way both Hume and Newton accept corpuscularianism depends on a purely inductive advance into smaller and smaller *observable* parts of matter rather than on a postulation beforehand of an unobservable hidden microstructure—a postulation characteristic of the mechanical philosophy of Descartes, Leibniz, Locke and Boyle.

qualities), he does not preclude that we may "enlarge" our perceptions, in the manifest phenomena themselves, beyond any given threshold that may arise (at some particular stage) in our ongoing inductive exploration of nature.

In common life and science the desideratum is to advance by induction from the observed to the as yet unobserved, but without postulating beforehand what the unobserved might be like in itself independently of what we can observe. Moreover, Hume's eight "Rules by which to judge of causes and effects" formulated in T 1.3.15 (SBN 173–176) amount to a normative methodological procedure by which we can correct the relatively crude inductive extrapolations of the vulgar in pursuit of greater and greater uniformity. Our inductive advance into more and more regions of nature thus takes place under the normative guidance of the principle of the uniformity of nature, as we go beyond the initial appearances and reduce irregularities to regularities by finding ways to explain away contrary observations.[13]

Hume provides a striking illustration of this methodological procedure in Section VIII, Part I of the *Enquiry*, "Of Liberty and Necessity" (EHU 8.13/SBN 86–87): "The vulgar, who take things according to their first appearance, attribute the uncertainty of events to such an uncertainty in the causes as makes the latter often fail of their usual influence; though they meet with no impediment in their operation. But philosophers, observing, that, almost in every part of nature, there is contained a vast variety of springs and principles, which are hid, by reason of their minuteness or remoteness, find, that it is at least possible the contrariety of events may not proceed from any contingency in the cause, but from the secret operation of contrary causes."[14] Hume's reference to "minuteness or remoteness" suggests that there might be a connection between this striking passage from the *Enquiry* and the discussion of his experiments to refute infinite divisibility at T 1.2.1.3–5 (SBN 27–28) (involving the grain of sand, the ink spot, the microscope and telescope, and the image of the parts of a mite).

[13] I introduced for the first time what I take to be Hume's two standpoints—the philosophical standpoint of radical skepticism and the standpoint of common like and science—in De Pierris (2001, 2002). I further develop this distinction in De Pierris (2015). In particular, I argue that Hume's radical skepticism concerning our causal reasoning and its presupposed principle of the uniformity of nature is directed against his own and Newton's employment of this reasoning, which is naturally unavoidable and normatively required by our best method of inquiry in common life and science.

[14] A virtually identical passage occurs at T 1.3.12.5 (SBN 132). As I discuss in detail in my article on Hume, Locke and Newton cited in note 4 above, Hume goes on to say that "this possibility is converted into certainty by farther observation," where the "certainty" in question is that of (inductive) proof as opposed to (mere) probability. The methodological principle that is most relevant here is Hume's sixth rule (T 1.3.15.8/SBN 174): "The difference in the effects of two resembling objects must proceed from that particular, in which they differ. For as like causes always produce like effects, when in any instance we find our expectation to be disappointed, we must conclude that this irregularity proceeds from some difference in the causes." As I explain in the cited article, this and the other rules are grounded in the desideratum of attempting to achieve, in the Newtonian style, full (inductive) proof or perfect uniformity (under the guidance, therefore, of the principle of the uniformity of nature).

Moreover, such a connection appears unavoidable when we take into consideration the immediately following passage from the *Enquiry* (EHU 8.14/SBN 87): "Thus, for instance, in the human body, when the usual symptoms of health or sickness disappoint our expectation; when medicines operate not with their wonted powers; when irregular events follow from any particular cause; the philosopher and physician are not surprized at the matter, nor are ever tempted to deny, in general, the necessity and uniformity of those principles, by which the animal œconomy is conducted." It appears that the reference of "animal œconomy" is to Ephraim Chambers's (*c.* 1680–1740) well-known *Cyclopædia*, as is the reference of "animal spirits" ("of an insect a thousand times less than a mite") at T 1.2.1.5 (SBN 28). In both passages, therefore, it appears that Hume is envisioning further (inductive) investigation into the minute parts of the animal body as essential to the science of medicine.[15]

4 The Inductive Application of Geometry to Nature

According to Hume, it is possible to attain some degree of exact and certain intuitive knowledge in geometry. However, in order to achieve this knowledge we must confine ourselves to "very limited portions of extensions, which are comprehended in an instant." Hume writes (T 1.3.1.3/SBN 70): "As to equality or any exact proportion, we can only guess at it from a single consideration; except in very short numbers, or very limited portions of extension, which are comprehended in an instant, and where we perceive an impossibility of falling into any considerable error. In all other cases we must settle the proportions with some liberty, or proceed in a more *artificial* manner."

Hume has here not yet distinguished between arithmetic and geometry, and he proceeds to do so in the following three paragraphs—where he reminds us that geometry "never attains a perfect precision and exactness" (T 1.3.1.4/SBN 71), that algebra and arithmetic are "the only sciences, in which we can carry on a chain of reasoning of any degree of intricacy, and yet preserve a perfect exactness and certainty" (T 1.3.1.5/SBN 71), and finally that it "is the nature and use of [scientific]

---

[15] The Annotations to EHU 8.14 in Hume (1999, p. 240), explain that "Chambers defined animal œconomy as 'the first branch of the theory of medicine; or that which explains the parts of the human body, their structure and use; the nature and causes of life and health, and the effects or phænomena arising from them'." The Editors' Annotations to T 1.2.1.5 in Hume (2000, p. 435), refer us, for "animal spirits," to their later annotation to 1.2.5.20 (p. 443): "Chambers describes animal spirits as 'an exceedingly thin, subtile, moveable fluid juice or humour separated from the blood in the cortex of the brain, hence received into the minute fibres of the medulla, and by them discharged into the nerves, by which it is conveyed through every part of the body, to be the instrument of sensation, muscular motion, *&c*'. Although he grants that the 'existence of the *animal spirits*' is controversial, Chambers contends that they provide the best explanation of bodily motion and function: 'the infinite use they are of in the animal œconomy, and the exceedingly lame account we should have of any of the animal functions without them, will still keep the greatest part of the world on their side' (*Cyclopædia*, 'Spirit')."

geometry, to run us up to such appearances, as, by reason, of their simplicity, cannot lead us into any *considerable* error" (T 1.3.1.6/SBN 72, emphasis added). Hume has begun this discussion earlier (at T 1.3.1.4/SBN 70–71) by remarking that "geometry, or the *art*, by which we fix the proportions of figures … much excels both in universality and exactness, the loose judgments of the senses and imagination." It is clear, therefore, that the "more *artificial* manner" for settling proportions Hume refers to at T 1.3.1.3 (SBN 70) is no other than the demonstrative science of geometry.

For Hume, therefore, the demonstrative science of geometry, despite its unavoidably lesser degree of exactness and certainty in comparison with algebra and arithmetic, is nevertheless much more exact and certain than the looser judgments of the vulgar. We begin with relatively limited regions of extension (neither too large nor too small) in which the immediate appearances ("in an instant" or "at one view") present us with the "easiest and least deceitful" intuitive apprehensions of geometrical figures or diagrams. These include such cases as the proposition "that we cannot draw more than one right line between two given points" (Euclid, Postulate 1)—where, although errors are certainly still possible because of the necessary inexactitude of *all* geometrical appearances, yet our "mistakes can never be of any consequence." What we perceive with full intuitive certainty, therefore, is only the "impossibility of falling into any *considerable* error" (emphasis added), and the function of geometrical demonstrations, as we have seen, is to transfer the *relatively* high degree of certainty and exactitude of the fundamental principles (axioms) to all of their abstruser consequences.[16] This is the crucial advantage of Euclidean diagrammatic reasoning in the demonstrative science of geometry in comparison with our cruder estimations and reasonings in common life.

Nonetheless, propositions such as "that we cannot draw more than one right line between two given points" are still susceptible to very small or *in*considerable errors, and such (very small) errors depend on the limited regions of extension that we can apprehend "at one view." Earlier, in Book One, Part Two, Sect. 4 of the *Treatise*, Hume challenges the supposed geometrical standard of perfect exactness in connection with this very example (T 1.2.4.30/SBN 51–52): "Now since these ideas are so loose and uncertain, I wou'd fain ask any mathematician what infallible assurance he has, not only of the more intricate and obscure propositions of his science, but of the most vulgar and obvious principles? How can he prove to me, for instance, that two right lines cannot have one common segment? Or that 'tis impossible to draw more than one right line betwixt any two points? Shou'd he tell me, that these opinions are obviously absurd, and repugnant to our clear ideas; I wou'd answer, that I do not deny, where two right lines incline upon each other with a sensible angle, but 'tis absurd to imagine them to have a common segment. But supposing these two lines to approach at the rate of an inch in twenty leagues, I perceive no absurdity in asserting, that upon their contact they become one. For, I

[16] The one clear exception to this conception of the fundamental principles or axioms is the parallel postulate (Euclid, Postulate 5), which cannot be grasped intuitively in a limited region of extension.

beseech you, by what rule or standard do you judge, when you assert, that the line, in which I have suppos'd them to concur, cannot make the same right line with those two, that form so small an angle betwixt them? You must surely have some idea of a right line, to which this line does not agree. Do you therefore mean, that it takes not the points in the same order and by the same rule, as is peculiar and essential to a right line? If so, I must inform you, that besides that in judging after this manner you allow, that extension is compos'd of indivisible points (which, perhaps, is more than you intend) besides this, I say, I must inform you, that neither is this the standard from which we form the idea of a right line; nor, if it were, is there any such firmness in our senses or imagination, as to determine when such an order is violated or preserv'd."

Hume's example of the two lines that "approach at the rate of an inch in twenty leagues" suggests that, in the context of the fundamental principles of geometry, a "limited portion of extension" means a medium-sized region of space easily accessible ("in an instant") to the naked human eye, entirely unaided by optical or other instruments. If this interpretation is correct, then Hume's discussion of the fundamental principles of geometry would be connected to his earlier discussion of exploring both very minute and very distant bodies with the microscope and telescope at T 1.2.1.4–5 (SBN 27–28). The corresponding passage from Book One, Part Three, Sect. 1 reveals such a connection (T 1.3.1.4/SBN 71): "[Geometry's] first principles are still drawn from the general appearance of the objects; and that appearance can never afford us any security, when we examine the prodigious minuteness of which nature is susceptible."

The explicit reference to the "minuteness of *nature*" (emphasis added) clearly links T 1.3.1.4 (SBN 71) to T 1.2.1.4–5 (SBN 27–28), and the immediately following sentence returns to the example of Postulate 1 of Euclid (T 1.3.1.4/SBN 71; emphasis added): "Our ideas seem to give a perfect assurance, that no two right lines can have a common segment; but if we consider these ideas, we shall find, that they always suppose a *sensible* inclination of the two lines, and that where the angle they form is extremely small, we have no standard of a right line so precise as to assure us of the truth of this proposition." By a "sensible inclination" Hume means a perceptible angle, such that, in a given visual presentation or field, it appears larger than a perceptible minimum. However, as I have argued, in the earlier discussion at T 1.2.1.4–5 (SBN 27–28) he suggests that further empirical advance in the study of nature (with the microscope and telescope, for example) allows us to obtain new impressions larger than indivisible minima that nonetheless *closely resemble* those which, at an earlier stage, were such minima (just as a barely visible spot of ink appears larger when one moves closer). In such cases, we compare closely resembling perceptible fields in a particular temporal sequence of observations, and our progression through such a sequence always involves a purely inductive advance from what has been perceived at one time to new sensible impressions at a later time.

It follows that our application of geometry to nature, for Hume, involves a corresponding purely inductive advance. We begin from medium-sized regions of space easily accessible to the unaided human eye. Because of the unavoidable inexactitude of the science of geometry, we cannot foresee how this science will apply in the future to any new regions not yet perceivable (whether too large or too small)—in precisely the same way that we cannot foresee beforehand what experience in general has in store for us. Similarly, we cannot demarcate now what new regions of space may become accessible in the future, for new techniques and instruments might always become available. There is no absolute limit, fixed in advance, concerning how the exactitude of geometry might possibly be improved. Nevertheless, geometry cannot set up an ideal standard of *perfect* exactness beyond that which, at any given time, we can actually attain.

Hume makes precisely this point at T 1.2.4.24 (SBN 48): "[W]e… suppose some imaginary standard of equality, by which the appearances and measuring are exactly corrected, and the figures reduc'd entirely to that proportion. This standard is plainly imaginary. For as the very idea of equality is that of such a particular appearance corrected by juxta-position or a common measure, the notion of any correction beyond what we have instruments and art to make, is a mere fiction of the mind, and useless as well as incomprehensible. But tho' this standard be only imaginary, the fiction however is very natural; nor is any thing more usual, than for the mind to proceed after this manner with any action, even after the reason has ceas'd, which first determin'd it to begin." We might think that geometry, as the theory of supposedly continuous magnitude, allows us to anticipate beforehand the microscopic spatial structure of nature (this is a fundamental assumption of the mechanical philosophy). But this, according to Hume, is an illusion. Rather, in so far as any new spatial regions do become perceptually accessible, geometry stands ready to apply to these regions as well—that is, to apply to them, in turn, without any "considerable error." For these new regions (whether very big or very small) present visual appearances to which the axioms of geometry conform to a very high degree of approximation, just like those already perceptible to the naked eye.

It would be natural to suppose that, as we reach successively smaller regions by optical instruments, we will necessarily obtain appearances corresponding to ever more exact applications of the axioms of the science of geometry. Yet this is certainly not Hume's view. We cannot conclude, for example, that two very slowly converging straight lines that appear to have a common segment near their point of intersection will appear to coincide less and less in a sequence of appearances obtained with a microscope. Rather, all we can conclude is that the axioms of geometry will hold in this new region in just the same sense in which they held in the old: nothing about *increasing* exactness in any sequence of *different* appearances follows.

Hume makes no realist assumptions about the existence of the same geometrical figures in different phenomenological appearances: all we can have are distinct but closely resembling appearances following one another in a temporal sequence. Moreover, since Hume explicitly denies the infinite divisibility of space (whether actual or potential), we cannot know in advance how any newly perceptible smaller region is geometrically related to the larger region perceived earlier of which it is now a perceptible part. All such relationships, for Hume, depend on matters of fact (the second kind of philosophical relations in the *Treatise*) and thus on inductive inferences. The demonstrative science of pure geometry, despite its intrinsic inexactitude (and like the inexact relationship of resemblance), depends only on "relations of ideas" in the idiom of the *Enquiry* (the "first kind of philosophical relations" in the *Treatise*)—what we can call "analytic" relations of containment between the internal properties of the ideas compared.[17] Nevertheless, applying this science to nature is unavoidably inductive.

In this way, Hume radically departs from the rationalist mechanical philosophers (such as Descartes and Leibniz), the empiricist mechanical philosophers (such as Gassendi, Locke, and Boyle), and even from Newton. For none of these thinkers ever question the perfect exactitude of the science of geometry, and all of them suppose that geometry applies exactly to the space of our experience in all regions no matter how large or how small. They all accept without qualification the infinite divisibility of space, and they could not rest content with the Humean view that the axioms of geometry only apply inexactly and approximately (even if only with "no *considerable* error"). For they all assume that the science of geometry is in place prior to the advance of our empirical inquiries into ever larger and ever smaller regions of space, and it can thereby serve as a powerful and indispensable means for guiding this advance. (Thus, for example, they would all agree that ever-smaller regions made visible with the microscope necessarily provide ever more exact instantiations of the axioms of geometry). This is exactly what Hume denies, and Hume's commitment to inductivism is thereby even more far-reaching than Newton's. Bracketing out his radical skeptical arguments concerning the lack of ultimate justification of induction and its implicit principle of the uniformity of nature (the first such skeptical argument starts at T 1.3.6), Hume holds that (in common life and the science of nature) only the purely inductive principle of the uniformity of nature could ground a reasonable belief that geometry is capable of successively better approximations.

References

De Pierris, G. (2001). Hume's Pyrrhonian skepticism and the belief in causal laws. *Journal of the History of Philosophy, XXXIX*(3), 351–383.

De Pierris, G. (2002). Causation as a philosophical relation in Hume. *Philosophy and Phenomenological Research, LXIV*(3), 499–545.

De Pierris, G. (2005). Hume and Descartes on skepticism with regard to demonstrative reasoning. *Análisis Filosófico, XXV*(2), 101–119.

De Pierris, G. (2006). Hume and Locke on scientific methodology: The Newtonian legacy. *Hume Studies, XXXII*(2), 277–329.

De Pierris, G. (2012). Hume on space, geometry, and diagrammatic reasoning. *Synthese, 186*(I), 169–189.

[17] See my discussion of Hume's distinction between what we would now call "a priori" and "a posteriori" methods of reasoning in De Pierris (2005). I further developed the discussion of this distinction in my 2015 book cited in note 13 above.

De Pierris, G. (2013). Hume's phenomenological conception of space, time and mathematics. In M. Frauchiger (Ed.), *Reference, Rationality, and Phenomenology: Themes from Føllesdal* (pp. 107–119). Heusenstamm: Ontos Verlag.

De Pierris, G. (2015). *Ideas, evidence, and method: Hume's skepticism and naturalism concerning knowledge and causation*. Oxford: Oxford University Press.

Hume, D. (1975). *Enquiries concerning human understanding and concerning the principles of morals* (3rd rev. ed.). In L. A. Selby-Bigge & P. H. Nidditch (Eds.). New York: Oxford University Press.

Hume, D. (1978). *A treatise of human nature* (2nd rev. edn.*)*. In L. A. Selby-Bigge & P. H. Nidditch (Eds.). Oxford: Oxford University Press.

Hume, D. (1999). *An enquiry concerning human understanding*. In T. L. Beauchamp (Ed.). New York: Oxford University Press.

Hume, D. (2000). *A treatise of human nature*. In D. F. Norton & M. J. Norton (Eds.). New York: Oxford University Press.

Kant on Geometry and Experience

Michael Friedman

Towards the end of the eighteenth century, at the height of the German Enlightenment, Immanuel Kant developed a revolutionary theory of space and geometry that aimed to explain the distinctive relation of the mathematical science of geometry to our experience of the world around us—both our ordinary perceptual experience of the world in space and the more refined empirical knowledge of this same world afforded by the new mathematical science of nature. From the perspective of our contemporary conception of space and geometry, as it was first developed in the late nineteenth century by such thinkers as Helmholtz, Mach, and Poincaré, Kant's earlier conception thereby involves a conflation of what we now distinguish as mathematical, perceptual, and physical space. According to this contemporary conception, mathematical space is the object of pure geometry, perceptual space is that within which empirical objects are first given to our senses, and physical space results from applying the propositions of pure geometry to the objects of the (empirical) science of physics—which, first and foremost, studies the motions of such objects in (physical) space. Yet it is essential to Kant's conception that the three types of space (mathematical, perceptual, and physical) among which we now sharply distinguish are necessarily identical with one another, for it is in precisely this way, for Kant, that a priori knowledge of the empirical world around us is possible.

By contrast, it is precisely by insisting on the sharp distinction in question that we now find a fundamental error in Kant's theory, and we are thereby inclined to deny the possibility of what he calls *synthetic* a priori knowledge of the empirical around us—the possibility of any substantive a priori knowledge of its factual

M. Friedman (✉)
Department of Philosophy, Stanford University,
Building 90, 450 Serra mall, Stanford, CA 94305-2155, USA
e-mail: mlfriedman@stanford.edu

V. De Risi (ed.), *Mathematizing Space*, Trends in the History of Science,
DOI 10.1007/978-3-319-12102-4_12

structure. From an historical point of view, however, it is important to acknowledge that our contemporary conception evolved in explicit reaction to Kant's and, as a result, that we are still very much in his debt. It is also important to appreciate the extent to which Kant's theory represents the culmination of the new view of space and geometry characteristic of the early modern period, according to which space—the very same space in which we live and move and perceive—is essentially geometrical, so that it can then serve as the basis for the new mathematical science of physical nature.[1] Kant's conception, in this sense, represents the culmination of such attempts to reconfigure our understanding of space and geometry as the Cartesian conception that the nature or essence of matter is pure (three-dimensional) extension, the Newtonian conception of absolute space developed in the famous Scholium to the Definitions in the *Principia*, and the Leibnizean conception of space as governing the "well-founded phenomena" described by the new mathematical science but not the underlying metaphysical reality knowable by the pure intellect alone.[2]

*

Kant's explains his distinctive conception of space and geometry in the *Critique of Pure Reason* (1st ed. 1781, 2nd ed. 1787) in terms of their roles in what he takes to be the a priori necessary conditions underlying all human experience. These conditions, for Kant, are two-fold. On the one side are the a priori *sensible* conditions of such experience: space and time, as what Kant calls our pure forms of outer and inner intuition. On the other are the a priori *intellectual* conditions: the pure concepts or categories of the understanding, under the four headings of quantity, quality, relation, and modality. Under the heading of quantity, for example, are the three categories of unity, plurality, and totality; under the heading of relation are the three categories of substance, causality, and community; and so on. Kant takes such pure concepts of the understanding to be derived from the forms of judgement of traditional logic: unity, plurality, and totality from the forms of universal, particular, and singular judgements; substance, causality, and community, from the forms of categorical, hypothetical, and disjunctive judgements; and so on.

The faculty of sensibility, for Kant, is our passive or receptive faculty for receiving sensory impressions. In sharp contrast with all forms of traditional rationalism from Plato through Leibniz, however, Kant takes this receptive faculty to be itself a source of significant a priori knowledge, notably, the science of

---

[1] Aside from the development of the new mathematical science of physical nature, the early modern conception of space as essentially geometrical has other important sources as well—notably, the development of linear perspective in the painting of the Italian Renaissance. For the latter see, for example, Edgerton (1991).

[2] I shall return in the final section of this essay to the sense in which Kant's conception represents the culmination these early modern attempts—and, at the same time, is also quite essential for understanding the later development of our contemporary (explicitly anti-Kantian) conception.

geometry as grounded in our outer (spatial) pure intuition. The Transcendental Aesthetic that begins the *Critique of Pure Reason* is devoted to our sensible faculty, and Kant there describes sensibility as having a *form*—"in which the manifold of all appearances can be ordered in certain relations" (A20/B34)—where this form (e.g., the spatial form of all outer appearances) is invariant under all changes in the *matter* (corresponding to sensation) that is taken up or received within it.[3] The faculty of understanding, by contrast, is our active or spontaneous faculty of thought, which, considered by itself, has no intrinsic relation to our spatio-temporal sensibility. In particular, the formal structure of the understanding is quite distinct from that of sensibility, and it comprises, in the first instance, just the structure of the traditional logic of concepts and judgements: the tree of Porphyry for concepts, for example, and the originally Aristotelian classification of the forms of judgement.[4]

Yet this purely logical structure, for Kant, is empty of all content, and it can only acquire objective meaning or what Kant calls "relation to an object" by being brought into necessary connection with our spatio-temporal faculty of sensibility. Despite the fact that the categories originate in the formal structure of traditional logic, and, in this respect, are entirely independent of all sensory experience, they nevertheless require what Kant calls a *schematism* in terms of our spatio-temporal sensibility in order to have objective meaning for us: substance in terms of the representation of spatio-temporal permanence, causality in terms of the represen-

[3] I cite the *Critique of Pure Reason* by the standard A/B pagination of the first and second editions. All translations from Kant's German are my own. The passage just quoted reads in full (ibid.): "I call that in the appearance which corresponds to sensation its *matter*, but that which brings it about that the manifold of appearances can be ordered in certain relations I call the *form* of appearance. Since that within which sensations can alone be ordered and arranged in a certain form cannot itself be sensation in turn, the matter of all appearance, to be sure, is only given to us a posteriori, but its form must already lie ready for it in the mind a priori and can therefore be considered separately from all sensation." (In B the words "can be" in the first sentence replace "are" in A).

[4] In the tree of Porphyry the highest genus Being is divided into the species Created and Uncreated, Created is divided into the lower species Material and Immaterial, Material into the lower species Animate and Inanimate, and Animate into the still lower species Rational and Irrational: Human Being is thus defined as Rational Animate Material Created Being. In the traditional Aristotelian classification of the logical forms of judgement the square of opposition depicts the logical relationships among judgements with respect to logical quantity (universal or particular) and quality (affirmative or negative), resulting in the four forms A, E, I, and O: A = Every S is P, E = No S is P, I = Some S is P, O = Some S is not P. Kant himself goes beyond the traditional Aristotelian classification, not only by adding the triads of categorical, hypothetical, and disjunctive judgements and problematic, assertoric, and apodictic judgements, but also by adding singular judgements to the two traditional forms of logical quantity (universal or particular) and "infinite" judgements to the two traditional forms of logical quality (affirmative or negative): see A70–76/B95–101. The resulting table of exactly twelve logical forms of judgement, together with the corresponding table of categories, has generated considerable scholarly controversy. I shall briefly touch on one such controversy below.

tation of temporal succession, and so on.[5] The categories thereby acquire a more than purely logical meaning and result in substantive a priori claims about the most fundamental structure of our (sensible) experience: that substance is permanent in space and time, for example, or that every alteration in the state of a substance must have a cause.[6] It is in precisely this way, for Kant, that all of our a priori knowledge —both sensible and intellectual—can only be ultimately understood in terms of the formal conditions of the possibility of experience, that is, of empirical knowledge.

These peculiarities of Kant's conception of the a priori and its necessary relation to empirical knowledge leads to the main problem addressed in the Transcendental Analytic of the *Critique of Pure Reason* (which is concerned with what Kant calls "transcendental" as opposed to "formal" logic). If the a priori concepts or categories of the understanding originate in the logical forms of judgement, entirely independently of all sensible experience, how can we show that they have more than a purely logical meaning—that they do relate a priori to objects and, in fact, to all possible objects of our (human) sense experience? The pure forms of sensibility are not subject to this difficulty, since, assuming that there are such forms, they are precisely the forms of what is sensibly received or given. They therefore relate, necessarily, to all possible objects of our senses, that is, to all possible objects in space and time. But the categories are pure forms of *thought*, not forms of *sensory perception*, and so, in this case, an additional step is needed. We need to show that the spatio-temporal schematization that they require in order to acquire objective meaning and ground empirical knowledge is indeed forthcoming.

The key step in solving this problem is taken in the Transcendental Deduction of the Categories, as becomes especially clear in the notoriously difficult § 26 of the (completely rewritten) version in the second edition of the *Critique*. The title of this section is "Transcendental Deduction of the Universally Possible use in Experience of the Pure Concepts of the Understanding," and Kant begins by describing the problem as one of explaining "the possibility of knowing a priori, *by means of categories*, whatever objects *may present themselves to our senses*—not, indeed, with respect to the form of their intuition, but with respect to the laws of their combination … [, f]or if they were not serviceable in this way, it would not become clear how everything that may merely be presented to our senses must stand under laws that arise a priori from the understanding alone" (B159–160). Kant then notes that, "under *the synthesis of apprehension*," he "understand[s] the composition [*Zusammensetzung*] of the manifold in an empirical intuition, whereby perception, i.e., empirical consciousness of [the empirical intuition] (as appearance), becomes

[5] See the discussion of substance and causality in the Schematism chapter (A144/B183): "The schema of substance is the permanence of the real in time, i.e., the representation of it as a substratum of empirical time determination in general—which therefore remains while everything else changes. The schema of cause and the causality of a thing in general is the real, upon which, whenever it is posited, something else always follows. It therefore consists of the succession of the manifold, in so far as it is subject to a rule".

[6] These are the first two Analogies of Experience. I shall return to the Analogies, together with other principles of the understanding, below.

possible" (B160). The synthesis of apprehension, therefore, is just the process of taking up the matter of sensible and empirical intuition (corresponding to sensation) into conscious awareness, and Kant calls this process "perception"—of empirical objects that may appear before our senses.

We now come to the argument proper. The synthesis of apprehension, Kant says, "must always be in accordance with" our "a priori *forms* of outer and inner sensible intuition in the representations of space and time" (B160). This is straightforward, because it merely reiterates that the representations in question *are* our two forms of outer and inner intuition. But the crucial (and extraordinarily difficult) point, around which the argument turns, immediately follows. Kant first reminds us that space and time are unified or unitary representations in a sense already articulated in the Aesthetic (ibid.): "[S]pace and time are represented a priori, not merely as *forms* of sensible intuition, but as *intuitions* themselves (which contain a manifold), and thus with the determination of the *unity* of this manifold (see the transcendental aesthetic*)." He then infers that precisely this unity must therefore govern the synthesis of perception (B160–161): "Therefore, *unity of the synthesis* of the manifold, outside us or in us, and thus a *combination* with which everything that is to be represented in space or time as determined must accord, is itself already given a priori, as condition of the synthesis of all *apprehension*, simultaneously with (not in) these intuitions." And it finally becomes clear, in the immediately following sentence, that this same unity is actually due to the understanding rather than sensibility (B161): "But this synthetic unity can be no other than that of the combination of the manifold of a given *intuition in general* in an original consciousness, in accordance with the categories, only applied to our *sensible intuition*."

This is why, in the second sentence, Kant insists that the synthetic unity in question is given "with" rather than "in" the intuitions of space and time themselves. Indeed, Kant has already suggested his doctrine of the *transcendental unity of apperception*—the highest and most general form of unity of which the understanding is capable—by using the term "combination [*Verbindung*]" here (which is then repeated in the third sentence). For "combination" is introduced as a technical term at the very beginning of the Deduction (§ 15) to designate the activity most characteristic of the understanding (B129–130): "[T]he *combination* (*conjunctio*) of a manifold in general can never come into us through the senses, and can thus not be simultaneously contained in the pure form of sensible intuition; for it is an act of the spontaneity of the power of representation, and since one must call this, in distinction from sensibility, understanding, all combination is an action of the understanding, which we would designate with the title *synthesis* in order thereby to call attention, at the same time, to the fact that we can represent nothing as combined in the object without ourselves having previously combined it."

Kant continues, in the following paragraph, by asserting that "[c]ombination is the representation of the *synthetic* unity of the manifold" (B130), and this unity, he says, "precedes all concepts of combination" and thus "is not, for example, the category of unity" (B131). Indeed, it cannot be the product of any of the categories, "for all categories are based on logical functions in judging, but in these combination, and thus unity of given concepts, is already thought" (ibid). Therefore, Kant concludes, "we must seek this unity … still higher, namely, in that which contains the ground of the unity of different concepts in judging, and thus of the possibility of the understanding, even in its logical use" (ibid). The required ground, according to the following section (§ 16), is "the original-synthetic unity of apperception"—namely, the representation "*I think*, which must be *able* to accompany all my representations" (ibid). And, according to the next section (§ 17), "[t]he principle of the synthetic unity of apperception is the highest principle of all use of the understanding" (B136).

There can be very little doubt, therefore, that Kant, in § 26, is asserting that the very same unity that was first introduced in the Aesthetic as characteristic of space and time themselves can—surprisingly—now be seen to be due to the understanding after all. Indeed, if he were not asserting this it would be extremely hard to see how Kant could conclude his argument with the claim that the *categories* are conditions of the possibility of experience (B161): "Consequently all synthesis, even that whereby perception becomes possible, stands under the categories, and, since experience is knowledge through connected perceptions, the categories are conditions of the possibility of experience, and thus are a priori valid for all objects of experience." The unity of apperception—"the highest principle of all use of the understanding"—is the ultimate ground of both the categories and the characteristic unity of space and time.

This conclusion, however, is puzzling in the extreme.[7] The crucial difficulty arises in the first sentence of the main argument of § 26 (B160), which contains the justificatory reference back to the Aesthetic. For the primary claim of the Aesthetic, in this connection, is that the characteristic unity of space and time is *intuitive* rather than *conceptual*. So how can we possibly begin with a unity that was earlier explicitly introduced as *non*-conceptual and conclude that this same unity is due to the understanding after all? Indeed, when we examine the footnote attached to Kant's reference back to the Aesthetic, we can appreciate even more how deeply problematic the situation appears to be, and I shall consider this in detail in what follows. For now, however, I shall simply observe that the first sentence of the footnote illustrates Kant's point by the example of "[s]pace, represented as *object* (as is actually required in geometry)" (B160n). Careful attention to this example,

---

[7] It is puzzling, in particular, because Kant here appears to take back his insistence that sensibility and understanding are two quite different faculties with two quite different a priori formal structures. Indeed, an important line of thought in post-Kantian German philosophy, including both the post-Kantian German idealists and the Marburg neo-Kantians, explicitly appeals to what Kant says in § 26 to motivate an "intellectualist" reading according to which the forms of intuition become absorbed into the more fundamental unity of the understanding. And another important line of thought, culminating in the notorious "common root" interpretation of Martin Heidegger, insists on the radical independence of sensibility—leaving us, in the end, with no plausible reading of § 26. (For further discussion and references relevant to these two lines of thought see Friedman (2015)). The interpretation I am developing here aims fully to incorporate the ineliminable role of the understanding in the characteristic unity of space and time appealed to in § 26 while simultaneously preserving the independent contribution of sensibility.

I shall argue, illuminates the precise character of the difficulty in question—and also points the way to its proper resolution.

*

Before focussing on the footnote and the example of geometry, let us first consider the earlier passage from the Aesthetic to which Kant apparently refers in the main text. The passage in question is the third paragraph of the Metaphysical Exposition of Space, where Kant appeals to the characteristic unity and singularity of our representation of space to argue that it must be an intuition rather than a concept—and, indeed, it must therefore constitute the a priori form of all outer intuition. The characteristic properties of this representation to which Kant appeals are, first, that "one can only represent to oneself a single [*einigen*] space, and if one speaks of many spaces, one understands by this only parts of one and the same unique [*alleinigen*] space" (A24/B39), and, second, that "these parts cannot precede the single all-encompassing [*einigen allbefassenden*] space, as it were as its constituents (out of which a composition [*Zusammensetzung*] would be possible); rather, they can only be thought *within it*" (ibid.). It is for this reason, in fact, that "the general concept of *spaces* (emphasis added)"—in the plural, that is, the finite spatial regions that are parts of the "single all-encompassing space"—"rests solely on limitations," which carve out such regions from the single infinite space that contains them all.[8]

The crucial point, therefore, is that space is a singular individual representation, whose whole-part structure is completely different from that of any general concept. In the former case the whole "all-encompassing" space precedes and makes possible all of its limited parts (finite spatial regions), whereas, in the latter, the "parts" of any concept—that is, the "partial concepts [*Teilbegriffe*]" which together constitute its definition (as Rational and Animal are parts of the definition of Human Being: compare note 4 above)—precede and make possible the whole. The unity of a general concept, in this sense, is essentially different from that of our representation of space (and similarly for our representation of time), and this is the primary

[8] The paragraph reads in full (A24–25/B39): "Space is not a discursive, or, as one says, general concept of relations of things in general, but a pure intuition. For, first, one can only represent to oneself a single space, and if one speaks of many spaces, one understands by this only parts of one and the same unique space. These parts cannot precede the single all-encompassing space, as it were as its constituents (out of which a composition would be possible); rather, they can only be thought *within it*. It is essentially single; the manifold in it, and the general concept of spaces as such, rests solely on limitations. From this it follows that an a priori intuition (that is not empirical) underlies all concepts of space. Thus all geometrical principles, e.g., that in a triangle two sides together are greater than the third, are never derived from general concepts of line and triangle, but rather from intuition, and, in fact, with apodictic certainty." The final sentence makes it clear that the science of geometry is implicated in the distinctive whole-part structure that Kant is attempting to delineate, a point to which I shall return below.

reason, in the Aesthetic, that space and time count as intuitive rather than conceptual representations for Kant.[9]

Yet the difficult passage from § 26 of the Deduction appears to appeal to precisely the characteristically non-conceptual unity of space and time to argue that this same unity is actually due to the "unity of synthesis" that is most characteristic of the understanding—"even in its logical use" (B131). And the appended footnote only compounds the appearance of paradox. The first sentence says that "[s]pace, represented as object (as is actually required in geometry)," contains a "*grasping together* [*Zusammenfassung*] of the manifold, given in accordance with the form of sensibility, in an *intuitive* representation, so that the *form of intuition* gives merely a manifold, but the *formal intuition* gives unity of representation" (B160n). So the "unity of representation" here appears to be the "all-encompassing [*allbefassenden*]" unity of space emphasized in our passage from the Aesthetic (A24–25/B39).

The second sentence confirms this idea, and also explains why Kant had previously, in sharp contrast with his present point, taken the unity in question to be sensible as opposed to intellectual (B160–161n): "In the Aesthetic I counted this unity [as belonging] to sensibility, only in order to remark that it precedes all concepts, although it in fact presupposes a synthesis that does not belong to the senses but through which all concepts of space and time first become possible." The third sentence, however, appears to take this back, and even to contradict itself (B161n): "For, since through it (in that the understanding determines sensibility) space or time are first *given*, the unity of this a priori intuition belongs to space and time, and not to the concept of the understanding (§ 24)." Thus, after reiterating that the synthetic unity in question is a product of the understanding, Kant appears explicitly to deny that it is due to the understanding after all.[10]

I have recently proposed a solution to these apparent paradoxes that emerged out of my evolving work on Kant's theory of geometry. My original interpretation of this theory in Friedman (1985) emphasized the importance of Euclidean constructive reasoning for Kant and, in particular, appealed to Kant's understanding of such reasoning to explain the sense in which geometry, for him, is synthetic rather than analytic—an essentially intuitive rather than purely logical science. Yet I did not there explain the necessary relation between the science of geometry and what Kant calls our pure form of outer intuition: the (three-dimensional) space of perception within which all objects of outer sense necessarily appear to us. I first proposed such an explanation, which establishes a link between geometry and our

---

[9] This crucial difference in whole-part structure is emphasized especially clearly in the immediately following fourth paragraph of the Metaphysical Exposition of Space in the second edition (B39–40): "Space is represented as an infinite *given* magnitude. Now one must certainly think every concept as a representation that is contained in an infinite aggregate of different possible representations (as their common mark), and it therefore contains these *under itself*. But no concept, as such, can be so thought as if it were to contain an infinite aggregate of representations *within itself*. However space is thought in precisely this way (for all parts of space *in infinitum* exist simultaneously). Therefore, the original representation of space is an a priori *intuition*, and not a *concept*."

[10] Compare note 7 above, together with the paragraph to which it is appended.

passage from the Aesthetic (A24–25/B39), in Friedman (2000a).[11] And I found the missing link, in turn, in Kant's discussion of the relationship between what he calls "metaphysical" and "geometrical" space in his comments on essays by the mathematician Abraham Kästner in 1790.[12]

Kant's comments first describes the relationship between the two types of space as follows:

> Metaphysics must show how one *has* the representation of space, but geometry teaches how one can *describe* a space, i.e., can present it in intuition a priori (not by drawing). In the former space is considered as it is *given*, prior to all determination of it in accordance with a certain concept of the object, in the latter a [space] is *made*. In the former it is *original* and only a (single [*einiger*]) *space*, in the latter it is *derivative* and here there are (many) *spaces* —concerning which, however, the geometer, in agreement with the metaphysician, must admit, as a consequence of the fundamental representation of space, that they can all be thought only as parts of the single [*einigen*] original space. (20, 419)[13]

So it appears, in particular, that the (plural) *spaces* of the geometer—i.e., the figures or finite spatial regions that are iteratively constructed in Euclidean proofs—are prominent examples of the *parts* of the "single all-encompassing space" according to our passage from the Aesthetic (A24–25/B39).[14]

Kant's comments go on to discuss the different types of infinity belonging to geometrical and metaphysical space:

> [A]nd so the geometer grounds the possibility of his problem—to increase a given space (of which there are many) to infinity—on the original representation of a single [*einigen*], infinite, *subjectively given* space. This accords very well with [the fact] that geometrical and objectively given space is always finite, for it is only given in so far as it is *made*. That, however, metaphysical, i.e., original, but merely subjectively given space—which (because there are not many of them) can be brought under no concept that would be capable of a construction, but still contains the ground of construction of all possible geometrical concepts—is *infinite*, is only to say that it consists in the pure form of the mode of sensible

[11] I thereby attempted to build a bridge between the "logical" interpretation of Kant's theory of geometry developed in my earlier paper (an approach that was first articulated by Jaakko Hintikka) and the "phenomenological" interpretation articulated and defended by Charles Parsons and Emily Carson. For further discussion of this issue see also Parsons (1992).

[12] Kästner's three essays on space and geometry were first published in J.A. Eberhard's *Philosophisches Magazin* in 1790. Eberhard's intention was to attack the *Critique of Pure Reason* on behalf of the Leibnizean philosophy, and Kästner's essays were included as part of this attack. Kant's comments on Kästner, sent to J. Schultz on behalf of the latter's defense of the Kantian philosophy in his reviews of Eberhard's *Magazin*, were first published by Wilhelm Dilthey in the *Archiv für Geschichte der Philosophie* in 1890. They are partially translated in Appendix B to Allison (1973), which also discusses the historical background in Chapter I of Part One. Kant's comments have played a not inconsiderable role in the subsequent discussion of space and geometry in § 26, and, after presenting my own interpretation, I shall touch on some of this discussion below.

[13] All references to Kant's works other than the first *Critique* are to volume and page numbers in Kant (1900-).

[14] This point becomes clearer in light of the final sentence of our passage from the Aesthetic—which brings Euclid's geometry explicitly into the picture (see note 8 above: the example there is Proposition I.20 of the *Elements*).

> representation of the subject as a priori intuition; and thus in this form of sensible intuition, as singular [*einzelnen*] representation, the possibility of all spaces, which proceeds to infinity, is *given*. (20, 420–421)

Thus, whereas geometrical space is only *potentially* infinite (as it emerges step-by-step in an iterative procedure), metaphysical space, in a sense, is *actually* infinite —in so far as the former presupposes the latter as an already given infinite whole. Geometrical construction presupposes a single "*subjectively given*" metaphysical space within which all such construction takes place.[15]

In Friedman (2000a) I interpreted the relationship between these two kinds of space as follows. Metaphysical space is the manifold of all oriented perspectives that an idealized perceiving subject can possibly take up. The subject can take up these perspectives successively by operations of translation and rotation—by translating its perspective from any point to any other point and changing its orientation by a rotation around any such point. In this way, in particular, any spatial object located anywhere in space is perceivable, in principle, by the same perceiving subject. The crucial idea is then that the transcendental unity of apperception—the highest principle of the pure understanding—thereby unifies the manifold of possible perspectives into a single "all-encompassing" unitary *space* by requiring that the perceiving subject, now considered as also a thinking subject, is able, in principle, to move everywhere throughout the manifold by such translations and rotations. But this then implies that Euclidean geometry is applicable to all such objects of perception as well, since Euclidean constructions, in turn, are precisely those generated by the two operations of translation (in drawing a straight line from point to point) and rotation (of such a line around a point in a given plane yielding a circle).[16]

---

[15] Immediately preceding this passage Kant illustrates the distinction by contrasting geometry with arithmetic (20, 419–420): "Now when the geometer says that a line, no matter how far it has been continually drawn, can always be extended still further, this does not signify what is said of number in arithmetic, that one can always increase it by addition of other units or numbers without end (for the added numbers and magnitudes, which are thereby expressed, are possible for themselves, without needing to belong with the preceding as parts to a [whole] magnitude). Rather [to say] that a line can be continually drawn to infinity is to say as much as that the space in which I describe the line is greater than any line that I may describe within it." Thus, while the figures iteratively constructed in geometry are only potentially infinite, like the numbers, the former, but not the latter, presuppose a single "all-encompassing" magnitude within which all are contained as parts: i.e., the space "represented as an infinite *given* magnitude" of note 9 above (B39).

[16] This connection between Euclidean constructions and the operations in question is suggested by Kant himself (20, 410–411): "[I]t is very correctly said [by Kästner] that 'Euclid assumes the possibility of drawing a straight line and describing a circle without proving it'—which means without proving this possibility *through inferences*. For *description*, which takes place a priori through the imagination in accordance with a rule and is called construction, is itself the proof of the possibility of the object. However, that the possibility of a straight line and a circle can be proved, not *mediately* through inferences, but only immediately through the construction of these concepts (which is in no way empirical), is due to the circumstance that among all constructions (presentations determined in accordance with a rule in a priori intuition) some must still be *the first* —namely, the *drawing* or describing (in thought) of a straight line and the *rotating* of such a line

I appealed to these ideas in proposing an interpretation of the problematic footnote to § 26 in Friedman (2012a). The "unity of representation" mentioned in the second sentence of this footnote is indeed that considered in our passage from the Aesthetic (A24–25/B39), and Kant is indeed saying that this unity is a product of the understanding. It does not follow, however, that it is a *conceptual* unity—that it depends on the unity of any particular concept. It does not depend on the unity of any geometrical concept, for example, for the schemata of all geometrical concepts are generated by Euclidean (straight-edge and compass) constructions, and these presuppose, according to Kant, the prior unity of (metaphysical) space as a single whole. Nor does it depend on the unity of any category or pure concept of the understanding. For, by enumeration, we can see that none of their schemata result in any such object, i.e., space as a singular given object of intuition.

Rather, the unity of space as a singular given whole results directly from the transcendental unity of apperception, prior to any particular category, in virtue of the circumstance that the former unity, as suggested, results from requiring that the perceiving subject (which has available to it the manifold of all possible perspectives) is also a *thinking* subject. For the latter, as Kant says in § 16, must be "one and the same" in all of its conscious representations (B132).[17] The unity of apperception, as Kant says in § 15, is not that of any particular category but something "still higher"—namely, "that which itself contains the ground of the unity of different concepts in judging, and hence of the possibility of the understanding, even in its logical use" (B131). This is why Kant can correctly say, in the last sentence of the footnote to § 26 (B161n; emphasis added), that "the unity of this a priori intuition belongs to space and time, and not to the *concept* [i.e., *category*—MF] of the understanding (§ 24)."

If we follow the reference of this sentence back to § 24, moreover, we find that Kant there describes the figurative synthesis or transcendental synthesis of the imagination as "an action of the understanding on sensibility and its *first* application (at the same time the ground of all the rest) to objects of the intuition possible for us" (B152; emphasis added). He then proceeds to illustrates this synthesis by Euclidean constructions and explains that it also involves motion "as action of the subject":

(Footnote 16 continued)
around a fixed point—where the latter cannot be derived from the former, nor can it be derived from any other construction of the concept of a magnitude."

[17] More fully (B132): "[A]ll the manifold of intuition has a necessary relation to the *I think* in the same subject in which this manifold is encountered. But this representation is an act of *spontaneity*, i.e., it cannot be viewed as belonging to sensibility. I call it *pure apperception*, in order to distinguish it from the *empirical*, or also *original apperception*, because it is that self-consciousness, which—in so far as it brings forth the representation *I think* that must be able to accompany all others, and in all consciousness is one and the same—can be accompanied in turn by no other." Thus, the *I think* is the subject of which all other representations are predicated, whereas it can be predicated of no other representation in turn, and it is in precisely this sense that the *I think* cannot itself be a concept.

> We also always observe this [the transcendental synthesis of the imagination] in ourselves. We can think no line without *drawing* it in thought, no circle without *describing* it. We can in no way represent the three dimensions of space without *setting* three lines at right angles to one another from the same point. And we cannot represent time itself without attending, in the *drawing* of a straight line (which is to be the outer figurative representation of time), merely to the action of synthesis of the manifold, through which we successively determine inner sense, and thereby attend to the succession of this determination in it. Motion, as action of the subject (not as determination of an object*), and thus the synthesis of the manifold in space—if we abstract from the latter and attend merely to the action by which we determine *inner* sense in accordance with its form—[such motion] even first produces the concept of succession. (B154–155)

Thus, Kant begins with the two fundamental geometrical constructions (of lines and circles), and, after referring to a further construction (of three perpendicular lines), he emphasizes the motion ("as action of the subject") involved in drawing a straight line (and therefore involved in any further geometrical construction as well).

In the appended footnote, finally, Kant says that the relevant kind of motion (as an action of the subject rather than a determination of an object), "is a pure act of successive synthesis of the manifold in outer intuition in general through the productive imagination, and it belongs not only to geometry [viz., in the construction of geometrical concepts—MF], but *even to transcendental philosophy* [presumably, in the unification of the whole of space, and time, as *formal intuitions*—MF]" (B155n; emphasis added).[18] And one should especially observe how the representation of time necessarily enters here along with that of space. For the motion involved "in the *drawing* of a straight line" is what Kant calls "the outer figurative representation of **time**" (B154; bold emphasis added).[19]

*

I shall return to our representation of time in the penultimate section of this essay. For now, however, I shall add some further reflections on what we have already learned about space. This will help to clarify the special role of space and geometry, for Kant, among the mathematical sciences. It will thereby clarify, as well, the distinctive contribution of space and geometry within his conception of the necessary a priori conditions underlying all human experience.

I have argued that an adequate understanding of the problematic footnote to § 26 involves the distinction Kant makes explicit in his comments on Kästner between

---

[18] The footnote reads in full (B155n): "*Motion of an *object* in space does not belong in a pure science and thus not in geometry. For, that something is movable cannot be cognized a priori but only through experience. But motion, as the *describing* of a space, is a pure act of successive synthesis of the manifold in outer intuition in general through the productive imagination, and it belongs not only to geometry, but even to transcendental philosophy."

[19] The precise relationship between the representation of *motion* ("as action of the subject") in § 24 and the representation of time as a *formal intuition* suggested in § 26 is a delicate and subtle matter into which I cannot delve more deeply here; I provide some further discussion in Friedman (2015).

metaphysical and geometrical space—where the latter is generated step by step via Euclidean constructions of particular figures (lines, circles, triangles, and so on), and the former is given all at once, as it were, as *actually* rather than merely *potentially* infinite. Metaphysical space is thus the single "all-encompassing" whole within which all Euclidean constructions—along with the schemata of all geometrical concepts—are thereby made possible. It is in precisely this way that its characteristic unity precedes and makes possible "all *concepts* of space" (B161n; emphasis added), that is, all concepts of determinate regions of space ("spaces" in the plural) constituting particular geometrical figures [*Gestalten*].

Kant first discusses the characteristic unity of concepts in relation to our cognition of their corresponding objects in § 17. The *understanding*, he says, is "the faculty of *cognitions*," where these "consist is the determinate relation of given representations to an object" (B137). But an *object*, Kant continues, "is that in whose concept a given intuition is *united*," and "all unification of representations requires the unity of consciousness in their synthesis" (ibid.). He illustrates these claims by the unification of a given spatial manifold under the concept of a line (segment), whose object is just the determinate spatial figure (the determinate line segment) thus generated (B137–138): "[I]n order to cognize anything in space, e.g., a line, I must *draw* it, and therefore bring into being synthetically a determinate combination of the manifold, in such a way that the unity of this action is at the same time the unity of consciousness (in the concept of a line), and only thereby is an object (a determinate space) first cognized."

Yet when Kant discusses "[s]pace, represented as *object* (as is actually required in geometry)" in the problematic footnote to § 26 (B160n), he does not mean an object in *this* sense: he does not mean the object of any particular geometrical concept (or, indeed, of any other concept). The single unitary space discussed in the first sentence of the footnote is not geometrical space but rather the metaphysical space that precedes and makes possible all (geometrical) "*concepts* of space" (B161n; emphasis added). Kant's philosophical (or "metaphysical") claim is then that these (geometrical) concepts, together with their finite bounded objects (particular spatial figures), are themselves only possible in virtue of the prior "all-encompassing" metaphysical space in which all such bounded objects appear as parts. This prior metaphysical space—the whole of space as a formal intuition—is not an object *of* the science of geometry but rather an object considered at an entirely different level of abstraction (peculiar to what Kant calls "transcendental philosophy"), which, from a philosophical as opposed to a purely geometrical point of view, can nevertheless be seen as *presupposed* by the science of geometry.[20]

[20] I observed that interpreters have appealed to Kant's comments on Kästner while discussing space and geometry in § 26 (see note 12 above): notably, Martin Heidegger, in his lecture course on *Phenomenological Interpretation of Kant's* ***Critique of Pure Reason*** in the winter semester of 1927–1928 (1977, § 9), and Michel Fichant (1997), published along with his French translation of Kant's comments. Both Heidegger and Fichant, however, interpret space as a "formal intuition" in the footnote to § 26 as *geometrical* space in the terminology of the comments on Kästner—so that, according to them, the formal intuition of space is derivative from the more original "form of intuition" within which geometrical construction takes place. But this reading is incompatible with

Kant's more general philosophical claim concerns the role of space as a condition of the possibility of experience (empirical cognition)—and therefore its relationship, more specifically, to the pure concepts or categories of the understanding. The relevant concepts here are the categories of *quantity* or *magnitude* [*Größe*], and Kant emphasizes their role in his first illustration following the main argument of § 26:

> Thus, e.g., if I make the empirical intuition of a house into perception through apprehension of the manifold [of this intuition], the *necessary unity* of space and of outer sensible intuition in general lies at the basis, and I draw, as it were, its figure [*Gestalt*] in accordance with this synthetic unity of the manifold in space. Precisely the same synthetic unity, however, if I abstract from the form of space, has its seat in the understanding, and is the category of the synthetic unity of the homogeneous in an intuition in general, i.e., the category of *magnitude* [*Größe*], with which this synthesis of apprehension, i.e., the perception, must therefore completely conform. (B162)

All objects of outer sense, in other words, occupy determinate regions of space, and are therefore conceptualizable as measurable geometrical magnitudes (in determining, for example, how many square meters of floor space there are in a particular house).

Yet the pure intellectual concept of magnitudes as such, in contrast to the subspecies of specifically spatial (geometrical) magnitudes, "abstracts" from the form of space and considers only "the synthetic unity of the homogenous in an intuition in general"—or, as Kant puts it in the Axioms of Intuition, it involves "the composition [*Zusammensetzung*] of the homogeneous and the consciousness of the synthetic unity of this (homogeneous) manifold" (B202–203).[21] By "the composition of the homogeneous" Kant has primarily in mind the addition operation definitive of a certain magnitude kind (such as lengths, areas, and volumes), in virtue of which magnitudes within a single kind (but not, in general, magnitudes from different kinds) can be composed or added together so as to yield a magnitude equal to the sum of the two. Kant has primarily in mind, in other words, the Ancient

(Footnote 20 continued)
Kant's claim in the footnote that space as a formal intuition is both unified and singular in the sense of the Aesthetic—and, most importantly, that it precedes and makes possible all *concepts* of space. Here I am in agreement with Béatrice Longuenesse: for her comments on Heidegger in this connection see Longuenesse (1998a, pp. 224–225); for her parallel comments on Fichant see Longuenesse (1998b/2005, pp. 67–69). I shall return to the relationship between my reading and Longuenesse's below. (I am indebted to Vincenzo De Risi for first calling the exchange between Fichant and Longuenesse to my attention and for prompting me to consider more carefully the relationship between my reading of § 26 and Longuenesse's).

[21] More fully (B202–203): "All appearances contain, in accordance with their form, an intuition in space and time, which lies at the basis of all of them a priori. They can therefore be apprehended in no other way—i.e., be taken up in empirical consciousness—except through the synthesis of the manifold whereby a determinate space or time is generated, i.e., through the composition [*Zusammensetzung*] of the homogeneous and the consciousness of the synthetic unity of this (homogeneous) manifold. But the consciousness of the homogeneous manifold in intuition in general, in so far as the representation of an object first becomes possible, is the concept of a magnitude (*quanti*)."

Greek theory of ratios and proportion (rigorously formulated in Book V of the *Elements*), but now extended well beyond the realm of geometry proper to encompass a wide variety of physical magnitudes (including masses, velocities, accelerations, and forces) in the new science of the modern era.[22]

Nevertheless, despite this envisioned extension, Kant takes specifically geometrical magnitudes to be primary. In the Axioms of Intuition he again appeals, in the first place, to the successive synthesis involved in drawing a line (A162–163/B203): "I can generate no line, no matter how small, without drawing it in thought, i.e., by generating all its parts successively from a point, and thereby first delineating this intuition." He then refers to the axioms of geometry (B163/B204): "On this successive synthesis of the productive imagination in the generation of figures is grounded the mathematics of extension (geometry), together with its axioms, which express the conditions of a priori sensible intuition under which alone the schema of a pure concept of outer intuition can arise." And he finally asserts that the axioms of geometry, in this respect, are uniquely privileged (ibid.): "These are the axioms which properly concern only magnitudes (*quanta*) as such."

The sense in which geometry is thereby privileged becomes clearer in the immediately following contrast with *quantity* (*quantitas*) and the science of arithmetic (A163–164/B204): "But in what concerns quantity (*quantitas*), i.e., the answer to the question how large something is, there are in the proper sense no axioms, although various of these propositions are synthetic and immediately certain (*indemonstrabilia*)." Kant illustrates the latter with "evident propositions of numerical relations," such as "7 + 5 = 12," which are "singular" and "not general, like those of geometry" (A164/B205).[23] The import of this last distinction, in turn, becomes clearer in Kant's important letter of November 25, 1788 (to his student Johann Schultz) concerning the science of arithmetic (10, 555): "Arithmetic certainly has no *axioms*, because it properly has no *quantum*, i.e., no object [*Gegenstand*] of intuition as magnitude as object [*Obiecte*], but merely *quantity* [*Quantität*], i.e., the concept of a thing in general through determination of magnitude." Instead, Kant continues, arithmetic has only "*postulates*, i.e., immediately certain practical judgements," and he illustrates the latter by the singular judgement "3 + 4 = 7" (10, 555–556).

Kant's claim, therefore, is that arithmetic, unlike geometry, has no proper domain of objects of its own—no *quanta* or objects of intuition as magnitudes. Arithmetic is rather employed in calculating the magnitudes of any such *quanta* there happen to be, but the latter, for Kant, must be given from outside of arithmetic

---

[22] For discussion of the Ancient Greek theory of ratios and proportion see Stein (1990). For further discussion of this theory in relation to Kant see Friedman (1990), and Sutherland (2004a, b); 2006).

[23] Kant here illustrates the generality of geometry by the Euclidean construction of a triangle in general (A164–165/B205): "If I say that through three lines, of which two taken together are greater than the third, a triangle can be drawn, I have here the mere function of the productive imagination, which can draw the lines greater or smaller, and thereby allow them to meet at any and all arbitrary angles." (This is Proposition I.22 of the *Elements*; compare note 14 above).

itself. Kant thus does not understand arithmetic as we do: as an axiomatic science formulating universal truths about the (potentially) infinite domain of natural numbers. Nor, in Kant's own terms, is arithmetic an axiomatic science like geometry, which formulates universal truths about the (potentially) infinite domain of geometrical figures generated by Euclidean constructions—which, as we have seen, can be given or constructed in *pure* (rather than *empirical*) intuition. In particular, the potential infinity of this domain is guaranteed by the single, all-encompassing, and actually infinite formal intuition of space, which, in the end, constitutes the pure form of all outer (spatial) perception.

My reading of how the transcendental unity of apperception originally unifies our pure form of spatial intuition into a corresponding "all-encompassing" formal intuition is thus essentially connected with the science of geometry—the most fundamental science of mathematical magnitude. For I understand the pure form of intuition of space as a mere (not yet synthesized) manifold of possible spatial perspectives on possible objects of outer sense, where each such perspective comprises a point of view and an orientation with respect to a local spatial region in the vicinity of a perceiving subject. The unity of apperception then transforms such a not yet unified manifold into a single unitary *space* by the requirement that any such local perspective must be accessible to the same perceiving subject via (continuous) motion—via a (continuous) sequence of translations and rotations. And this implies, as we have seen, that the science of geometry must be applicable to all outer objects of perception. Space is thereby necessarily represented as comprising all specifically *geometrical* mathematical magnitudes.

It does not follow, however, that the transcendental unity of apperception and the pure concepts of the understanding take over the role of our pure forms of intuition, that there is no independent contribution of sensibility as in "intellectualist" readings like that of the Marburg neo-Kantians (see note 7 above). Rather, the representation of space as a formal intuition—as a single unitary (metaphysical) space within which all geometrical constructions take place—is a direct realization, as it were, of the transcendental unity of apperception within our particular pure form of outer intuition. For this form of intuition originally consists of an aggregate or manifold of possible local spatial perspectives, which the transcendental unity of apperception then transforms into a single, unitary, geometrical (Euclidean) space in the way that I have sketched above. Whereas our original form of outer intuition does not have the (geometrical) structure in question independently of transcendental apperception, it is equally true that no such realization of the latter can arise independently of our original form of outer intuition: this *particular* realization of the unity of apperception can by no means be derived in what Kant calls a manifold of intuition in general.[24]

[24] As we have seen, the pure intellectual concept of magnitudes in general abstracts from the structure of specifically spatial (geometrical) magnitudes and involves only "the synthetic unity of the homogenous in an intuition in general" (B162; compare note 21 above, together with the paragraph to which it is appended).

That there is a uniquely privileged mathematical science, the science of geometry, which establishes universal truths about a special domain of magnitudes (spatial regions as *quanta*) constructible in pure intuition, therefore depends on the existence of *our* pure form of outer intuition. Yet it also depends—mutually and equally—on the action of the transcendental unity of apperception (understood, in the first instance, in terms of a manifold of intuition in general) on this particular form of sensibility. Sensibility does make an independent contribution to the synthetic determination of appearances by the understanding, but it cannot make this contribution, of course, independently of the understanding. In particular, the distinctively geometrical structure realized in our pure form of outer intuition, on my reading, is the one and only realization of the unity of apperception in a domain of objects or magnitudes constructible in pure intuition.[25] And the understanding, on my reading, can only subsequently operate on empirical intuition through the mediation of the resulting formal intuition of space. The fundamental aim of the understanding, in this context, is to secure the possibility of the modern mathematical science of nature—which, as suggested, essentially involves a greatly expanded domain of physical magnitudes extending far beyond those traditionally considered in geometry.[26] It is in this way, as I shall argue in the penultimate section of this essay, that we can finally secure the possibility of what Kant calls experience.

*

It is illuminating, at this point, to compare my reading of § 26 of the Deduction with that developed by Béatrice Longuenesse. For, as suggested, I am in agreement with Longuenesse concerning fundamental issues surrounding the interpretation of this crucial section (see note 20 above). In particular, I agree with her that that the synthetic unity in question proceeds from the transcendental unity of apperception itself, prior to the synthetic unity of any particular concept or category, and we also agree that the unity characteristic of space as a formal intuition is precisely the same as that which was earlier characterized as intuitive as opposed to conceptual unity in discussing our form of outer intuition in the Transcendental Aesthetic. Yet the understanding originally affects sensibility, for Longuenesse, in *empirical* rather than *pure* intuition: in the process of "comparison, reflection, and abstraction" by which we ascend from what is sensibly given in perception to form ever more general empirical concepts. Moreover, it turns out, for Longuenesse, that arithmetic is prior to geometry in our application of the categories of quantity to objects of outer intuition—while, as we have just seen, the priority on my reading is exactly the reverse.[27]

---

[25] I am here indebted to a very helpful conversation with Graciela De Pierris concerning the precise connection between the transcendental unity of apperception and geometry in my reading.

[26] Compare the paragraph to which note 22 above is appended.

[27] As I have said, I first arrived at my reading of the role of the transcendental unity of apperception in § 26 in the course of the ongoing development of my work on Kant's theory of geometry.

Arithmetic becomes prior, for Longuenesse, in the following way. In her general conception of how the understanding originally affects sensibility the logical forms of judgement on which the categories are based are fundamentally forms of empirical concept formation by which we ascend from the sensible given in forming ever more general empirical concepts that can then figure in full-fledged acts of judgement. The understanding as the mere “capacity to judge [*Vermögen zu urteilen*],” however, can originally affect the sensible given in empirical intuition prior to all actual conceptualization and judgement. And, in this way, the understanding can “generate” our pure forms of outer and inner intuition (the formal intuitions of space and time) “as the necessary *intuitive* counterpart of our discursive capacity to reflect *universal* concepts, concepts whose extension (the multiplicities of singular objects thought under them) is potentially unlimited.”[28] Space and time thus become a priori “prepared,” as it were, for a corresponding potentially unlimited process of forming ever more general empirical concepts, and the application of the categories, as derived from the logical forms of judgement, to any sensible object given in space and time is thereby secured.

As we have seen, the logical forms associated with the categories of quantity are universal, particular, and singular.[29] For Longuenesse, subsumption of empirical intuitions under these forms involves a movement of thought from singular sensible items thought under some concepts (‘This body is heavy’, ‘This other body is heavy’, etc.) to the particular judgement involving the same concepts (‘Some [a number of] bodies are heavy’), to the corresponding universal judgement (‘All bodies are heavy’).[30] The resulting role for the categories of quantity is then to think

(Footnote 27 continued)
In my critical study of Longuenesse (1998a), Friedman (2000b), I discussed her interpretation of the categories of quantity, but I did not consider her views on the “pre-conceptual” synthesis of space and time (which views were then opaque to me). Only after I arrived at my own interpretation of § 26, in Friedman (2012a), did I appreciate what I now see as her important insight. My differences with Longuenesse concerning the formulation and articulation of this insight will emerge in what follows.

[28] See Longuenesse (1998a, p. 224): “[I]f one reads ‘the understanding’ as *das Vermögen zu urteilen*, the capacity to judge, then one can understand, as I suggested earlier, that the capacity to form judgments, ‘affecting sensibility,’ generates the pure intuitions of space and time as the necessary *intuitive* counterpart to our discursive capacity to reflect *universal* concepts, concepts whose extension (the multiplicities of singular objects thought under them) is potentially unlimited. When this original intuition is produced, no concept is thereby *yet* generated. Everything, as it were, remains to be done. But part of the minimal equipment that a human being, capable of discursive thought, has at his disposal, is the capacity to generate the ‘pure’ intuitions of space and time as that in which empirical objects are instances of *concepts* (i.e., universal representations, representations whose logical extension is unlimited).”

[29] See note 4 above, together with the paragraph to which it is appended and the preceding paragraph.

[30] Longuenesse thus follows a well-known paper by Frede and Krüger (1970) in taking the correspondence between categories and forms of judgement to be reversed in Kant’s published tables. This correspondence, according to Frede and Krüger, should align singular judgements with unity and universal judgements with totality rather than the other way around. (This is the controversy to which I allude at the end of note 4 above).

the extension of some general concept—the aggregate of singular items thought under this concept—in such a way that it can now be counted or enumerated. For Longuenesse, therefore, Kant's conception of quantity turns out to be very similar to Frege's later conception of number, as a notion that attaches primarily to the extension of a given concept and assigns a (cardinal) number to such a collection of objects so conceptualized. Finally, the figurative synthesis or transcendental synthesis of the imagination corresponding to these categories—the *schema* of the categories of quantity—is that spatio-temporal synthesis by which extensions of concepts in general are placed in space and time as "homogeneous multiplicities" countable or enumerable by successively iterated synthesis ('Here-now is one body', 'Here-now is a second body', 'Here-now is a third body', etc.). The pure intuitions of space and time, as synthesized in accordance with the logical forms of quantity, thereby provide something like a primitive mathematical theory of sets or aggregates: a general theory of the possible extensions of concepts, which, in particular, guarantees that such extensions can be potentially infinite in principle.[31]

Longuenesse is explicit that by "homogeneous multiplicity" she always means a number of distinct (sensible) particulars falling under a shared general concept.[32] The point, it appears, is that space and time serve to guarantee that, no matter how many concepts two sensible particulars share, it is always possible for them to be distinct: it is always possible for there to be *numerical* diversity together with *generic* identity, no matter how extensive the latter may be.[33] Thus the arithmetical units that are now to be successively enumerated and thereby added together are homogeneous in the purely logical sense: they are simply objects falling under the same general concept. The specifically mathematical notion of homogeneity figuring in the traditional theory of ratios or proportion—the property of being "composable" elements of a single magnitude kind or "dimension" (such as lengths, areas, or volumes)—is not yet at issue.[34] Indeed, there is no need at all for this specifically mathematical notion if the point (at least so far) is solely to guarantee that the extensions of empirical concepts, in general, can always be potentially infinite.

[31] See Longuenesse (1998a, p. 276): "For this to be fully clear, Kant should have said that the concept of number is not an ordinary concept, that is, not a 'common concept' that can be predicated, as mark or a combination of marks, of another concept. It is different from 'common concepts,' since it reflects as such (as multiplicities or, as Cantor will say, as sets having a determinate 'power') sets of objects defined by a concept."

[32] See Longuenesse (1998a, p. 252; bold emphasis added): "I maintain that according to Kant, even the category of quantity is originally acquired insofar as the power of judgment, reflecting on the sensible given in order to subordinate representations to empirical concepts combined in judgments, generates the *schema* of quantity—that is, a successive synthesis of homogeneous elements (**where 'homogeneous' means 'reflected under the same concept'**)."

[33] Kant discusses this situation in the Amphiboly in the course of criticizing Leibniz's doctrine of the identity of indiscernibles (A263–264/B319–320, A271–272/B327–328). Compare Longuenesse's discussion of the Amphiboly in her chapter on "Concepts of Comparison, Forms of Judgment, Concept Formation" (1998a, pp. 132–135).

[34] See note 22 above, together with the paragraph to which it is appended. For discussion of a variety of different notions of homogeneity—both logical and mathematical—see especially Sutherland (2004b).

Now, as Longuenesse emphasizes, arithmetical enumeration is certainly applicable to quantities that are mathematically (dimensionally) homogeneous in terms of the traditional theory of continuous magnitudes. For, in order to measure such a magnitude and assign it a numerical value, we can arbitrarily choose a unit (meters or feet in the case of length, for example) and then count or enumerate the number of such units composing the given magnitude in question. In particular, we may (exhaustively) decompose a given length into a multiplicity of (non-overlapping) line segments and take the measure of this length to be simply the number (in the Fregean sense) of the objects falling under the concept 'segments in the decomposition equal to the chosen unit'.[35] Yet Longuenesse's prioritizing of arithmetic over geometry may lead to misunderstandings concerning the fundamental mathematical differences between the two cases—between arithmetic as the science of discrete magnitude and geometry as the science of continuous magnitude. As a result, Kant's conception of the relationship between the categories of quantity and the science(s) of continuous magnitude may also become obscured.

There are two mathematical points worth noting here. In the first place, arithmetical units in the Fregean sense are merely logically homogenous or "equal" to one another: all they need have in common is that they fall under the same general concept, and, in particular, they need not have the same "size" or "magnitude" in any specifically mathematical sense. By contrast, the geometrical units applicable to continuous magnitudes (meters, square meters, cubic meters, and so on) necessarily have the same size or magnitude themselves, and they must all be congruent or equal to one another in precisely this sense—which, as such, goes far beyond the purely logical notion of conceptual homogeneity. In other words, although there is no doubt that such geometrical units *can* be taken to be homogeneous in the logical sense, the notions of equality and proportion relevant to continuous magnitudes have additional specifically mathematical structure. And it is precisely this additional structure that is required for their quantitative comparison.

A second and closely related point makes the fundamental difference between arithmetic and geometry—discrete and continuous magnitude—even more evident. One of the most significant discoveries of Ancient Greek mathematics was the phenomenon of *incommensurability* for continuous magnitudes: the existence of pairs of such magnitudes (like the side and diagonal of a square) that are not jointly measurable by any common unit, no matter how small a unit we choose. So the ratio or proportion between such magnitudes is not representable as a ratio of two whole numbers (as a rational fraction), but only, as we would now put it, by an irrational number (in our example $\sqrt{2}$). In this precise sense, therefore, it is possible

[35] See Longuenesse (1998a, p. 265): "[T]he same capacity to judge that makes us capable of reflecting our intuitions according to the logical form of quantity also makes us capable of recognizing in the line a plurality of homogeneous segments, thought under the concept 'equal to the segment *s*, the unit of measurement.' To 'subsume under the concept of quantity' is to count these segments, that is, to reflect the unity of this plurality of homogeneous elements." By contrast, the notion of homogeneity used in the traditional theory of proportion, as noted, has an essentially dimensional significance: lengths may be composed with lengths but not areas, areas with areas but not volumes.

to demonstrate that equalities and proportions between continuous magnitudes are incapable of purely arithmetical representation: geometry, in this sense, is demonstrably irreducible to arithmetic. Thus, while arithmetic can be subsumed under the general theory of proportion (now restricted to commensurable magnitudes), the latter can by no means be subsumed under the former. And it was just this mathematical discovery that initiated the first rigorous formulation of the general theory of proportion (including specifically incommensurable magnitudes) in Book V of the *Elements*.[36]

I am not claiming that Longuenesse ignores these mathematical differences between discrete and continuous magnitudes. Indeed, she provides an interesting discussion of specifically continuous magnitudes (*quanta continua*), and she also considers Kant's own treatment of irrational magnitudes (like $\sqrt{2}$) in his correspondence with August Rehberg.[37] The problem, rather, is that Longuenesse's treatment of the categories of quantity may easily suggest that the transition from the discrete to the continuous case is smoother and more natural than it actually is. It may suggest, in particular, that one can begin with a conception of these categories tailored to the enumeration of discrete aggregates of individuals, which are homogeneous only in the sense of falling under the same general concept, and then move to their application to truly continuous magnitudes without sufficiently attending to the quite different notion of (mathematical) homogeneity characteristic of the traditional theory of proportion. I want also to insist, accordingly, that Kant's own conception of the categories of quantity is not originally tailored to discrete aggregates of individuals at all, but rather to precisely the traditional general theory of proportion for continuous magnitudes.

What, for Kant, is the relationship between the categories of quantity, on the one side, and the representations of space and time as continuous magnitudes, on the other? For Longuenesse, as observed, space and time function in this context as principles of individuation for discrete sensible particulars falling under a common general concept. They secure the possibility of numerical diversity together with generic identity and thereby secure the possibility that the extension of any given empirical concept is potentially infinite. Moreover, they then make it possible to count or enumerate the individuals in any such extension via successively iterated synthesis. Note, however, that we have no reason so far to take space and time to be *continuous* magnitudes. We have no reason to require that there exists a congruence relation between spatial and temporal parts or regions (lengths, areas, and volumes), that the relations between such parts are thereby representable in the traditional theory of proportion, or, in modern terms, that the parts in question have any metrical properties whatsoever. It is only necessary, rather, that the parts of both space and time be sufficiently distinct from one another and sufficiently numerous

[36] A specifically arithmetical development of the theory of proportion, devoted to *commensurable* magnitudes (numbers), follows in Book VII. See the "Introductory Note" to Book V in Heath's edition of the *Elements* (1926, vol. 3, pp. 112–113) for discussion of how the discovery of incommensurables is reflected in its structure.

[37] See Longuenesse (1998a, pp. 263–271 and 262–263, respectively).

in a purely mereological sense: they may even be wholes consisting of discrete aggregates of parts.

For Longuenesse, that space and time are continuous magnitudes appears to be a brute fact about our two forms of sensible intuition. In particular, it has no essential connection with either our original conception of how the categories of quantity operate or our original conception of how the transcendental unity of apperception first gives unity to our forms of sensible intuition. For, if we begin by conceiving the operation of the categories of quantity in terms of the possibility of enumerating discrete individuals falling under the same general concept, then, as we have just seen, there is no need for our original conception of space and time to involve anything more than the possibility of potentially indefinite numerical diversity together with generic identity. Moreover, as we have also seen, Longuenesse conceives the original act of the understanding by which it first unifies sensibility to be that action of the "capacity to judge" by which, in empirical intuition, it "generates" our pure forms of outer and inner intuition (the formal intuitions of space and time) "as the necessary *intuitive* counterpart of our discursive capacity to reflect *universal* concepts" (see note 28 above). And so, once again, no further properties of space and time beyond their purely mereological (discretely individuative) properties are involved.

On my reading, by contrast, there is an essential connection between our forms of intuition and the relevant activities of the understanding. For I have suggested that, when Kant explains the category of magnitude in terms of "the composition [*Zusammensetzung*] of the homogeneous and the consciousness of the synthetic unity of this (homogeneous) manifold [in intuition in general]" (B202–203), he has specifically in mind the traditional theory of proportion articulated for continuous magnitudes.[38] To be sure, he is considering this theory in "abstraction" from our particular forms of spatio-temporal intuition and, in this sense, only as applied to a manifold of intuition in general. Yet Kant is not, on my view, beginning from a Fregean conception of number as enumerating an aggregate of discrete individuals falling under a common general concept, and he does not need to make a transition from discrete magnitude (arithmetic) to continuous magnitude (geometry and the theory of proportion). On the contrary, specifically continuous magnitudes are his focus from the very beginning, and so it is no wonder, in particular, that all of Kant's examples of magnitudes (*quanta*) in both the Transcendental Deduction and the Axioms of Intuition are continuous rather than discrete.[39]

---

[38] See note 21 above, together with paragraph to which it is appended. If, however, we take the traditional theory of proportion as Kant's model instead of arithmetic, we need to develop an alternative account to Longuenesse's of the correspondence between the categories of quantity and the (quantitative) logical forms of judgement (see note 30 above). Thompson (1989) has developed such an account, although I believe that it needs more work. Compare the discussion of Thompson in Longuenesse (2005, pp. 45–46)—which, in particular, suggests that she may be willing to revise her account of this correspondence accordingly.

[39] Longuenesse acknowledges—and even emphasizes—that the application of the categories of quantity to continuous magnitudes is most important to Kant. For example, in her section on continuous magnitudes (see note 37 above) she says that "the most important aspect of the

It is even more significant, however, that my conception of how the transcendental unity of apperception originally unifies our pure form of spatial intuition into a corresponding singular formal intuition is also essentially connected with geometry—the most fundamental science of continuous magnitude. For, as explained at the end of the previous section, it follows from my reading of § 26 that the representation of space as a formal intuition—as a single unitary (metaphysical) space within which all geometrical constructions take place—is a direct realization of the transcendental unity of apperception within our pure form of outer intuition. Space, in this sense, is necessarily represented as a specifically geometrical—and therefore continuous—mathematical magnitude, and it is so represented, in particular, on behalf of the transcendental unity of apperception. Nevertheless, as also explained at the end of the previous section, my reading preserves the independent contribution of sensibility, for this particular realization of the unity of apperception can by no means be derived in what Kant calls a manifold of intuition in general.[40]

In the end, neither my view nor Longuenesse's is subject to the standard criticism of "intellectualist" readings of § 26, according to which Kant's original conception of the faculties of sensibility and understanding as having distinct and independent a priori formal structures is thereby necessarily subverted (compare again note 7 above). Indeed, Longuenesse's view that the application of the categories of quantity to specifically continuous magnitudes is a result of the entirely contingent fact—so far as the understanding is concerned—that space and time happen to be continuous (rather than discrete) magnitudes makes the independence of the two faculties, on her view, especially clear. On my view, by contrast, we do not have radical independence in *this* sense because the categories of quantity have specifically continuous magnitudes in view from the very beginning, and sensibility

(Footnote 39 continued)
category of quantity" is "the role it plays in the determination of a *quantum*" (1998a, p. 265), and she goes on to single out continuous *quanta* in particular (p. 266): "[T]he category of quantity (*Quantität*) finds its most fruitful use when it serves to determine the *quantitas* of a *quantum*, that is, the *Größe*, the *magnitude* of an objects itself given as a continuous magnitude, a *quantum continuum* in space and time." Yet, because of her overriding emphasis on arithmetic and discrete magnitude in the original application of the categories of quantity, she is also led to the surprising claim that in the Axioms of Intuition "appearances are treated essentially as aggregates, namely discrete magnitudes" (2005, p. 50). One can see what she means by this from her earlier discussion of continuous magnitude. In particular, she there (1998a, p. 264) focusses on the notion of a "*quantum* 'in itself' *continuum*, but which I can represent as *discretum* by choosing a unit of measurement to determine the *quantitas* of this *quantum*—that is, *quoties in eo unum sit positum*, how many times a unit is posited in it." However, by failing to emphasize here that this arithmetical "representation" of a continuous magnitude (as discrete) is limited, and is in fact impossible in the comparison of *incommensurable* magnitudes, her discussion may easily give the impression that continuous magnitude can be considered as a species of discrete magnitude.

[40] See again note 24 above, together with the paragraph to which it is appended and the preceding paragraph.

makes an independent contribution by providing objective reality for these categories so understood.[41]

The crucial difference between Longuenesse's and my reading, therefore, is that I do not conceive the original action of the understanding on sensibility as aiming, in the first instance, to secure the possibility of "reflecting on the sensible given in order to subordinate representations to empirical concepts."[42] Rather, this action of the understanding on sensibility originally operates on the *pure* form of our outer intuition, by requiring that the perceiving subject—considered also as a thinking subject—is necessarily "one and the same" in all of its conscious representations (B132).[43] This action of the understanding, once again, results in the representation of space as a single, unified, *geometrical* (and therefore continuous) mathematical magnitude.[44] And it is only through the mediation of this representation, in the end,

---

[41] As observed (note 27 above), I did not sufficiently appreciate the importance of Longuenesse's reading of § 26 of the Deduction in Friedman (2000b). In particular, I did not then sufficiently appreciate the way in which her account preserves the independent contribution of sensibility by viewing continuity as a de facto property of space and time as our two forms of pure intuition—and, as a result, I did not sufficiently appreciate the way in which Longuenesse can and does make a transition from considering what she takes to be the original application of the categories of quantity in the enumeration of discrete particulars to the measurement of continuous magnitudes in space and time (see notes 37 and 39 above, together with the paragraphs to which they are appended). So Longuenesse (2001/2005) is perfectly correct to rectify these oversights and to emphasize, accordingly, that a proper appreciation of her reading requires one "to pay attention to the distinct and complementary roles Kant assigns to the logical forms of judgement, on the one hand, and to the pure forms of intuition and synthesis of the imagination, on the other" (2005, p. 53). I now agree with Longuenesse in emphasizing the "distinct and complementary roles" of the understanding and sensibility, but, at the same time, I differ with her on two remaining issues: (i) the structure of the categories of quantity, (ii) the relation of the transcendental unity of apperception to space (and time) in the transcendental synthesis of imagination. With respect to the first issue, the main point is that I take the application of these categories, first and foremost, to be to continuous rather than discrete magnitudes, and so, on my reading, there is no transition from the discrete to the continuous case at all. For the second issue see note 44 below.

[42] Compare, once again, note 28 above, together with the paragraph to which it is appended.

[43] See note 17 above, together with the paragraph to which it is appended.

[44] I am not suggesting that Kant appeals to the understanding to explain the continuity of space and time. In particular, it makes perfect sense, from the point of view of contemporary mathematics, to assign the property of continuity to the original form of spatial intuition (as a manifold of perspectives) and to appeal to the possibility of continuous motions (translations and rotations) only to explain the resulting *metrical* structure. The point, on my reading, is rather, from a contemporary point of view, that Kant assumes a full Euclidean (metrical) structure for space from the beginning—in virtue of which all its parts (lengths, areas, and volumes) then counts as three-dimensional continuous magnitudes in the sense of the traditional theory of proportion. He appeals to the understanding, however, as that which is originally responsible for the status of space as a unified and unitary formal intuition (metaphysical space) within which all (Euclidean) geometrical constructions take place. The fundamental difference between myself and Longuenesse here is that it is precisely this pure geometrical structure, on my view, that most directly realizes the transcendental unity of apperception in sensibility. (I am indebted to Vincenzo De Risi for raising the particular question of continuity in this connection).

that the application of the transcendental unity of apperception—and therefore the categories—to *empirical* intuition is then possible.[45]

*

When discussing the question "How is pure mathematics possible?" in the *Prolegomena to any Future Metaphysics* (1783), Kant isolates three principal mathematical sciences, namely, geometry, arithmetic, and "pure mechanics" (4, 283): "Geometry takes as basis the pure intuition of space. Even arithmetic brings its concepts of numbers into being through the successive addition of units in time; above all, however, pure mechanics can bring its concept of motion into being only by means of the representation of time." This passage suggests that it is pure mechanics, rather than arithmetic, which relates most directly to time. The reason, as Kant explains in the letter to Schultz, is that numbers are not themselves temporal objects (10, 556–557). Numbers are only "pure determinations of quantity," and not, like "every alteration (as a *quantum*)," properly temporal objects.[46] Whereas all calculation with numbers takes place *within* our pure intuition of time, the numbers themselves do not relate to parts of time or temporal intervals in the way in which the science of geometry—through the construction of figures—relates to parts of space or spatial regions corresponding to these figures.[47]

In the second edition of the *Critique* Kant added two new sections to the Transcendental Aesthetic: a "transcendental exposition of the concept of space" (§ 3) and a "transcendental exposition of the concept of time" (§ 5). The first argues

[45] See again the paragraph to which notes 25 and 26 above are appended. As I shall argue in the next section, it turns out that the mediation in question essentially involves the Newtonian mathematical theory of motion. In Longuenesse's reading, by contrast, there are two distinct aspects to Kant's grounding of experience—one involving the medium-sized sensible particulars of our ordinary perceptual experience, the other involving the objects of Newtonian mathematical natural science (2005, p. 54): "[I]f my reading is correct, Kant's argument is an attempt to account both for the pull of Aristotelianism in our ordinary perceptual world and for the truth of Newtonianism." I, for one, find much less of a contrast in Kant between ordinary and scientific experience, and much less room, accordingly, for an independent grounding of the former. This does not mean, however, that the argument of the first *Critique*, on my view, simply collapses into that of the *Metaphysical Foundations of Natural Science*: see the Conclusion to Friedman (2013) for my account of the fundamentally different perspectives (on the same phenomenal world) represented in these two works.

[46] More fully (ibid.): "Time, as you correctly remark, has no influence on the properties of numbers (as pure determinations of magnitude), as [it does], e.g., on the properties of every alteration (as a *quantum*), which is itself only possible relative to a specific constitution of inner sense and its form (time), and the science of number, regardless of the succession that every construction of magnitude requires, is a pure intellectual synthesis, which we represent to ourselves in thought."

[47] The crucial difference is that, whereas the science of number or arithmetic, for Kant, certainly presupposes the possibility of indefinite iteration (succession) in time, it does not yet constitute the determination of parts of time as mathematical magnitudes or *quanta*. As explained below, time only acquires what we would now call a metrical structure by means of precisely the mathematical theory of motion, whereby parts of time in particular can now be considered as *quanta*.

that space is indeed a pure or a priori intuition by appealing to the synthetic a priori science of geometry. The second, however, focusses on what Kant calls the "general doctrine of motion [*allgemeine Bewegungslehre*]." "[T]he concept of alteration," Kant says, "and, along with it, the concept of motion (as alteration of place) is possible only in and through the representation of time" (B48). The reason, he continues, is that "[o]nly in time can two contradictorily opposed determinations in one thing be met with, namely, *successively*" (B48–49). Therefore, he concludes, "our concept of time explains as much synthetic a priori knowledge as is set forth in the general doctrine of motion, which is by no means unfruitful" (ibid.). This strongly confirms the idea that it is the mathematical science of motion ("pure mechanics"), not arithmetic, which relates to time as geometry does to space—as the latter science, in particular, relates to the parts of space or spatial regions corresponding to geometrical figures (see again note 47 above).

It is in the *Metaphysical Foundations of Natural Science* (1786) that Kant develops in detail what he takes to be the synthetic a priori principles contained in the general doctrine of motion. He explains in the Preface that the natural science for which he is providing a metaphysical foundation is "either a pure or an applied *doctrine of motion* [*reine oder angewandte Bewegungslehre*]" (4, 476). Moreover, he concludes the Preface by saying that he wants to bring his enterprise "into union with the mathematical doctrine of motion [*der mathematischen Bewegungslehre*]" (478) and suggesting that he has Newton's *Principia* most centrally in mind.[48] Then, in the first chapter or Phoronomy, Kant characterizes the object of this synthetic a priori science as "the movable in space" (480) and remarks that "this concept, as empirical, could only find a place in a natural science, as applied metaphysics, which concerns itself with a concept given through experience, although in accordance with a priori principles" (482). Nevertheless, he also suggests that he is envisioning a *transition* from what § 24 of the B Deduction will call the pure act of motion of the subject—"as the *describing* of a space" (B155n)—to the motion of an empirically given object (a perceptible body) considered in the *Metaphysical Foundations*. For he begins the Phoronomy by considering moving matter as an abstract mathematical point—whereby "motion can only be considered as the *describing of space*" (489)—and reserves its subsumption under the more empirical concept of an extended (massive) body for later.[49]

The *Metaphysical Foundations* is organized into four main chapters—the Phoronomy, Dynamics, Mechanics, and Phenomenology—in accordance with the four headings of the table of categories (quantity, quality, relation, and modality).

[48] For further discussion of the relationship between Kant's "mathematical doctrine of motion" and Newton's *Principia* see Friedman (2012b) and (more fully) Friedman (2013).

[49] The quoted passage reads more fully (ibid.): "In phoronomy, since I am acquainted with matter through no other property but its movability, and thus consider it only as a point, motion can only be considered as the *describing of a space*—in such a way, however, that I attend not solely, as in geometry, to the space described, but also to the time in which, and thus to the velocity with which, a point describes a space. Phoronomy is thus just the pure doctrine of magnitude (*Mathesis*) of motion." For further discussion of the transition from pure to empirical motion see again Friedman (2012b) and (more fully) Friedman (2013).

In the third chapter Kant formulates his own three "Laws of Mechanics," which he employs in the fourth chapter to determine the true or actual motions in the cosmos from the merely apparent motions that we observe from our parochial position here on the surface of the earth.[50] He thereby shows how we can move from the mere "appearance [*Erscheinung*]" of motion to a determinate "experience [*Erfahrung*]" thereof (554–555). Moreover, whereas Kant's three Laws of Mechanics are derived as more specific realizations or instantiations of the three Analogies of Experience, his procedure for determining true from merely apparent motions involves a more specific realization or instantiation of the three Postulates of Empirical Thought. He determines the true from the merely apparent motions, in other words, by successively applying the three modal categories of possibility, actuality, and necessity. I have argued elsewhere, and in great detail, that Kant's model for this procedure is precisely Book 3 of the *Principia*, where Newton determines the true motions in the solar system from the initial "Phenomena" encapsulated in Kepler's laws of planetary motion and, at the same time, thereby establishes the law of universal gravitation.[51]

It is especially significant that Kant's Laws of Mechanics are more specific realizations of the Analogies of Experience. For the latter are characterized in the first *Critique* as the fundamental principles for the determination of time:

> These, then, are the three analogies of experience. They are nothing else but the principles for the determination of the existence of appearances in time with respect to all of its three modes, the relation to time itself as a magnitude (the magnitude of existence, i.e., duration), the relation in time as a series (successively), and finally [the relation] in time as a totality of all existence (simultaneously). This unity of time determination is thoroughly dynamical; that is, time is not viewed as that in which experience immediately determines the place of an existent, which is impossible, because absolute time is no object of perception by means of which appearances could be bound together; rather, the rule of the understanding, by means of which alone the existence of the appearances can acquire synthetic unity with respect to temporal relations, determines for each [appearance] its position in time, and thus [determines this] a priori and valid for each and every time. (A215/B262)

Just as we need the transcendental unity of apperception, in connection with the categories of quantity, to secure the application of the mathematical science of

[50] Kant's three Laws of Mechanics are the conservation of the total quantity of matter, the law of inertia, and the equality of action and reaction; compare the discussion (and illustration) of the synthetic a priori propositions of pure natural science in the Introduction to the second edition of the *Critique* (B20–21). I (briefly) comment on the relationship between these laws and the Newtonian Laws of Motion in Friedman (2012b) and (more fully) in Friedman (2013).

[51] See, e.g., Friedman (2012c) and (more fully) Friedman (2013). The successive determination of true from merely apparent motions, for Kant, involves a nested sequence of ever more comprehensive rotating systems—as we proceed from our parochial perspective here on earth, to the more comprehensive perspective of the center of mass of the solar system, to the even more comprehensive perspective of the center of mass of the Milky Way galaxy, and so on ad infinitum. Kant thereby reinterprets Newtonian absolute space (and extends the Newtonian determination of true from merely apparent motions far beyond the solar system) as the regulative idea (which can never be actually attained) of the ideal limit of this procedure: the perspective (which can never be actually attained) of the center of gravity of all matter.

geometry to all objects that may be presented within this form, we need the same transcendental unity of the understanding, in connection with the categories of relation, to generate a parallel mathematical structure (for duration, succession, and simultaneity) governing all objects that may be presented to us in time—that is, all objects of the senses whatsoever. And it is only at this point, in particular, that parts of time (temporal intervals) are themselves determined as mathematical magnitudes (*quanta*).[52]

But there is a crucial disanalogy between the two cases. The objects or magnitudes (*quanta*) considered in geometry, as explained, can be given or constructed in pure intuition—which, in turn, is the necessary form of all empirical intuition of outer objects. The Axioms of Intuition, therefore, are constitutive of such objects as appearances. The Analogies of Experience, however, as what Kant calls "dynamical" rather than "mathematical" principles, are concerned with "*existence* [*Dasein*] and the *relation* among [the appearances] with respect to [their] existence" (A178/B220). Further, because "the existence of appearances cannot be cognized a priori" (A178/B221), because "[existence] cannot be constructed" (A179/B221), the latter principles, unlike the former, cannot be constitutive of appearances (A180/B222–223): "An Analogy of Experience will thus only be a rule in accordance with which from perceptions unity of experience may arise (not, like perception itself, as empirical intuition in general), and it is valid as [a] principle of the objects (the appearances) not *constitutively* but merely *regulatively*."[53]

I can now delineate more exactly the uniquely privileged role of the mathematical science of geometry in Kant's conception of the possibility of experience. Geometry, for Kant, involves a procedure whereby all the objects of this science—all the figures considered in Euclid's geometry—are constructed step-by-step in pure intuition within space as a singular and unitary formal intuition (metaphysical space). Since all (outer) appearances as empirical intuitions are also given within this space, geometry necessarily applies to all objects of outer sense merely considered as objects of perception or appearance. The mathematical structure of time resulting from the general doctrine of motion, by contrast, can by no means be constructed in pure intuition. It can only arise within the context of the relational

[52] Compare again note 47 above. In applying the Analogies of Experience to the mathematical science of motion, in particular, we determine the magnitudes of temporal intervals by reference to idealized perfectly uniform motions, which then set the standard for correcting the actually non-uniform motions found in nature. In Newton's famous remarks concerning "absolute, true, and mathematical time" in the *Principia* (1999, p. 408), for example, we thereby correct the common "sensible measures" of time such as "an hour, a day, a month, a year" (ibid.), and I argue in Friedman (2013) that Kant takes this procedure as his model for time determination in the above passage from the Analogies (A215/B262). I also argue that Kant has the same procedure in mind in his Second Remark to the Refutation of Idealism, according to which, for example, we "undertake [*vornehmen*]" such time determination from the observed "motion of the sun with respect to objects on the earth" (B277–278).

[53] Although the Analogies of Experience are thus not constitutive of *appearances*, they are (of course) constitutive of what Kant calls "experience." Compare Kant's discussion of this distinction in the Appendix to the Transcendental Dialectic (A664/B692).

categories, and it thus involves a crucial transition from objects of perception or appearance to objects of what Kant calls experience. So we can only determine objects within this structure as objects of experience by beginning our determination in empirical rather than pure intuition.

This crucial asymmetry, for Kant, between the mathematical structure of time and that of space sheds further light on the independent contribution of the faculty of sensibility to the determination of the objects of experience by the understanding. I have explained the independent contribution of space as the form of outer sense in terms of the circumstance that geometry is the only mathematical science whose objects (as magnitudes) are determinable in pure intuition. It is now clear, however, that it is only by taking account of the characteristic structure of both space and time—the structure of our *spatio-temporal* sensibility—that we can fully appreciate the way in which our understanding can similarly determine the objects of experience. For the latter objects can only be so determined in empirical rather than pure intuition, and, for this purpose, we need to make a transition from *perception* (in accordance with the mathematical principles) to *experience* (in accordance with the dynamical principles).

The *Metaphysical Foundations*, I have suggested, takes the argument of Book 3 of the *Principia* as its model for determining true from merely apparent motions, and thus for determining "experience" from "appearance." In this procedure Kant substitutes his own Laws of Mechanics for Newton's Laws of Motion, where these Laws of Mechanics, in turn, are more specific realizations or instantiations of the Analogies of Experience. The determination in question, moreover, proceeds in accordance with the modal categories of possibility, actuality, and necessity, and thus by a more specific realization or instantiation of the Postulates of Empirical Thought.[54] So at the end of Kant's procedure, in particular, we have determined the resulting causal interactions between each body and every other body subject to the law of universal gravitation as *necessary* in the sense of the third Postulate (A218/B266): "That whose coherence [*Zusammenhang*] with the actual is determined in accordance with the universal conditions of experience, is (exists as) *necessary*." Indeed, as I have argued in detail elsewhere, it turns out that the law of universal gravitation itself (in sharp contrast with the Keplerian Phenomena from which it is inferred) is thereby determined, at the same time, as a universally valid and necessary law—as opposed to a merely inductive regularity or general rule.[55]

It follows, more generally, that the transition from what Kant calls "perception" to what he calls "experience" is also a transition from that which is merely *actual* (in the sense of the Postulates) to that which is *necessary* (in the same sense). For Kant says of the Postulates as a whole that they "together concern the synthesis of mere intuition (the form of appearance), of perception (the matter of appearance), and of experience (the relation of these perceptions)" (A180/B223). Indeed, in the second edition Kant reformulates the general principle governing all three

[54] See the paragraph to which note 51 above is appended.

[55] Friedman (2012c) is my most recent detailed discussion of this point.

Analogies so as, in effect, to explain "experience" in terms of such necessity (B218): "*Experience is only possible through the representation of a necessary connection* [*Verknüpfung*] *of perceptions*." And this explanation, in turn, reflects the intervening discussion in the *Prolegomena* of how "judgements of experience" differ from "judgements of perception" by the transformation of a merely inductive general rule ("If a body is illuminated long enough by the sun it becomes warm") into a genuine causal law ("The sun through its light is the cause of the warmth").[56]

The discussion in the *Prolegomena* concludes by illustrating Kant's conception of universally valid and necessary laws of nature more precisely. Kant says that he will illustrate his fundamental claim, that "*the understanding does not extract its laws* (a priori) *from, but prescribes them to, nature*" (§ 36; 4, 320), with "an example, which is supposed to show that laws which we discover in objects of sensible intuition, especially if these laws have been cognized as necessary, are already held by us to be such as have been put there by the understanding, although they are otherwise in all respects like the laws of nature that we attribute to experience" (§ 37; ibid.). And the example of such a law considered in the immediately following section is none other than the law of universal gravitation (§ 38; 4, 321): "a physical law of reciprocal attraction, extending to all material nature, the rule of which is that these attractions decrease inversely with the square of the distance from each attracting point."

It is striking, therefore, that this part of the *Prolegomena* is also echoed in § 26 of the Deduction. In particular, the conclusion of § 36 of the *Prolegomena* is echoed by the introductory remarks of § 26 of the Deduction where Kant announces the goal of the argument to follow: namely, to explain "the possibility of knowing a priori, *by means of categories*, whatever objects *may present themselves to our senses*, not, indeed, with respect to the form of their intuition, but with respect to the laws of their combination—**and thus to prescribe the law to nature and even make nature possible**" (B159; bold emphasis added). Moreover, after the conclusion of the main argument in § 26—"Consequently all synthesis, even that whereby perception becomes possible, stands under the categories, and, since experience is knowledge through connected [*verknüpfte*] perceptions, the categories are conditions of the possibility of experience, and thus are a priori valid for all objects of experience" (B161)—Kant finally arrives at the claim that the understanding is thus "the original ground of [nature's] necessary lawfulness (as *natura formaliter spectata*)" (B165). And this, in turn, echoes the second introductory question posed in § 36 of the *Prolegomena* (4, 318): "How is nature possible in the *formal* sense, as the sum total of the rules to which appearances must be subject if they are to be thought as connected [*verknüpft*] in one experience?"

It is not unreasonable to suppose, therefore, that Newtonian natural science in general and the law of universal gravitation in particular are just as relevant to the conception of experience articulated in the second edition Deduction as they are

[56] This is the famous response to Hume's "*crux metaphysicorum*" in the *Prolegomena* (§ 29; 4, 312). For a detailed discussion see again Friedman (2012c).

(explicitly) in the corresponding sections of the *Prolegomena*.[57] And it is quite clear, in any case, that Kant's treatment of the possibility of experience in the Deduction is just as involved with the question of how pure natural science is possible. The formal intuition of space as a whole highlighted in the footnote to § 26—"[s]pace represented as *object* (as is actually required in geometry)" (B160n)—is the three-dimensional, infinite, essentially geometrical space central to the new science of nature. It is that space in which all of nature is contained so as thereby to subject it to a unified system of mathematically formulated universally valid laws.[58] This modern conception of the laws of nature, Kant sees, has been finally successfully realized by Newton, who shows, for the first time, how we can thereby rigorously treat temporal duration as a mathematical magnitude as well. Kant incorporates this insight into his own revolutionary conception of transcendental time determination in accordance with the Analogies of Experience, whereby the universally valid and necessary laws of nature turn out to be prescribed to nature by us. Nature, on this conception, is nothing more nor less than the sum total of sensible objects in space and time, as necessarily subject to the lawgiving activity of the understanding. And it is in precisely this way that nature itself, for Kant, becomes the necessarily correlative object of our (human) experience.

*

At the beginning of this essay I suggested that Kant's revolutionary conception of the relationship between geometry and experience is centrally situated between the early modern conception of space—the very same space in which we live and move and perceive—as essentially geometrical and our contemporary conception, where the latter, by contrast, is based on a clear and sharp distinction between mathematical, perceptual, and physical space. I shall conclude, accordingly, by providing a bit more detail concerning the way in which Kant's conception is so situated.

---

[57] The relevance of the law of universal gravitation in particular is suggested in § 19 of the Deduction, which develops an account of the "*necessary unity*" belonging to the representations combined in any judgement as such—"i.e., a relation that is *objectively valid*, and is sufficiently distinguished from the relation of precisely the same representations in which there would be only subjective validity, e.g., in accordance with laws of association" (B142). Kant then illustrates his point by the relation between subject and predicate in the judgement "Bodies are heavy" (ibid.). This discussion continues the discussion of "judgements of experience" developed in the *Prolegomena*, and the example Kant chooses invokes universal gravitation as discussed in both § 38 of the *Prolegomena* and the *Metaphysical Foundations of Natural Science*. See Friedman (2012c) and Friedman (2015) for further relevant details concerning the relationship between the second edition Deduction and the *Prolegomena*.

[58] That the space in question functions as a foundation for both the mathematical science of (Euclidean) geometry and the modern conception of universally valid mathematical laws of nature sheds light on the sense in which the original act of the understanding responsible for the necessary unity of this space is more general than the unifying activity expressed in any *particular* category. We are thereby not confined, for example, to either the categories of quantity (as realized in the science of geometry) or the categories of relation (as realized in the universal laws of nature); rather, the original act in question insures the application of all of the categories together.

The early modern conception begins, as I suggested, with the Cartesian view that the essence or nature of matter, and thus of all of physical nature, is pure (three-dimensional) extension, so that physical nature itself consists of the objects of pure geometry actualized or made real. But this elegant and austere view was rejected by virtually all post-Cartesian thinkers, who took some additional, characteristically physical property of matter—such as impenetrability, mass, or force—to be unavoidably required. Thus Newton explicitly opposed the Cartesian metaphysics of space and matter with his own metaphysical conception of (absolute) space as a divine emanation, where properties such as impenetrability, mass, and force are then bestowed on particular regions of space by a separate act of divine creation.[59] But Leibniz, who of course entirely rejected Newtonian absolute space, took the space of the new mathematical science of physical nature to govern only the "well-founded phenomena" described by this science but not the more fundamental metaphysical reality underlying the phenomena.[60] And, as I have argued in detail elsewhere, Kant himself should be read as articulating a kind of synthesis of Leibniz and Newton—where, on the one side, Kant agrees with Leibniz that the new mathematical physics requires a metaphysical foundation in the purely intellectual (and so far non-spatio-temporal) concepts of substance, action, and force, and, on the other side, he agrees with Newton that such concepts only have sense and meaning when realized or instantiated (in Kantian terminology "schematized") within geometrical space.[61]

The primary aim of the present essay has been to explain in detail exactly how, for Kant, purely intellectual concepts—and, indeed, the pure intellect itself—are realized or instantiated in geometrical space, and exactly how, as a result, they then make our experience of nature possible. So what remains is briefly to consider exactly how Kant's late eighteenth-century account of these matters led from the early modern conception of space and geometry that was his starting point to our contemporary conception. The basic point I want to emphasize is that, although the latter conception, as explained, is indeed diametrically opposed to his, there is nonetheless a continuous conceptual evolution from Kant's late eighteenth-century account to the early twentieth-century developments in both geometry and physics that had our contemporary conception as a central result.

For, as I have argued in detail elsewhere, Kant's conception of space as both the pure form of our outer sensible experience and the basis for the mathematical science of geometry was taken up and generalized in the nineteenth century by first

---

[59] This conception is explicitly developed in Newton's unpublished *De Gravitatione*, but it also surfaces in some of his best-known published works, such as the General Scholium to the *Principia* and the Queries to the *Opticks*. See Stein (2002) for a detailed discussion, and compare Janiak (2008) for a somewhat different perspective.

[60] This, in any case, is the traditional understanding of Leibniz, which was certainly shared by Kant. Two recent more sophisticated interpretations—which argue for greater continuity between Leibniz's "phenomenalism" and Kant's—are Adams (1994) and De Risi (2007).

[61] See again Friedman (2013) for a detailed development of this reading, and compare Friedman (2009) for Kant's relationship, in particular, to Newton's metaphysics of space.

Hermann von Helmholtz and then Henri Poincaré, in an effort to accommodate what they took to be correct in Kant's conception to the new mathematical discoveries in non-Euclidean geometry.[62] During approximately the same time, moreover, Ernst Mach (and others) reconsidered the conceptual foundations of Newtonian physics, with the result that Newtonian absolute space was replaced by what we now call inertial frames of reference.[63] And Albert Einstein was explicitly indebted to both of these sets of developments (in geometry and physics respectively) in his revolutionary early twentieth-century theories of relativity.[64] In particular, as Einstein himself tells us in his aptly titled lecture, "Geometry and Experience," he could never have developed his general theory of relativity—which replaced the Newtonian theory of universal gravitation with a theory of dynamical variably-curved space-time—without the intervening work of Helmholtz and Poincaré.[65] It was in just this lecture, finally, that Einstein canonically articulated our contemporary (explicitly anti-Kantian) conception of the relationship between geometry and experience for us.[66]

References

Adams, R. M. (1994). *Leibniz: Determinist, theist, idealist*. Oxford: Oxford University Press.

Allison, H. E. (1973). *The Kant-Eberhard controversy*. Baltimore, London: Johns Hopkins University Press.

De Risi, V. (2007). *Geometry and monadology: Leibniz's analysis situs and philosophy of space*. Basel, Boston, Berlin: Birkhäuser Verlag.

[62] I discuss the relationship between Helmholtz's and Poincaré's conception of geometry as based on the principle of free mobility and Kant's "perspectival" conception of space as our pure form of outer sensible intuition in Friedman (2000a).

[63] For the relationship between Kant's reinterpretation of the Newtonian conception of absolute space (compare note 51 above) and the concept of what we now call an inertial frame of reference see Friedman (2013, pp. 503–509), where I also refer to DiSalle (2006) in the same connection. For Mach and the concept of inertial frame see DiSalle (2002).

[64] My most detailed discussion of the conceptual development from Kant through Helmholtz, Mach, and Poincaré to Einstein is Friedman (2010, pp. 621–664).

[65] See Einstein (1921); for further discussion see Friedman (2002).

[66] See Einstein (1921, pp. 3–4, 1923, pp. 28–29; my translation): "In so far as the propositions of mathematics refer to reality they are not certain; and in so far as they are certain they do not refer to reality. Full clarity about the situation appears to me to have been first obtained in general by that tendency in mathematics known under the name of 'axiomatics'. The advance achieved by axiomatics consists in having cleanly separated the formal-logical element from the material or intuitive content. According to axiomatics only the formal-logical element constitutes the object of mathematics, but not the intuitive or other content connected with the formal-logical elements." Although Einstein does not explicitly mention Kant here, these famous words were clearly intended and standardly taken as a rebuttal of the Kantian conception that mathematics (especially geometry) is paradigmatic of synthetic a priori knowledge. They were so standardly taken, in particular, by the logical empiricists beginning with Moritz Schlick—who is in turn favorably cited in precisely this connection by Einstein. For further discussion see again Friedman (2002).

DiSalle, R. (2002). Reconsidering Ernst Mach on space, time, and motion. In D. B. Malament (Ed.), *Reading natural philosophy: Essays in the history and philosophy of science and mathematics* (pp. 167–191). Chicago and La Salle, Ill.: Open Court Publishing Company.

DiSalle, R. (2006). *Understanding space-time: The philosophical development of physics from Newton to Einstein*. Cambridge: Cambridge University Press.

Edgerton, S. Y. (1991). *The heritage of Giotto's geometry: Art and science on the eve of the scientific revolution*. Ithaca, London: Cornell University Press.

Einstein, A. (1921). *Geometrie und Erfahrung. Erweiterte Fassung des Festvortrages gehalten an der Preussischen Akademie der Wissenschaft zu Berlin am 27. Januar 1921* (G. Jeffrey & W. Perrett (Eds.), Trans.). Berlin: Springer. (*Sidelights on Relativity,* pp. 27–55. London: Methuen, 1923).

Fichant, M. (1997). 'L'Espace est représenté comme une grandeur infinie donnée'. La radicalité de l'Esthétique. *Philosophie, 50*, 20–48.

Frede, M., & Krüger, L. (1970). Über die Zuordnung der Quantitäten des Urteils und der Kategorien der Größe bei Kant. *Kant-Studien, 61*, 28–49.

Friedman, M. (1985). Kant's theory of geometry. *Philosophical Review* 94 (pp. 455–506).

Friedman, M. (1990). Kant on concepts and intuitions in the mathematical sciences. *Synthese* 84 (pp. 213–257).

Friedman, M. (2000a). Geometry, construction, and intuition in Kant and his successors. In G. Scher & R. Tieszen (Eds.), *Between logic and intuition: Essays in honor of Charles Parsons* (pp. 186–218). Cambridge: Cambridge University Press.

Friedman, M. (2000b). Logical form and the order of nature: Comments on Béatrice Longuenesse's Kant and the capacity to judge. *Archiv für Geschichte der Philosophie, 82*, 202–215.

Friedman, M. (2002). Geometry as a branch of physics: Background and context for Einstein's 'geometry and experience'. In Malament, *op. cit.,* (pp. 193–229).

Friedman, M. (2009). Newton and Kant on absolute space: From theology to transcendental philosophy. In M. Bitbol, P. Kerszberg, & J. Petitot (Eds.), *Constituting objectivity: Transcendental perspectives on modern physics* (pp. 35–50). Berlin: Springer.

Friedman, M. (2010). Synthetic history reconsidered. In M. Domski & M. Dickson (Eds.), *Discourse on a new method: Reinvigorating the marriage of history and philosophy of science* (pp. 571–813). Chicago, La Salle, Ill.: Open Court Publishing Company.

Friedman, M. (2012a). Kant on geometry and spatial intuition. *Synthese, 186*, 231–255.

Friedman, M. (2012b). Newton and Kant: Quantity of matter in the metaphysical foundations of natural science. *Southern Journal of Philosophy, 50*, 482–503.

Friedman, M. (2012c). The *Prolegomena* and natural science. In H. Lyre & O. Schliemann (Eds.), *Kants Prolegomena. Ein kooperativer Kommentar* (pp. 299–326). Frankfurt am Main: Klostermann.

Friedman, M. (2013). *Kant's construction of nature: A reading of the metaphysical foundations of natural science*. Cambridge: Cambridge University Press.

Friedman, M. (2015). Space and Geometry in the B Deduction. In C. Posy & O. Rechter (Eds.), *Kant's philosophy of mathematics, vol. I: The critical philosophy and its background.* Cambridge: Cambridge University Press.

Heath, T. L. (Ed.). (1926). *The thirteen books of Euclid's elements*. Cambridge: Cambridge University Press.

Heidegger, M. (1977). *Phänomenologische Interpretation von Kant's Kritik der reinen Vernunft.* Frankfurt am Main: Klostermann). English edition: (1997). *Phenomenological Interpretation of Kant's Critique of Pure Reason.* Bloomington and Indianapolis: Indiana University Press.

Janiak, A. (2008). *Newton as philosopher*. Cambridge: Cambridge University Press.

Kant, I. (Ed.). (1900-). *Kant's gesammelte Schriften* (Vol. 29). Berlin: Royal Prussian (later German and Berlin-Brandenburg) Academy of Sciences, Berlin: Georg Reimer (later Walter de Gruyter).

Longuenesse, B. (1998a). *Kant and the capacity to judge: Sensibility and discursivity in the transcendental analytic of the critique of pure reason*. Princeton: Princeton University Press.

Longuenesse, B. (1998b). Synthèse et donation (Réponse à M. Fichant), *Philosophie* 60 (pp. 79–91). (Translated in B. Longuenesse 2005, pp. 64–78).

Longuenesse, B. (2001). Synthesis, logical forms, and the objects of our ordinary experience: Response to Michael Friedman. In *Archiv für Geschichte der Philosophie* 83 (pp. 199–212). (Reprinted in B. Longuenesse 2005, pp. 39–63).

Longuenesse, B. (2005). *Kant on the human standpoint*. Cambridge: Cambridge University Press.

Newton, I. (1999). *The principia*: *Mathematical principles of natural philosophy* (I. B. Cohen & A. Whitman, Ed., & Trans., assisted by J. Budenz). Berkeley and Los Angeles: University of California Press.

Parsons, C. (1992). The transcendental aesthetic. In P. Guyer (Ed.), *The Cambridge companion to Kant* (pp. 62–100). Cambridge: Cambridge University Press.

Stein, H. (1990). Eudoxus and Dedekind: On the ancient Greek theory of ratios, and its relation to modern mathematics. *Synthese, 84*, 163–211.

Stein, H. (2002). Newton's metaphysics. In I. B. Cohen & G. E. Smith (Eds.), *The Cambridge companion to Newton*. Cambridge: Cambridge University Press.

Sutherland, D. (2004a). The role of magnitude in Kant's critical philosophy. *Canadian Journal of Philosophy, 34*, 411–442.

Sutherland, D. (2004b). Kant's philosophy of mathematics and the Greek mathematical tradition. *Philosophical Review, 113*, 157–201.

Sutherland, D. (2006). Kant on arithmetic, algebra, and the theory of proportion. *Journal of the History of Philosophy, 44*, 533–558.

Thompson, M. (1989). Unity, plurality, and totality as Kantian categories. *The Monist, 72*, 168–189.

Index

V. De Risi (ed.), *Mathematizing Space*, Trends in the History of Science,
DOI 10.1007/978-3-319-12102-4