TRENDS IN THE HISTORIOGRAPHY OF SCIENCE

# BOSTON STUDIES IN THE PHILOSOPHY OF SCIENCE



VOLUME 151

# TRENDS IN THE HISTORIOGRAPHY OF SCIENCE


Edited by

KOSTAS GAVROGLU
*National Technical University, Athens*

JEAN CHRISTIANIDIS
*Greek Naval Academy, Athens*

*and*

EFTHYMIOS NICOLAIDIS
*National Hellenic Research Foundation, Athens*


SPRINGER-SCIENCE+BUSINESS MEDIA, B.V.

**Library of Congress Cataloging-in-Publication Data**

Trends in the historiography of science / edited by Kostas Gavroglu,
Jean Christianidis, Efthymios Nicolaidis.
p. cm. -- (Boston studies in the philosophy of science ; v.
151)
Includes index.
ISBN 978-90-481-4264-4 ISBN 978-94-017-3596-4 (eBook)
DOI 10.1007/978-94-017-3596-4
1. Science--Historiography. 2. Mathematics--Historiography.
I. Gavroglou, Kōstas. II. Christianidis, Jean. III. Nicolaidis, E.
IV. Series.
Q126.9.T74 1993
509--dc20 93-7415

ISBN 978-90-481-4264-4

*Printed on acid-free paper*

# TABLE OF CONTENTS

## PART I

PART II

*Historiography of Mathematics*

## PART III

### *Historiography of the Sciences*

# PREFACE

The articles in this volume have been first presented during an international Conference organised by the Greek Society for the History of Science and Technology in June 1990 at Corfu. The Society was founded in 1989 and planned to hold a series of meetings to impress upon an audience comprised mainly by Greek students and scholars, the point that history of science is an autonomous discipline with its own plurality of approaches developed over the years as a result of long discussions and disputes within the community of historians of science.

The Conference took place at a time when more and more people came to realise that the future of the Greek Universities and Research Centres depends not only on the progress of the institutional reforms, but also very crucially on the establishment of new and modern subject areas. Though there have been significant steps towards such a direction in the physical sciences, mathematics and engineering, the situation in the so-called humanities has been, at best, confusing. Political expediencies of the post war years and ideological commitments to a glorious, yet very distant past, paralysed the development of the humanities and constrained them within a framework which could not allow much more than a philological approach.

Like in many other countries, the establishment of the history as well as the philosophy of science in Greece has been faced with strong opposition coming from two seemingly different sides. The first was from some quarters of the humanities, who wondered why all these questions cannot be part of the ongoing discussions in philosophy. It may have been a legitimate and arguable viewpoint, if the same people did not have exactly the same attitude for psychology, sociology, political science etc. The second opposition came from people who professed that history and philosophy of science give a bad image to the sciences, since these subjects are nothing but popularisations of science, and that, furthermore, they are practised by scientists who had been failures in their scientific careers. Though this was dominant attitude, there were notable exceptions whose unswerving stand was rather decisive in the developments concerning the founding of these disciplines in Greece.

The great majority of the arguments advanced in the papers of this volume adhere to what has been codified as the internalist approach to the history of science. Such a choice does, by no means, imply that the organisers preferred this particular approach over any other. It was rather thought that the Corfu meeting will be followed by another one where the predominant *problematique* would be the externalist approach. There is no need whatsoever to elaborate on the differences between these two approaches. We do, though, want to stress that the differentiation has more of a historical and conventional justification and that neither of the two approaches can claim to be exclusively used for the analysis of any period or event in the history of science. What is to our mind a significant characteristic of these two approaches is their interdependence and the inherent pluralisms expressed in each one separately. This is why we would like to emphasise the word *Trends* in the title of the volume.

We have made an effort to include a fair number of papers examining the issues concerning the relation between philosophy and history of science. We feel that the discussion of this symbiotic relationship has always been beneficial in elucidating problems faced by each subdiscipline separately. Not all the approximately fifty papers presented at the meeting are included here. One in particular, Dr. Jens Hoyrup's paper "changing trends in the historiography of Babylonian mathematics" was far too long to be able to accommodate it in this volume.

The help of the Governing Body of the Ionian University and its President, Professor Elli Giotopoulou, had been absolutely decisive for the success of the Corfu meeting. We also thank the help we have received from the Technical Chamber of Greece and, especially, by its Secretary Mr. Stephanos Ioakimidis. The General Secretariat for Research and Technology, the Ministry of Education, The National Technical University of Athens and the University of Thessaloniki have also been generous in their help. Finally, we would like to thank two persons in particular who have been so very much supportive of all our activities concerning the history and philosophy of science in Greece. From our very first attempts to try to establish these disciplines in Greece, we have found an enthusiastic ally in Dr. Spiro Latsis with his continuous interest about the various developments and funding of some of our activities. The second person is Professor Robert S. Cohen who has managed from his rather unassuming headquarters at the University of Boston, to open channels of communication among so many members of the international community and to be so catalytic for the developments concerning the institutionalization of history and philosophy of science in so many countries. Finally we wish to thank our editors Annie Kuipers and Evelien

Bakker of Kluwer Academic Publishers for being so patient with us and so very efficient in the production of the volume.

*October 1993*

KOSTAS GAVROGLU,
JEAN CHRISTIANIDIS,
EFTHIMIOS NICOLAIDIS.

RICHARD S. WESTFALL

# CHARTING THE SCIENTIFIC COMMUNITY

Instead of presenting and defending a thesis in this paper, I will describe the research project on which I am engaged. Let me begin first with a brief justification of it. In broad terms, the project is a social history of the scientific community of the 16th and 17th centuries, the period of the Scientific Revolution. The study rests on the premise that the modern scientific community dates from those years as surely as modern science itself does. My goal is to chart the parameters of the social existence of those engaged in the study of nature at that time. The project is not Boris Hessen revisited. Neither for that matter is it Schaffer and Shapin rewarmed. I am doing my best to avoid the question of what caused the rise of modern science. The question appears to me as a trap, or if you will a morass in which historical research bogs down, and instead of illumination we get ideological assertions that determine the conclusions of research conceived for the purpose of supporting them. I am incurably empirical in outlook and uneasy with historical research that departs far from its empirical base. In this study I seek to explore the parameters of the social existence of those engaged in the study of nature during the 16th and 17th centuries; I am not concerned with any hypothesis about the origins of modern science. As I have not failed to recognize, there are points at which the study verges, whatever my intent, toward issues of causation. I am not so sanguine as to think that I succeed any better than others in eschewing apriori judgments in these cases.

When I started, three major issues or questions stood at the center of my attention. I was confident that more would appear as I proceeded, and in this I have not been disappointed. Nevertheless the three original ones continue to be central. First is the issue of support. The study of nature is an expensive enterprise. Usually it requires books and equipment. Frequently it requires expeditions. Always its requires leisure from remunerative employment. In the 16th and 17th centuries, moreover, the study of nature lacked all, or virtually all, of that demand in the market place, arising from the merging of technology and science in our age, which works to obscure for us the fact that support of science is an issue. How was it possible that men – in the 16th and 17th centuries the practices



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 1-14.

of society determined that the study of nature was almost exclusively a male occupation – were able to devote themselves to this pursuit? Patronage was a major factor in the support of science. It has been one focus of my concern during the past few years, and I will return to it later in this paper. Here let me insist only that patronages situates itself, in my view, within the broader problem of support in general.

A second issue is the technological involvements of scientists. I am not interested here in Baconian talk about useful knowledge. Everyone knows there was plenty of that during the 17th century. I am concerned rather with the involvement of scientists (if I may use this anachronistic word to avoid continually referring to those engaged in the study of nature) in projects of real utility. The centrality of this issue to any study of the social history of the scientific community seems obvious to me, and I shall not pause further to justify it.

Third, I am interested in the evolution of societies through which scientists established and maintained communication with each other. Recently, when I presented a paper drawn from this study at a meeting, a question was raised about whether I can legitimately claim that scientists of the 16th and 17th centuries constituted a community. (I have used the word "community" in the title of this paper, and, for lack of a better word, have not been loth to use it in discussing my project.) I am unable to see that the question of whether scientists of that period constituted a true community matters for the research in which I am engaged. As I said above, I am attempting to chart the parameters of the social existence of those engaged in the study of nature. One of the parameters was their communication with others engaged in similar studies whereby they began to knit themselves into communities. There were not many of these outside the ranks of university-based Scholastic philosophers when the 16th century dawned. By the end of the 17th, we can trace the existence of a considerable number, and the communities themselves were beginning to attach themselves together in the larger one, now in the singular, that I call the scientific community. It has gone on growing in size and importance since that time to the point of summoning into existence in our day a subdiscipline of sociology to study it.

As part of the larger project, I am collecting what I call a catalogue of the scientific community. It will be my topic in the rest of the paper. I indicated above that I am incurably empirical in outlook. I tend to think that something can be learned by counting what can legitimately be counted; one, though not the sole, function of the catalogue is to count or measure, in rough form, certain dimensions of the scientific community. I have based the catalogue on the *Dictionary of Scientific Biography*. The advantage of using the DSB is that it provides me with a list of scientists assembled by those most competent to know, but a list that I myself did not compile and cannot then have slanted, consciously or unconsciously, to some preconceived end. I am concerned, of course, with the Western scientific community, and without further ado I eliminated from my consideration the Eastern and Arab scientists included in the DSB. I started with those born in the decade of Copernicus' birth, the 1470s, and continued to those born in 1680, who thus reached the age of twenty in the final year of the 17th century.

In all, the DSB lists 651 Western scientists born within this span of two hundred and ten years. I purged 21 from that list, not because of the level of their scientific work – this seems like a subjective judgment to me, and I am striving to minimize subjective elements in the catalogue – but because they do not appear to me to have been scientists at all. Those purged include religious figures such as Roberto Bellarmino, Jacob Boehme, and Sebastian Franck, and others like Zacharias Jansen (in the DSB solely because of a claim, now universally rejected, that he invented the telescope), and Jean Nicot (a philologist in the DSB solely because his name became the root of our word "nicotine"). I am left then with 630 scientists whom I am using, in some sense, as representative of the scientific community of the 16th and 17th centuries. It will be obvious to everyone, and it is obvious to me, that there are problems here. If those who selected the list for the DSB did their job well – and though I may dissent on a small number of cases, I am convinced that they did – the 630 constitute the best scientists of the age and cannot then be representative of the others. At the moment I have no indication of how representative the 630 may be in the topics on which I am gathering information. I do need to confess, however, that I am not unduly dismayed by this issue. The scientific community of the 16th and 17th centuries did not include the enormous numbers which would make the question, of how representative the sample is, critical for the 20th century. I have long been convinced that the Scientific Revolution was the product of a small handful of natural philosophers; no aspect of my present research has called that conviction into question. I invite anyone to contemplate the peripheral significance, not just of a few, but of an extensive number of scientists included in the DSB (which he or she will not need to look far in any volume to find), and to remain deeply concerned about some imagined population of scientists not represented.

The DSB furnished the list. For information about the 630 I have gone far beyond the DSB and tried to consult the best secondary literature available on each of them. The research has yielded two products: a report or sketch of each of the scientists, two to three pages long on average, all of them organized under the same format of ten headings; and a dBase file into which I enter the information in the reports. The purpose of the dBase file, of course, is to count and correlate the information entered into it, and I will proceed in this paper to cite numbers. The longer I have worked on this project, however, the more uneasy I have become with the dBase counts, and I do not wish to conceal the fact. There is the problem of representativity mentioned above. I sometimes calculate percentages to one decimal point, though I will not do so in this paper. If the information entered in the file is correct, percentages to one decimal are perfectly legitimate for the set of 630, but they cannot be accurate for the community as a whole, even though I do not believe the issue of representativity is acute. More problematic in my eyes is the need to categorize in order even to organize a dBase file of this sort. The infinite variety of human life seems to resist reduction to rigid categories; and having spent long hours collecting the information, I cannot be unaware of how much some categories lack in homogeneity and how uncertain the demarcations between some are. As I proceed, the reports and the information they contain appear more and more as the products of my study that have the greatest significance. Increasingly I see the dBase counts primarily as heuristic indications of questions that would bear further investigation. Even when I calculate percentages to one decimal, I understand them only as gross indications of the extent of the social phenomena in question.

When I started the catalogue I expected to complete it, with the aid of a couple of assistants, in two years of only part time effort on my part. Six years later it remains the principal focus of my activity. It is now tolerably complete, however, and what I am presently doing is expanding some of the reports in which I want more detail. The results that I shall cite here are then preliminary, indications if you will of the potential that I see in the information.

Collecting the catalogue has been a good exercise. It has introduced me to aspects of the Scientific Revolution I had not known. As far as I can recall, I had not heard of Juan Caramuel y Lobkowitz. I do not see how I could have heard of someone named Caramuel y Lobkowitz and then forgotten his name. Caramuel was born in 1606 of a Spanish mother and a Czech father, who was an engineer in the service of the Hapsburg monarch. Caramuel was the author of some seventy works on a variety of subjects. As a mathematician he invented a system of logarithms recognizable to mathematicians today as the first incarnation of cologarithms. He made contributions as well to astronomy, physics, and natural philosophy. When he was teaching at Louvain during the Thirty Years War, he planned the defenses that held off the Prince of Orange. Later, in Prague, he was present when a surprise incursion of Swedes nearly seized the city; it was Caramuel who rallied the defenders to drive the invaders out. In the Low Countries he had made the acquaintance of the Papal Nuncio, Cardinal Chigi. Chigi appears to have been of mixed minds about Caramuel. On the one hand he thought he was half mad; on the other he was greatly impressed. The later view prevailed, and when Chigi became Pope Alexander VII, he appointed Caramuel to a bishopric in Italy, where he ended his days.

Leonty Filoppovich Magnitsky presents a rather different case. The son of impoverished Russian peasants, and the only Russian among the group of 630, he was sent by his parents to a monastery. He surprised the monks by his self-taught ability to read, and eventually they sent him on to the order's school in Moscow. Magnitsky became the source of modern Russian mathematics. He also caught the eye of Peter the Great who set him up, first as a teacher in his Navigation School, and then as its director. As happened with those known to be proficient in mathematics, Magnitsky found himself called, again by Peter, to plan the defenses of Tver against the Swedish invaders during the Great Northern War. I had not met either Caramuel or Magnitsky, and I like to think that my understanding of the Scientific Revolution has been enhanced by acquaintance with them and with others like them.

My goal, however, is not anecdotal information about individuals but quantified social phenomena, however rough the quantification may be. My first heading is demographic data, the years of birth and death. For this heading alone the information in the DSB sufficed, although in a small number of cases other sources led me to alter a date. The first thing then that impressed me as I began to collect the catalogue was the longevity of scientists in that period. The average lifespan of the 630 (or better 599, since even approximate lifespans cannot be determined for 31 of the group) was only a fraction less than sixty-five. As soon as I mention that to anyone, they respond with the same thought that occurred immediately to me. By definition, everyone in the set survived the perils of infancy and childhood. I have not yet found time to consult works in historical demography to see if an average lifespan of sixty-five was exceptional for those who survived initially to twenty. Whatever I find in that regard, I do think there is something of importance in this data. One of the hoariest generalizations of our discipline is the proposition that scientists do their creative work at an early age. The assertion does not appear compatible with my data. Only one of the 630, Jeremiah Horrocks, failed to live beyond twenty-five. Only six (including Horrocks) failed to live beyond thirty. Only thirty-four (including the six) failed to live beyond forty. Every indication is that those who failed to live beyond forty were most unlikely to have produced scientific work that would appear sufficient, three and four centuries later, for their inclusion in the DSB. It is worth noting that Galileo, Descartes, and Newton would not be there, and Kepler just barely.

My second heading concerns the father – his occupation or status and his financial standing. On no subject were my sources more apt to be silent. Moreover, the information on the family's relative wealth appears especially suspect to me, a subjective judgment in the first place, interpreted subjectively by me three or four centuries later. Nevertheless I collected what data I could find, though I intend to use it with extreme caution. One thing that does not appear ambiguous did catch my eye, the number of sons of clerics. If we confine ourselves to those born after 1540 and to Protestants, of 271, forty-two in all, or more than one of seven, was the son of a cleric. When we take into account the fact that there is no information at all about the fathers of nearly a quarter of the set, the proportion was about one out of five. As I realize, I should not have been surprised. The sons of clerics grew up in an atmosphere in which learning was both present and prized. My catalogue suggests that the presence of learning in the family was roughly twice as important as the presence of wealth, as far as later scientific achievement is concerned.

I have recorded nationality third, nationality of birth, national setting of the career, and national setting of death. As with the majority of my categories, national setting of career is not exclusive; a few (such as Caramuel) pursued their careers in as many as four settings. Although the information in this case is not ambiguous, I intend to use it sparingly. The DSB is the product largely of Anglo-American scholars, and I would regard an effort to measure, say, the relative size of national scientific communities from the number of entries in the DSB as wholly meaningless. The information about nationality was the most readily available; it seemed foolish not to record it. It does appear to me that correlations and comparisons within national groups (for example, the proportion in the medical sciences, the proportion supported primarily by patronage, the proportion engaged in technological enterprises) would not be meaningless where the differences are pronounced.

One phenomenon that I find interesting emerged from my use of nationality. There are fifteen Danes among the 630. Eight of the fifteen belonged to one extended family. Thomas Finke, a physician first active in the late 16th century (though known to the history of science primarily in mathematics) was related by the marriage of his cousin to Peder Soerenson, another physician. Finke's two daughters married Caspar Bartholin and Ole Worm. A sister of Bartholin married Christian Soerensen, known to historians of astronomy as Longomontanus. Caspar Bartholin's two sons, Erasmus and Thomas, were both scientists of European eminence. I pause to note that Thomas was well enough entrenched to install four sons in chairs at the University of Copenhagen, not all in sciences and none of sufficient accomplishment to be in the DSB. The daughter of Erasmus Bartholin married Ole Roemer. I should add that a granddaughter of Ole Worm was engaged at one time to Ole Borch (or Borrichius), who was another member of the Danish contingent in the DSB, though the engagement was later broken off, and that two others among the Danish contingent were related to each other. Jacob Winslow was the grand nephew of Niels Stensen (Steno), who in turn had been very close, though not related, to Thomas Bartholin. I do not know of any other national scientific community so tightly organized around one family. I am unable to read Danish; I have relied here on the work of a graduate assistant who does. Perhaps these facts are well known to Danish historians; if they are not, I am convinced that a fascinating and important story remains to be explored.

The next heading is university education. I have been accustomed over the years to saying that the scientific community of that period was not located in the universities but that it was university educated. Both statements are now appearing dubious to me. I shall come back to the first later. For the moment let me note that 247, or about two out of five, of the 630 did not obtain a bachelor's degree. The number needs further analysis to be meaningful. A university degree held no significance for a wealthy man or an aristocrat, and a number of them attended universities for a time without bothering to take a degree. Others, such as Robert Boyle and Blaise Pascal, never attended a university, although they were exposed to the same body of learning through private tutors. Both of these factors reduce the true size of the group without university education. Nevertheless, after all the analysis, the number of contributing scientists of the 16th and 17th centuries without university education remains larger than I had thought, and not all of them were minor figures.

I was impressed also by the quantitative identification of certain universities as clear foci of scientific life. Ninety-one different institutions appear in the sources I have consulted as "universities" where individuals in the set studied, though I am sure I will not wish to call every one of them a university if I take the time to find more about them. Add to the ninety-one the educational establishments of the religious orders, especially the Jesuits. The overwhelming majority of the institutions had fewer than five scientists (I mean from among the 630, of course) as students; spread out over a period of two hundred and ten years, the number does not suggest a center of scientific work. A few had larger concentrations. Not surprisingly, Padua had the largest number, sixty-six; among the other Italian universities Bologna had twenty-five. In France, Paris had a concentration of fifty-three; Montpellier followed with twenty-seven. Germany had many universities but no single center. Twenty-five studied at Basel, though the number appears to be inflated by several who took medical degrees by examination without significant periods of study there. There were twenty-one at Wittenberg and fifteen at Jena and Leipzig. The two English universities had large concentrations – forty-six at Cambridge and forty-four at Oxford. In both cases the students were exclusively British, however; universities such as Padua and Paris drew an international clientel. Most important of all was Leyden. Founded in 1575, it existed for only half the span of my study; during that time forty-eight scientists among the 630 studied there, by no means only Dutch students. Add to this the fact that the large numbers at Padua and to a lesser degree at Paris belonged primarily to the earlier years of my study, echoes if you will from the medieval past. In so far as we measure by the number of future scientists of note who studied there, Leyden was the undoubted scientific center of Europe during the 17th century. I would add that the data about universities does not embody the ambiguities found in some of the others. It would bear further analysis according to disciplines.

I also collected information about religion. I have no intention of entering the debate about Calvinism and science, but the data about religious affiliation was generally available and could hardly be omitted. Let me note only that over half of the 630 were Catholic.

My sixth heading is scientific discipline. It seemed essential that I know what sciences the men in the set pursued. When I started, I was determined to keep the number of disciplines I recognized small in order that numbers not become too fragmented. In the end I was unable to do so. The richness of the scientific enterprise during the 16th and 17th centuries, a richness I admit to not having fully appreciated, forced me to expand the list, and I ended by including forty-three different disciplines. A number of these can be grouped into larger categories – mechanics, optics, and the like into physics; anatomy, physiology, and the like into medical sciences; and so forth. The categories are not exclusive; in the dBase file I have made room to list three primary disciplines and three subordinate ones, and a significant number fill all six slots, in itself a datum worth knowing.

I come then, after the various categories of basic information, to one of my central questions, means of support. Again the categories are not exclusive. I have arranged to list as many as three primary sources of support and three secondary, and quite a few of the scientists require all six. Only a small minority drew their support from a single source. Some of the categories are quite small. Fifteen were artisans (not a highly homogeneous category), twelve engineers, seven lawyers. Twenty-one were merchants, again a category broadly defined, in this case without concern for the precise meaning of the word "merchant". It includes men such as Georgius Agricola, who made a fortune speculating in mining shares, John Graunt, a haberdasher, and Otto von Guericke, who ran a brewery. Three groups that I chose to list separately because of their special relevance to science, apothecaries (20), publishers (22), and instrument makers (11), bear some analogy to merchants, though each of them is heavily populated with men we would not readily think of as merchants. Galileo and Torricelli, for example, both earned some of their support by making and selling instruments. Four of the "merchants" appear also in one of these three additional categories, reducing the total number involved.

The category of merchant is one of the places at which I begin to verge into questions of causation as existing literature has defined them. Compare the number of merchants with the number of ecclesiastical appointments, 131 (or more than a fifth), and the number who held governmental positions, again broadly defined to include all levels of government, 198 (not far short of a third). 224 (or about three-eights) drew at least part of their support from a medical practice. 123 had personal means; I am convinced that personal means are underreported, but I have not wanted to guess. 250 (about two out of five) held academic positions. Almost no one drew support from an academic appointment alone; it was almost always combined with other sources, frequently a medical practice, and frequently patronage. For all that, as I suggested above, the number is much larger than I had been led to expect. The largest category of all was patronage, with 272 (well over two out of five). I have not considered isolated gratuities, in return, say, for the dedication of a book, as a significant source of support, even under my category of secondary support. I have in mind rather continuing patronage that provided an enduring source of support; Galileo's position as mathematician and philosopher to the Grand Duke is a familiar example.

My eighth heading is patronage itself. Under means of support, patronage is one category. Under this heading I am concerned with the sources of patronage, regardless of its amount or its continuity. A recent article on patronage argues that when a service was performed, as in the case of a personal physician to a king, we should consider the relationship as employment rather than patronage. This seems entirely mistaken to me. Under my heading of support, I have listed personal physicians to monarchs and wealthy individuals in the category of patronage; usually these physicians maintained an ordinary practice as well and appear also in that category. Patronage has been described as a lop-sided friendship, characterized by an exchange of benefits different in kind. Thus the dedication of a book conferred prestige; it was almost always repaid in one way or another, usually with cash or an object of value such as a gold chain. Viewed from this perspective, a personal physician appears no different than an artist or a musician, who also performed a service. No one questions treating artists and musicians as the clients of patrons. As I have become increasingly convinced of the pervasive presence of patronage throughout early modern society, it has seemed far better to me to define it in broad terms apt to capture its full extent. Patronage for me, then, includes a broad spectrum that stretches from a permanent pension such as Galileo enjoyed to the conferring of a knighthood, which is the most tenuous form of patronage I have included.

When patronage is defined in these terms, only sixty-seven (or only about one-tenth of the 630) appear in the exclusive category, none. That is, virtually everyone was involved, one way or another, with patronage, and most of them had multiple patrons. Here is one of the places where the ambiguity of categories most troubles me. Among others, I use the categories of court, aristocrat, governmental official, and ecclesiastical official. Consider Cardinal Richelieu, who along with much else patronized some natural philosophy. I could make a case for listing him under any of those four categories. Although I have attempted to distinguish when a man patronized as an agent of a court, as a governmental official, or as an aristocrat, the information available is often sketchy and the distinctions are always tenuous. Suffice it to say that my numbers indicate that 373 of the group (about three out of five) were patronized by a court, 243 (about two out of five) by an aristocrat, 101 by a governmental official, and 184 by an ecclesiastical official. Compare those numbers with the number patronized by a merchant (still defined broadly), 21. As I said, it is difficult wholly to avoid the question of causation in the form that it has been defined. I cannot see how to reconcile this data with well known assertions about modern science and capitalism.

My ninth heading is technological involvements. I have long maintained that in the 16th and 17th centuries science had little connection with technology. I was surprised then to find that only 155 appeared in the exclusive category, none. That is, more than three-quarters of the 630 were involved in some technological enterprise. Incidentally, Francis Bacon appears among those in the category, none. Note that my categories make no effort to measure the extent of involvement, and my statement is in no way equivalent to saying that three-quarters of the scientific activity during this period was directed toward technological application. Near the end of his career, Galileo received a summons from the Grand Duke's government to give an opinion about a flood control project west of Florence. In contrast, his student and disciple Benedetto Castelli devoted his life to such matters. Both appear equally under the category, hydraulics. Castelli suggested to Galileo the method of projecting the image of the sun through a telescope onto a screen in order to observe sun spots. In contrast, Galileo developed the astronomical telescope, a microscope, a thermoscope, and a calculating device he called the geometric and military compass. He also contributed significantly to the precision clock. Both appear equally under the category, instrumentation. It is clear from those examples that only the category, none, is exclusive; Simon Stevin and Philippe de la Hire appear in seven different ones.

The figure of seventy-five percent demands further analysis. Roughly three-eighths of the 630 scientists were physicians, and I have treated medical practice, as surely I had to, as an application of scientific knowledge to practical utility. If the physicians are excluded, the percentage of the remainder that involved themselves in some technological enterprise inevitably drops. Nevertheless, about sixty percent of the truncated community had some technological involvement. This remains a much larger figure than I had expected.

The figure invites still further analysis. Four of the categories of technology that I have used are quite large – military engineering, 48; hydraulics, 50; navigation, 42; cartography, 86. These four categories had at least two things in common. They were all heavily mathematical, and like medicine, an even larger category as I have indicated, none of them represented a new undertaking during the Scientific Revolution. In all these cases the application of science, or to be more precise, mathematics, to technological practice reached some distance into the past.

Consider cartography. If we accept portolan charts as maps – at the very least they were first cousins to maps – a new beginning in Western cartography dates back to the late 13th century. The origin of the portolan charts remains a mystery, and it is not clear that science or scientists had anything to do with them. Terrestrial maps, by which I mean a representation of some territory to scale, so that medieval mappae mundi do not count, began to appear at the very end of the 15th century. From the beginning scientists were central to the enterprise. The most developed scientific technology during the 16th and 17th centuries, in my opinion the first truly scientific technology, was cartography. Every major name in the history of cartography would appear in the DSB for other reasons if he had not done any cartography. I think of Reiner Gemma Frisius, Willebrord Snel, Philippe de la Hire, Jean Picard, the two Cassinis, and other lesser ones. All of the important steps in the development of a scientific cartography, such as the method of triangulation, the determination of latitude by celestial observation, the determination of longitude by means of the satellites of Jupiter, came from these men. Any person known to be skilled in mathematics was apt to find some chore in cartography thrust upon him. For the 630 as a whole, about one out of eight engaged in some cartography. If we eliminate the physicians, who did very little cartography, the figure was more than one in five.

In 1949 Benjamin Farrington published a book with the title, *Francis Bacon, Philosopher of Industrial Science*. It does appear to me that the image of Francis Bacon and his talk about useful knowledge, along with the image of the industrial revolution, has dominated discussions of science and technology during the Scientific Revolution. A third idea that

Farrington did not express in his title though it informs the whole book, the proposition that the Scientific Revolution constituted a new beginning in the application of science to practical utility, equally influences the discussions. None of these notions is supported by my data. Hydraulics and navigation were certainly economic, though one could hardly call them industrial, and it is not clear that military engineering and cartography can even be considered as undertakings primarily related to economic enterprise. Other categories under technological involvement are more related to industrial science. All of them are much smaller. Twenty men developed some new mechanical device, though a number of these, such as Stevin's wind-driven carriage, were not applied to any purpose even remotely industrial. Sixteen applied chemistry to some practical use; nine furthered metallurgy. Let me add that only ten among the 630 undertook to improve agricultural practice. More than the size of the four categories I cited, the traditions antedating the Scientific Revolution seem critical to me. There is no doubt that my data has convinced me that we need to approach the whole issue of science and technology in a way different from that of the past. Every historian of the Scientific Revolution ought to know something about cartography; in fact very few do. Until very recently I was among those who did not.

There is one other large category of technology that I have not discussed, instrumentation. 144 scientists from the set of 630, nearly one out of four, developed some new instrument or technique. They appeared in every domain of science, and were not confined to the well known optical devices and clock. I think of Jan Swammerdam's fine scissors for dissecting insects and Frederick Ruysch's method of preserving anatomical specimens. If anyone is in doubt, this efflorescence of instruments and techniques across the spectrum of scientific disciplines, a phenomenon without serious antecedent during preceding centuries, ought to convince us anew that there was indeed a Scientific Revolution during those years. Although a new industry of instrument making came into being, the primary beneficiary of this application of science to utility was science itself.

Tenth and finally, I have a heading for membership in scientific societies. It is not clear to me, I must admit, what use the dBase counts of membership of the Royal Society, the Académie des Science, and the like will be, though I could not imagine omitting that readily available data. On the individual reports, however, I have collected extensive information about informal groups and networks of correspondence, and they seem of immense importance for the formation of true scientific communities. Increasingly I am amazed at how little use historians have made of the correspondence of scientists in the 16th and 17th centuries. To the best of my recollection I have seen exactly one reference to the edition of Mersenne's correspondence in my whole career of reading about science during that age. I need to state that I have not used the Mersenne correspondence any more than others have. In Galileo's correspondence, which I have studied, we can trace the formation of a circle of his young followers. This circle seems more important for the future development of the scientific community than the Accademia dei Lincei, which is studied at length and cited times without number as the first modern scientific society.

I am not unhappy with the body of information I have collected. To me it seems to offer a path around the ideological morass in which studies of science and society appear to be mired, a path leading toward the solid ground of history.

*Indiana University*

STILLMAN DRAKE

# THEORY AND PRACTICE IN EARLY MODERN PHYSICS

From the time of Aristotle to that of Galileo, physics had remained pure theory unmixed with knowledge gained from experience. Mechanics as a mixed science being regarded as subordinate and inferior to physics, the concept of a domain for the practice of physics did not exist until the seventeenth century. The practice of astronomy existed separately from the science of cosmology, and a certain tension between them held ever since Hipparchus and Geminus. A similar tension had held since antiquity between music theorists and practitioners of music. The same is true of architecture and hydraulics. But physics, by Aristotle's definitions, had no practical component.

A principal mark of physics after the time of Galilleo was that it became a useful science, a contradiction in terms for strict Aristotelians. In order for this to happen, elements of utility and of applicability to experience had to come into physics from some source other than existing natural philosophy. Yet for a long time no other precursor discipline for physics has won favour from historians. The required additional source for utility in physics could have been music, architecture, or hydraulics. For Galileo it appears to have been all three, at various times.

Galileo achieved the correct statement of the law of fall – the times-squared rule for distances fallen from rest, set forth in his last book – though there is still some contention among historians whether he adequately verified it in practice. There is even more contention about the process by which he arrived at the law of fall. A number of historiographic styles have been applied by historians to early modern physics over the past half-century. Before reviewing the variety of their approaches. I shall first state briefly the ten steps taken by Galileo for this fundamental discovery. Galileo measured distances in a unit he called the *punto*, equal to 0.94 mm. His timings were based on the weight of water flowing in a fine stream, from which he derived a unit called the *rempo* fo 1/92 second. His surviving records show that he was correct within 3 units of either kind.

1 Timing of a fall of 4000 *punti* as 1,337 grains of flow.
2. Timing of a fall of 2000 *punti* as 903 grains of flow.



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 15-30.

3. Finding 870 *punti* as the length of a pendulum making a quarter-swing during flow of 1,337/2 grains weight.
4. Finding time of a pendulum of 2 × 870 *punti* as 942 grains.
5. Tabulating lengths and times of pendulums.
6. Adoption of 16 grains flow for one *tempo*.
7. Finding the mean proportional of 118 *tempi* (2 × 942 grains) and 167 *tempi* (2 × 1,337 grains) to be 140 *tempi*.
8. Verifying the mean proportional rule with a 30-foot pendulum (9,843 *punti*).
9. Use of 942/850 as ratio of times for pendulum and fall.
10. Calculation of 48,143 *punti* for fall in 280 *tempi*.

In the historiography of Koyré no such detailed steps to the discovery of the fall law could even be imagined, for experimental measurement is absolutely denied to Galileo. Nor did Koyré even attempt to formulate a plausible reconstruction of the discovery of the law of fall. That historiography was nevertheless adopted by a whole generation of historians of science. Even a brief glance reveals various prejudices attributed to Galileo when it would be just as logical to ascribe corresponding preconceptions to the historians themselves. Instead, I shall assume no preconceptions but limit myself to characterizing the images of Galileo that emerge from their historiographies.

Around 1900 Ernst Mach offered a picture of Galileo as a modern, positivist scientist. For Mach, Galileo recognized from casual observation that speed of fall increases during fall and then undertook a mathematical analysis to determine whether speed was a function of time or of distance. Having settled on time, Galileo then derived a relation he could test experimentally. By rolling balls down an inclined plane and timing motions with a stream of water, Galileo was able to confirm the relation he had derived, according to Mach.

In contrast, Pierre Duhem soon attributed most of the achievement to medieval predecessors of Galileo. In Duhem's picture, there seems little left for Galileo to do beyond elaborating and extending results already obtained by Leonardo, Cardano, and Benedetti. Duhem made no reference to Galileo's experimental activities.

By the late 1930's the chief contention seemed to be whether the period of Galileo had witnessed any genuine novelties in science, or whether the scientific revolution was merely the culmination of attacks on Aristotelian principles that had been accumulating over centuries. It should be noted that the field of history of science was still largely in the hands of amateurs, scientists with an interest in the history of their subjects, and historians of philosophy who were curious abount the influence of men like Galileo and Newton on philosophers of their own time.

Alexandre Koyré came to the history of science from the history of philosophy. His trilogy on Galileo's role in the seventeenth century, *Etudes Galiléennes* was published in 1939. Koyré dismissed the major part of medieval influences and associated Galileo explicitly with the birth of modern physics. That event was for Koyré nothing short of a complete change in point of view and style in science; for him, Galileo's achievement consisted in the geometrization of motion. Instead of looking to Galileo's immediate predecessors, Koyré leapfrogged him back to ancient Greece, making of him a Platonic mathematician who extended the statistical researches of Archimedes into the realm of motion. Of course for Koyré, this decision by Galileo to mathematicize nature must have had a purely intellectual origin with no taint of utility or experience.

So profound was Koyré's conviction of the exclusively intellectual nature of the scientific revolution that he could not (as Duhem had done) merely neglect Galileo's experimental activities. He seems to have felt forced to *deny* them – to deny that any such thing *could* have contributed to the scientific revolution, and to deny that Galileo could ever have performed them.

In denying that any of the "numerical data invoked by Galileo relate to measurements actually made," Koyré heaped praise upon him, writing: "I do not reproach him on this account; on the contrary, I should like to claim for him the glory and merit of having known how to dispense with experiments (shown to be nowise indispensable by the very fact of his having been able to dispense with them); yet the experiments were unrealizable in practice with the facilities at his disposal."[1]

Now, our discipline began to become professionalized at the end of World War II. In the United States, Great Britain, and France, a growing number of scholars determined to devote their careers to the history of science. For a couple of decades the scientific revolution of the seventeenth century commanded a large share of their interest. As they searched for an appropriate style of approach, many of them responded enthusiastically to Koyré. He seemed to them to have laid a foundation of erudite scholarship that would enhance the respectability of their nascent discipline.

Koyré's historiography swiftly became widespread. As taught by the first generation of professional historians of science like I. Bernard Cohen and A. Rupert Hall, the Koyrean interpretation began to be elaborated by their pupils. William Shea, Winifred Wisan, Raymond Fredette, and Ronald Naylor have delved more deeply into particular aspects of the Galilean corpus and confirmed and extended the picture of Galileo as a mathematizing intellectual intent on overturning the Aristotelian natural philosophy still current at his time.

Concurrently, several older scholars were reviving a form of the continuist thesis of Duhem. Marshall Clagett in his long study of medieval manuscripts on mechanics and on motion tried to fit Galileo into the tradition of the Paris and Oxford schools of natural philosophy. William Wallace has demonstrated that Galileo's early writings on natural philosophy are closely related to lectures given by Jesuits at the Collegio Romano. Implicitly or explicitly, Clagett and Wallace sought to show that Galileo's mature work derived in a linear way from his early work.

While these and other scholars were following trails blazed by Koyré or Duhem, still others were disturbed by the picture of Galileo that was thus being drawn. It ill fitted the lively and energetic person indicated by biographical studies of mine. That picture seems not at all consonant with either Koyré's intellectual Galileo or with Duhem's medievalist Galileo. However, it did fit with the counterattack by the young scholar Thomas Settle against Koyré on experiments.

In order to test Koyré's rejection of Galileo's published experiments, Settle built apparatus to re-enact Galileo's findings with the inclined plane. Without modern techniques, Settle found that the success reported by Galileo was easily obtainable in the way he had described. This was in 1961.

In 1972 a Guggenheim fellowship gave me the opportunity to study at first hand the working papers on motion that are preserved almost intact at Florence. Most of those were transcribed and published by Antonio Favaro around 1900, but he omitted pages bearing few or no words, but only diagrams and calculations. Arranged in chronological order those show step by step the measurements made by Galileo, confirming Settle's work and extending it until the discovery of the law of fall and the pendulum law in 1604 no longer present any puzzles.

With this background it is appropriate now to outline the picture from which the role of practice in Galilean physics can be fully understood.

The concept of science itself was defined by Aristotle as an understanding of natural phenomena in terms of causes hidden from our senses. Greek philosophers had begun that enterprise the word "physics" to designate the science of nature. He later examined its principles in a book he called "first philosophy," but which was renamed "metaphysics" by his successors. That, with his books on physics, on the heavens, and on meteorology, continued to form the essential basis of natural philosophy throughout the Middle Ages, and until nearly the end of the Renaissance.

From the very beginnings of universities, mainly in the 12th and 13th centuries, Aristotelian natural philosophy dominated all education in science. Astronomy was usually taught by professors of mathematics, not philosophers. Whether astronomy was *science* under Aristotle's definition is questionable, for about 150 B.C. the Greek astronomer Hipparchus showed that the earth cannot be at the exact center of the sun's apparent motion. A compromise, now attributed to Geminus, was soon reached by which astronomers would refrain from considering causes of celestial motions, contenting themselves with framing mathematical hypotheses to fit with their observations and leaving causal explanation to philosophers. The compromise worked splendidly until the Copernican revolution.[2]

Astronomy without causal explanations was philosophically a merely practical discipline; not, strictly speaking, a science at all. While recognizing the existence of knowledge gained through practice, Aristotle explicitly excluded it from truly scientific understanding, or νπιστεμε. For practical knowledge he reserved the distinguishing name, τεχνε. The sharp distinction between science and the practical arts was still maintained by scholars during the time of the scientific revolution, and not astronomy but cosmology constituted for them the science of the heavens.

But beginning from the age of Kepler and Galileo, physics and astronomy were combined in one unified science in the modern sense of the word, while the ancient philosophical separation of physics from useful knowledge vanished from the scientific scene. The meaning of the word "science" before the 17th century was no longer the same after Sir Isaac Newton published his *Principia* in 1687. From the standpoint of semantics that explains why it has been historically unfruitful to look for the roots of the modern physical sciences only in the principles of ancient and medieval philosophy.

Of course, natural philosophy by no means ceased to dominate science in the universities, which were conservative institutions from the very beginning. And even outside the universities, where most of the work was done in developing the new sciences, a kind of counter-revolution was led by René Descartes. He offered a new natural philosophy to replace that of

Aristotle, and at the same time to remedy a serious defect in Galileo's physics, as Descartes saw it, because it neglected causal explanations. Writing about Galileo's new science of motion. Descartes held that to be built without foundations, because it did not start from the *cause* of motion.[3] Newton, on the contrary, saw in it the anticipation of his own laws of inertia and of force, the true basis of modern dynamics.

It was the Dutch enginner Simon Stevin who was the first to put to experimental test G. B. Benedetti's proposition that speeds in fall are not governed by the weights of the falling bodies.[4] Stevin's tests, from a height of thirty feet, were conducted in 1585–86 and published in Dutch four years before equal speeds in fall were exhibited by Galileo from the Leaning Tower of Pisa.

It is an illuminating fact about the state of physics in the latter half of the 16th century that neither Benedetti, who first published his demonstration of equal speeds in fall in 1553, nor any of his supporters or adversaries in this matter over the next three decades, appears to have put his innovative conclusion to actual test, which was an easy and seemingly obvious thing to do. The question was put to nature not by the challenged Aristotelian natural philosophers, but by the mathematical physicist Stevin. Galileo, then a young man who had just completed his years as a student at the University of Pisa, was probably still unaware of Benedetti's proposition. He had reached the same conclusion from the same book that had inspired Benedetti long before, a work on raising sunken ships by the mathematician Nicolò Tartaglia, first printed at Venice in 1551. On the very first page, Tartaglia had remarked that speeds of sinking through water are as the densities of the materials, from which Benedetti drew his idea of equal speeds through air by all bodies of the same material.

While Galileo was completing his final book, in 1636–7, a treatise on music which incorporated critical discussions of the newly emerging physics was just being published at Paris – Marin Mersenne's *Harmonie Universelle*. In that treatise, and in his later works, Mersenne did more to propagate emerging new sciences of acoustics, pneumatics, and ballistics than anyone else of his time, though he is remembered mainly in music. Mersenne's own original contributions to science were modest. It was chiefly as spokesman and translator of Galileo in France, and as friend and loyal supporter of Descartes, that Mersenne furthered the spread of modern sciences, by means not only of books but of voluminous correspondence with savants all over Europe. Lacking the flair for mathematics shared by Stevin, Kepler, Galileo, Descartes and Huygens, Mersenne distinguished himself as a tireless, original, and resourceful experimentalist, at first in the field of musical acoustics and subsequently in physics. Skill in the design and conduct of experiments was replacing speculative philosophy as a guarantee of correct analyses of nature.

Here I should say that the word *experiments* is too broad in scope to throw light on the historical origins of recognizably modern physics, and it is also out of place when we turn to astronomy. The crucial word is *measurements*, a particular kind of experimental activity that was pioneered by Galileo in the study of natural motions; that is, motions undertaken by heavy bodies spontaneously upon simple removal of contraints. No one is known to have attempted measurements of natural motions up to the time of Galileo, even rough measurements. His working papers reveal the *punto*, his unit of distance in 1604 to have been no more than 0.94 mm, a fact from which a new historiography of physics took its rise in 1973.

Mersenne took up experiments described by Galileo and added observations and measurements of his own. Actual measurements of motion had no place in Aristotelian natural philosophy, since they could not reveal hidden causes behind the phenomena. Still less could careful measurement have had any place in the philosophy of Plato, who forbade careful attention to sensible phenomena as a potentially misleading distraction from the archetypal world that he believed superior to the changing world of sensible experience. Archetypes became the favored study of Kepler, but Galileo spoke of them only once, in a letter written in 1633, to reject them.[5]

Like Kepler, Mersenne had a lifelong interest in philosophy, hoping by that to explain the source of truth in science. Unlike Kepler, however, Mersenne was less impressed by speculations of ancient philosophers than he was by some novel ideas of his own contemporaries, especially those of Descartes, with whom he often corresponded on matters concerning science and philosophy.

Like Galileo, Johann Kepler, who founded modern astronomy, received no support from any recognized philosopher of his time, which coincided very nearly with that of Galileo. And although Kepler (unlike Galileo) was not born into a family of musicians, he was unusually well-informed in the classical musical theories, which were entirely mathematical. Kepler became an enthusiast for possible applications of harmonic theory to the Copernican astronomy. His very first book embodied a scheme of the celestial spheres circumscribed around the five

Platonic solids, nested in a certain order around the central Sun.[6] His later revolutionary discovery that planetary orbits are not circular, but elliptical – marking the veritable beginning of modern astronomy – failed to dim Kepler's earlier enthusiasm. He saw elliptical orbits as relieving the music of the spheres from dull monotony. Ellipses produced scale-passages and chords to replace the sustained tones that would inevitably result from perfectly circular motions.

That Kepler's debt to music in science was different in kind from that of Galileo resulted from the fact that it stemmed from musical theory, while Galileo's came from musical practice. That best shows why, in my opinion, the birth of modern science cannot be fully explained without considering the role of music in it.

Although Kepler was indebted to music for his cosmological schemes, he was hardly less deeply influenced by philosophy, and particularly by the Platonism which conferred on mathematics the highest rank of all among the sciences Galileo differed. That is hardly surprising when we recall that Galileo's contributions to astronomy were chiefly observational, whereas Kepler's were entirely theoretical. Observation does not require a philosophy, as theorizing does.

Theoreticians classified music as one branch of mathematics, rooted in arithmetic. In classical Greek mathematics there exists an unbridgeable gulf between arithmetic, which involves only the discrete, and geometry, which involves also continuous magnitudes. Astronomy being the branch of pure mathematics that in classical times belonged with geometry, Kepler's linkage of it through music with arithmetic contradicted the ancient separation between that and geometry.

Like musical practice, observational astronomy was hampered by an ancient tradition – that the heavens, being perfect, could have no motions that were not perfectly circular motions, and that celestial bodies must likewise be perfectly spherical in shape. In 1609 Kepler published his discovery that planetary orbits are elliptical, and the next year Galileo announced his new telescopic discoveries. Discovery of mountains and craters on the moon met with more open hostility from philosophers than even the finding of new planets, as Galileo called Jupiter's satellites. After the two-pronged attack of 1609–10 by Kepler and Galileo, the ancient world-view was doomed to collapse, though not without a struggle.

Stevin's original contributions to mathematics, both pure and applied, are less known but no less important to science than the analytical geometry of Descartes. It was Stevin who in 1585 first narrowed the classical gulf between the discrete and the continuous in mathematics by his invention of decimal fractions.

Of outstanding fundamental importance was Stevin's title for the first chapter of his *L'Arithmetique*, published in 1585. There he stated that one is a number, contradicting the definition of number by Euclid as "multitude of units." The unit itself could not be a number under that definition, though Stevin did not offer a new definition of number to replace it.

The supposed irreconcilability of any discrete and countable quantities with all continuous and infinitely divisible magnitudes and their ratios was, of course, theoretical, and of no practical concern. That is why this tradition holds the key to the musical dispute between Vincenzio Galilei and Gioseffo Zarlino, a quarrel anticipated in Greek antiquity by the position of Aristoxenus in opposition to classical arithmetical musical theory. No matter what the mathematicians said, the ear of a musician can accurately divide musical intervals in ratios that cannot be expressed in terms only of the numbers by which things are in fact counted.

The practical inadequacy of arithmetic alone was also the key to the new science of motion created by Galileo. For that reason I have stressed this close relation between the dead hand of theory that held back progress in music until the age of Galileo's father and that which delayed the birth of modern physical science until the age of Galileo.

Stevin, along with Nicolò Tartaglia, Benedetti, and Galileo, was a principal founder of modern hydrostatics and of theoretical as well as practical mechanics. Had it not been for his publishing chiefly in Dutch, Stevin would doubtless have become much more widely known as a pioneer modern scientist than is presently the case. Curiously enough, Stevin took the position that Dutch was the only language fully suited to the science of nature,[7] because it allowed the coining of new words whose precise meanings would be clear at once to others. In his treatise on music, it was to lack of the Dutch language that Stevin ascribed the failure of all ancient Greek writers to arrive at a fully correct musical theory. But Stevin himself had also a preconception – that all mathematics must in principle be ultimately reducible to the numbers that are used in counting, assumed by Arabs who garbled in translation the Euclidean general theory of proportion for continuous magnitudes.

Whether mathematics is in fact so reducible is completely irrelevant to the practice of music, and to useful science, though until the age of Galileo that was not perceived. Even today this preconception tends to cloak the refutation of medieval impetus theory that was brought about by Galileo's mathematical physics. His new physics owed its origin to two Euclidean definitions, those of "having a ratio to one another", and of "same ratio" as applied to mathematically continuous magnitudes. The first of these was omitted, and the second became hopelessly garbled, in the standard medieval Latin translation of Euclid's *Elements* taken from Arabic (and not authentic Greek) texts. Neither definition was entirely re-established until 1543, and at first was limited to the Italian translation of Euclid's *Elements* by Tartaglia.[8] As a result, the Italians enjoyed a half-century head-start over the rest of Europe in the creation of recognizably modern mathematical physics, most especially Italians who could not read Greek or Latin; for in the universities no attention whatever was paid to Tartaglia because he had not published in an academically respectable language.

It is clearly as a result of overlooking the mid-16th-century revival of Euclidean proportion theory that historians of science still imagine that recognizably modern science must have come from speculative philosophy. As to that, Galileo sarcastically asked, "What has philosophy got to do with measuring anything?"[9] His use of measurements made with all possible precision as the main basis of his new science required such measurements to be subjected to a mathematically rigorous theory of ratios and proportionality, and that had been non-existent in Europe from the fall of Rome until 1543. As Tartaglia said on the title-page of his translation, it was made in order to put into the hands of any person of average intelligence the whole body of mathematical knowledge. Nothing like that was ever the intention of ancient or medieval natural philosophers, whose monopoly on science ended with the invention of printing from movable type and its early 16th-century sequel, the first appearance of inexpensive books in living languages.

Astronomy already had a two-millenium history of accurate measurements of actually observed motions before the first known measurements of pendulums, falling bodies, descents on inclined planes, and projectile motions were made by Galileo in 1604–08. Now, by that time, a profound revolution in musical practice and theories of music was already well under way, one that seems to have originated mainly in resentment of restraints put upon the practice of music by long-accepted theories of musical consonance. Ancient tradition decreed consonance to depend only on ratios of the smallest numbers, a metaphysical conception unduly limiting practice that was utterly rejected by Vincenzio Galilei. A closely parallel conception still delayed the rise of modern science, but was soon to be thoroughly refuted by his son Galileo.

As the musicologist Professor Claude Palisca wrote, in 1961:

> By creating a favorable climate for experiment and the acceptance of new ideas, the scientific revolution greatly encouraged and accelerated a direction that musical art had already taken.[10]

It is the other side of that coin of which I am about to speak. Vincenzio Galilei appears to me to have been the first person ever to have discovered a law of physics by experimental measurements involving motion. Late in his long controversy with Zarlino he found that the ratio 3:2 does not hold for the perfect fifth when sounds are produced by *tensions* in strings, rather than by their *lengths*. He published an account of his experiments in 1589,[11] and various circumstances support my belief that those were carried out in 1588. In that year Vincenzio's son Galileo, then teaching mathematics privately at Florence, was probably residing with his parents. In his notes for a treatise on motion written in 1588, Galileo alluded in passing to the motion of a pendulum, a form of "natural motion" as spontaneous descent was called at that time that had escaped attention by natural philosophers generally. Now, Vincenzo's study of tensions in strings required weights to be attached to them, whether hanging freely or suspended over the bed of a monochord, and in either case a pendular motion would be observably imparted to them. It is thus probable that the young Galileo was present at Vincenzio's experimental measurements.

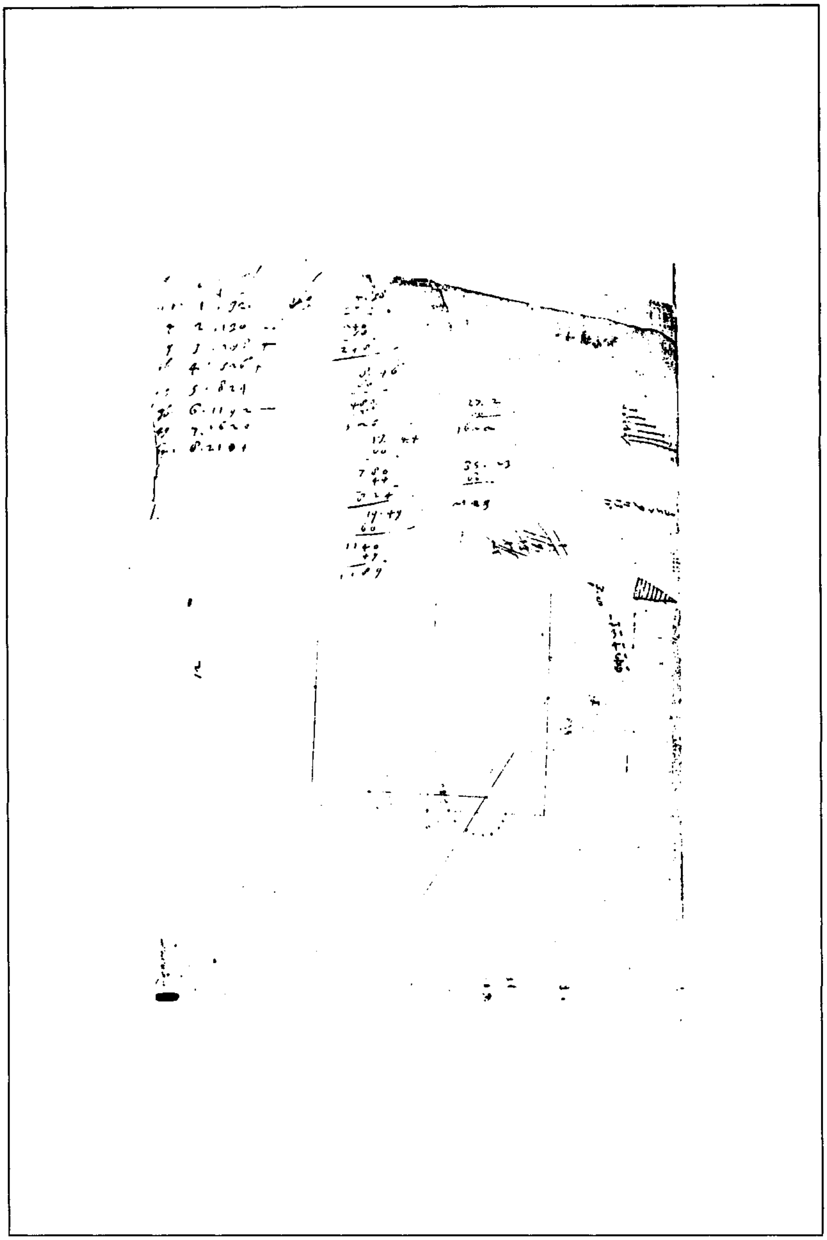
In those years, though Galileo was already in disagreement with some fundamental propositions about motion that were then taught as being Aristotle's – whether or not they were in fact – he did not yet doubt that physics must concern itself mainly with causal inquiry. Years later, in 1602, Galileo's working papers show him to have been making careful experiments with very long pendulums, which led him to a correct and important theorem about motions along inclined planes, and an incorrect conjecture about their relation to motions of pendulums. Within two years he was to discover first, the law of the pendulum; from that, the law of falling bodies; and next, that this same law applied to descents along inclined planes. Galileo's physics from then on concerned only *laws* of nature, not *causal* inquiries of the kind dominating physics for the past 2,000 years. No such revolutionary change in the very nature of science itself would have occurred to Galileo had the musical measurements of his father not first interested him in the motions of pendulums.

Galileo's working papers on motion from 1602 to 1637 still survive nearly complete at the Biblioteca Nazionale Centrale in Florence, though now chaotically bound together in Volume 72 of the Galilean manuscripts. Those that bear theorems, solutions of problems, or enough other words to form one complete sentence or more, were transcribed and published in the definitive Edizione Nazionale of Galileo's works around 1900. It happened, however, that Galileo's experimental measurements, being recorded on pages with few or no words, had gone unnoticed by historians of science until recently. Without taking them into account, it was not possible to reconstruct the experiments underlying these papers, and it remained mere speculation to debate how Galileo discovered the law of falling bodies, opening the road to modern physics.

The first page of those notes to be identified and dated was associated with Galileo's discovery of the parabolic trajectory of a horizontally-launched projectile, in 1608. That left still unknown the manner of his discovering the law of fall, achieved no later than 1604. It did, however, give the name that Galileo used for his unit of length in making measurements, the *punto*. A note on another page, written probably in 1605 or 1606, made it possible to convert the *punto* into metric units. It was, to my surprise, less than one millimeter; to two places it was 0.94 mm.

Knowing this unit of length made it possible to reconstruct the uses made of it. In 1975 I published in *Scientific American* an article entitled "The Role of Music in Galileo's Experiments," analyzing a page numbered *f* 107*v*. For a decade or so I regarded that page as the discovery document for Galilieo's law of fall.

A set of calculations in the middle of *f* 107*v* shows how Galileo had arrived at the eight distances he tabulated there. In every case he had first multiplied a number by sixty and then had added a number less than sixty to it, showing that Galileo owned a ruler divided accurately into 60 equal parts, which he called *punti*. His measurements were made along an inclined plane, grooved to guide a rolling ball, and they represented the places of the ball at the end of each of eight equal intervals of time. It was not difficult then to reconstruct the experimental set-up behind the measurements. The plane was tilted by raising one end 60 *punti* above the horizontal. Because it was about 2,000 *punti* long, its slope was 1.7°. At that slope, a ball rolling the full length of the plane will take 4 seconds, permitting eight half-second marks. Calculation shows Galileo's accuracy to have been within 1/64th of a second for every mark except the last, when the ball was moving about a thousand *punti* per second. Interestingly, that



Mss Galileianc vol. 72 *f* 107*v*, reproduced with permission of Biblioteca Nazionale Centrale, Florence was the only measurement that he subsequently altered. His final entry for it was almost exactly correct, as calculated from modern physical equations.

Historians of science have questioned Galileo's ability to have timed half-second intervals accurately to 1/64 of a second. As I reconstruct his procedure, he tied frets around the plane, so that the ball would make audible bumping sounds as it passed over the frets, which were then adjusted patiently until every bump coincided with a note of some crisply-sung marching song.

When Galileo readjusted the lowest fret, he also placed a plus or a minus sign on four other measurements. It is a great nuisance to adjust any fret but the last, because this requires moving all the frets below it; and in any case, the differences were not worth the bother. Galileo being a good amateur musician, my reconstruction plausibly accounts for everything on the page.

Nevertheless, as I eventually found out, it was not *f* 107*v* that was the discovery document for the law of falling bodies. Into the narrow left-hand margin seen on it, Galileo had squeezed the first 8 square numbers, in bluish ink and slightly smeared. That was odd, as everything else on the page is in black ink and in small, neat writing. If Galileo already knew the times-squared law of fall from his work on *f* 107*v*, as I supposed, then it was a puzzle why the square numbers had not been entered at once.

In fact, all that Galileo found from his first experimental measurements was the rule that speeds grow from rest, in *equal* times, as do the odd numbers 1,3,5,7... Since it had been only a rule for speeds that he was looking for, he laid *f* 107*v* aside before he came to discover the times-squared law of fall. But at that time he recognized that if a precise rule could be found by equalizing eight times musically, much more might be learned by accurately measuring brief times, and not just equalizing them. At the end of the page he drew a preliminary sketch for a timing device that he described years later, in *Two New Sciences* of 1638. A bucket of water with a tube through its bottom was hung up. The water flowing during a fall or pendulum swing was collected; that water was weighed, and these weights became his measures of times. His first recorded weighing was 1,337 grains during fall of 4,000 *punti*, about 12 feet. That was Galileo's poorest timing, high by 1/30 second. Next he timed half this fall, at 903 grains weight of flow, correctly within 1/100 second. He then adjusted the length of a five-foot pendulum until its swing to the vertical accompanied flow of water to the previous mark on his collection vessel. His measured length for this pendulum, 1,590 *punti* was exactly correct, as shown by modern calculation. Galileo next concentrated his attention on pendulums, and found the rule that doubling the length quadruples the time. Choosing 16 grains of flow for a new unit, to fit with proportion theory, he named this the *tempo* (= 1/92 second.) He then calculated length and time for a very long pendulum, about 10 meters, and verified his result by hanging and timing such a pendulum, implying the general pendulum law that the times are always as the square roots of the lengths. From that, he found the law of fall, as was seen on the discovery document when that was finally identified as *f* 189*v*.[12] using his fall law, Galileo corrected his one poor timing and turned back to his data on *f* 107*v* to test whether the law of fall held true also for descents on inclined planes. Writing the square numbers on it, he multiplied each one by the distance to his first mark, and saw that those products were almost identical with the 8 distances he had previously measured. Hence the seeming puzzle (that had been pointed out in 1975) vanished. The reconstruction of *f* 107*v* was correct, but that was not the discovery document for the law of free fall. On the actual discovery page *f* 189*v*1, Galileo next found the rule for descents along two planes differing in both slope and length, and verified numerically a theorem he had found experimentally in 1602 – that although the distance is greater, the time is less for motion along two conjugate chords to the low point of a vertical circle than along the single chord joining the same endpoints. Thus was modern mathematical physics born – not of metaphysical principles, or philosophical speculations – but of accurate measurements inspired by those which had already refuted the ancient philosophical theory of musical consonance.

Without Galileo's having been present at his father's musical experiments in 1588, he probably would not have gone on to his own study of pendulum motion. Without musical training, Galileo would hardly have been able to make his very first timings nearly exact. Music played not only a unique, but an essential role in leading Galileo to his new physics, a science of precise measurements, for music is an art demanding precise measurement and exact divisions.

ACKNOWLEDGEMENT

I am indebted to my friend James H. MacLachlan for many suggestions concerning the text.

*University of Toronto*

## NOTES

[1] Galileo's Treatise *de motu gravium*, reprinted in *Metaphysics and Measurement* (London, 1968), p. 75.
[2] Details are given in my 'Hipparchus - Geminus - Galileo', *Stud. Hist. Philos. Science* **20**: 1 (1989), pp. 47–56.
[3] Descartes to Marin Mersenne, 11 October 1638. Translated in S. Drake, *Galileo at Work* (Chicago 1978), pp. 387–8.
[4] Benedetti's proposition is translated in I. E. Drabkin & S. Drake, *Mechanics in Sixteenth-Century Italy* (Madison 1960), pp. 147–53 as it first appeared at Venice in 1553.
[5] Translated in S. Drake, 'A neglected Galilean letter', *Jrnl. Hist. of Astron.* **17** (1987), pp. 93–105.
[6] Johannes Kepler, *Mysterium Cosmographicum*, tr. A. M. Duncan (New York 1981), esp. pp. 85–105.
[7] *Principal Works* ... cited above, Vol. 1 (Amsterdam 1955), pp. 85–65.
[8] *Euclide ... diligentemente rassettato, et alla integrita ridotto ... talamente chiara, che ogni mediocre ingegno, senza la notitia, over suffragio di alcuna altra scienza con facilita sara capace a poterlo intendere.*
[9] S. Drake, *Galileo against the Philosophers* (Los Angeles 1976), p. 38.
[10] Claude Palisca, 'Scientific Empiricism in Musical Thought', in H. H. Rhys, ed., *Seventeenth Century Science and the Arts* (Princeton 1961), p. 137.
[11] *discorso intorno all'opere di messer Gioseffo Zarlino da Chioggia* (Florence 1589; facs. reprint Milano 1933).
[12] Details are given in the second edition of my translation of Galileo's *Two New Sciences* (Toronto, Wall & Thompson, 1989).

IAN HACKING

# STYLES OF SCIENTIFIC THINKING OR REASONING: A NEW ANALYTICAL TOOL FOR HISTORIANS AND PHILOSOPHERS OF THE SCIENCES

A conference on the historiography of the sciences inevitably includes papers that edgily notice the arrangements between philosophers and historians of the sciences. Let us not squander our time in global metatalk. Friendship, collaboration, appropriation and even mutual indifference will work themselves out unaided by all-purpose generalities. Instead I shall introduce you to a new analytic tool and explain two quite distinct uses of it, one by an historian who originated the idea, and one by a philosopher who picked it up. They are complementary and at first sight asymmetric. The historian may conclude that the philosopher's use of the tool is irrelevant to understanding the past, but the philosopher needs the history, for it the tool does not provide a coherent and enlightening ordering of the record, then it has no more place in sound philosophy than any other phantasy.

The analytical tool in question is concerned less with what we find out than with how we find out, less with the content of the sciences than with their methods. Its author is A. C. Crombie, whom I heard lecture on it some thirteen years ago, in 1978 (Crombie 1981, Hacking 1982). It is out of step with present fashion, which teaches us so much about the intricate details of incidents and relationships. It derives from a conception of the entire Western scientific tradition. We cannot help but recall that Spengler too spoke of the "Western style" – so much so, indeed, that an embarrassed translator: "The word 'Stil' will therefore not necessarily be always rendered 'style'." (1926, vol. 1, p. 108, n. 2). Styles get turned for example into souls, as when "die Expansionkraft der abendländischen Stile" (Spengler 1918, vol. 2, p. 55) is translated as "the expansionpower of the Western Soul" (1926, vol. 2, p. 46). Phrases like "style of thinking" or "reasoning" occur naturally enough without such grandiose implications. That is to be expected with a word like "style" that already has so many connotations. For example the cosmologist Stephen Weinberg and the theoretical grammarian Noam Chomsky wrote about "the Galilean style of reasoning in physics, that is, making abstract mathematical models of the universe...." (Chomsky 1980, p. 9, citing Weinberg 1976, p. 28). Both



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 31-48.

authors attributed this idea of the Galilean style to Husserl. I. B. Cohen (1982, p. 49) gave a more detailed account of the same kind of reasoning; he called it "the Newtonian style," a way of combining "two levels of ontology," the mathematical and the measurable. He also said that "Edmund Husserl has written at large concerning the 'Galilean' style, essentially the mode of modern mathematical physics; from this point of view, the Newtonian style can be seen as a highly advanced and very much refined development of the Galilean." All these authors are referring to Husserl (1954, Part 2, Section 9). Husserl certainly wrote "at large," in this section, about Galileo as the discover of a new kind of science, but I don't think he called it, in so many words, "the Galilean style." Indeed his use of the word "style" seems different once again from that of Chomsky, Weinberg, Cohen, Crombie or myself. For example it is used six times on one page (1970, p. 31), but always to refer to a feature of the "empirically intuited world".

Thus even among philosophers and historians the word "style" has experienced many a use or abuse. Looking further afield, it is well known how literary critics have long distinguished a "generalizing" and a "personalizing" use of the word "style." There is a Balzacian style and there is Balzac's style. Equally, in swimming, there is the Australian crawl and freestyle, as opposed to the style of Patti Gonzalez, that can be imitated but is inimitably hers. It is entirely natural to talk of the style of an individual scientist, research group, programme or tradition. Kostas Gavroglu, although taking the word "style" from myself and by derivation from Crombie, has quite legitimately put the word to its personalizing usage, for he contrasts the "style of reasoning" of two low temperature laboratories, and indeed of two men, Dewar and Kaemerlingh Onnes. Crombie and I instead intend something more attuned to Cohen, Chomsky and Weinberg than to Gavroglu. And even if we put aside all obviously personalizing uses of "styles" of thinking there are plenty of generalizing uses in the history or philosophy of science that differ from Crombie's. The most famous instance is in Ludwig Fleck's fundamental book of 1935, subtitled *Introduction to the Theory of the Thought Style and the Thought Collective*. By a thought style Fleck meant something less sweeping than Crombie, more restricted to a discipline or field of inquiry. Nevertheless a thought style is impersonal, the possession of an enduring social unit, the "thought collective." It is "the entirety of intellectual preparedness or readiness for one particular way of seeing and acting and no other" (1979/1935, p. 64.) Fleck intended to limn what it was possible to think; a *Denkstil* makes possible certain ideas and renders others unthinkable. Crombie and I fix on an extreme end of the spectrum of such permissible uses, and accordingly enumerate very few styles of thinking or reasoning. This is partly because our unit of analysis is very large in scope. There are many other units of analysis comparable to Fleck's, and which also deal what it is possible to say. They are thoroughly impersonal, but more restricted in scope, in time and in space. For many purposes they may be, for that very reason, more instructive than something along the lines of Spengler or Cohen or Weinberg and Chomsky. We think for example of Michel Foucault's episteme and discursive formation.

As I have said, I first attended to a very specialized idea of a style of thinking or reasoning after hearing Crombie's talk. Even then he had a very long book manuscript in hand, which has now grown to three thick volumes, scheduled for publication in 1994 (Crombie, forthcoming). It combines a rather profound analysis with an incredibly rich array of citations spanning three millennia, plus dense references to secondary studies – the life-time collection of an erudite.

In contrast to Crombie I speak of "styles of (scientific) reasoning" rather than "thinking" but that is a matter of philosophical taste. Thinking is too much in the head for my liking. Reasoning is done in public as well as in private, by thinking, yes, but also by talking and arguing and showing. The difference between Crombie and me here is only one of emphasis. He writes that "the history of science has been the history of argument" – and not just thinking. We agree that there are many doings in both inferring and arguing. Crombie's book describes a lot of them, and his very title happily ends not with science but with "Sciences and Arts." He has a lot to say about architecture, clock making and the doctrine that "knowing is making." Nevertheless there may still be a touch too much thinking for my pleasure. He gave his prospectus for the book the title, "Designed in the Mind" (Crombie 1988). Does one hear the resonance of Crombie's Koyréan roots? Even my word "reasoning," although it recalls talk, and argument, and all things more public than the mind, does not, I regret, sufficiently invoke the manipulative hand and the attentive eye. Crombie's final title-word is "Arts." Mine would be "Artisan."

But there's more than that in my preference for reasoning over thinking. It recalls me to my roots – I am talking about just what Aristotle called rational, even if my analysis is better suited to the temper of our times than his. "Reasoning" recalls the *Critique of Pure Reason*. Our study is a continuation of Kant's project of explaining how objectivity is possible.

He proposed preconditions for the string of sensations to become objective experience. He also wrote much about science, but only after his day was it grasped how public an activity is the growth of knowledge. Kant did not think of scientific reason as an historical and communal product. We do. My styles of reasoning are eminently public. They are also part of what is required for us to understand what, in the West, is called objectivity. This is not because they are objective, but because they determine what it is to be objective.

Crombie does not expressly define "style of scientific thinking in the European tradition." He explains it ostensively by pointing to six styles which he then describes in painstaking detail. "We may distinguish in the classical scientific movement six styles of scientific thinking or methods of scientific inquiry and demonstration. Three styles or methods were developed in the investigation or individual regularities and three in the investigation of the regularities of populations ordered in space and time." These are (I combine several of his expositions):

(a) The simple method of postulation exemplified by the Greek mathematical sciences.
(b) The deployment of experiment both to control postulation and to explore by observation and measurement.
(c) Hypothetical construction of analogical models.
(d) Ordering of variety by comparison and taxonomy.
(e) Statistical analysis of regularities of populations, and the calculus of probabilities.
(f) The historical derivation of genetic development.

You will see why Crombie divides these into three plus three. This is not quite the contrast of the predictive/experimental and the historical sciences, according to the division which (as Professor Hiebert told us yesterday) is now enforced at Harvard. Crombie includes mathematics among the sciences, which is where they belong, whatever some of my recent philosophical predecessors may have thought. And statistical reasoning is not historical, but belongs among the wider class of scientific work that deals in populations, not individuals. For my purposes the Oxford historian is a more sure guide than the Harvard bureaucrats.

Note that styles do not determine a content, a specific science. We do tend to restrict "mathematics" to what we establish by mathematical reasoning, but aside from that, there is only a very modest correlation between (a)–(f) and a possible list of fields of knowledge. A great many inquiries use several styles. The fifth, statistical, style for example is now used, in various guises, in every kind of investigation, including some branches of pure mathematics. The paleontologist uses experimental methods to carbon date and order the old bones. The "modern synthesis" of evolutionary theory is among other things a synthesis of taxonomic and historico-genetic thought.

It is of great value to have a canonical list of styles descriptively determined by the historian who, whatever his axes, is not grinding any of mine. As a philosopher I then have to discover, from his examples, at least a necessary condition for being such a "style." Moreover we are not bound to accept Crombie's preferred descriptions, nor even to conclude with exactly his arrangements of styles. I shall give some reasons for this and then give two examples.

Crombie offers an account of the "classical scientific movement" and tailors his characterizations to the long period of time in which that movement was formed and firmed up. He tends to leave a given style at the date when it is securely installed. His discussions of mathematics end with revivals of Greek maths by Kepler. His exposition of the first three styles together dries up at the end of the seventeenth century. Only the final style is developed for the nineteenth century, with Darwin being the major figure. But I as philosopher am properly Whiggish. The history that I want is the history of the present. That's Michel Foucault's phrase; I can be as "archaeological," in his sense, as I want, but the selection of an historical object is always made with a present aim in view. Hence I may modify Crombie's list not to revise his history but to view it from a present standpoint.

In the same vein, a quick review of Crombie's (a)–(f) will convince you that it is itself an historical progression, each style beginning later than its predecessor in the list, and Crombie's presentation of each style concluding closer to the present than his descriptions of preceding styles. What strikes me, however, is the ahistorical point that all six styles are alive and quite well right now. I am writing about what styles of scientific reasoning do for *us*. What's important now may be a little different from what was important in the early days.

Taking an opposite tack, we also notice that Crombie cannot have intended his list as an exhaustive list of mutually exclusive styles. Quite aside from any styles that we might properly want to call scientific, and which evolved largely outside the West, there might also be yet earlier styles of "science" found say in records of Babylonian computations. To judge by his exposition yesterday, I do not think that Jens Høyrup would want these to be identified with a mere anticipation of (a). Looking forward rather than backwards, new styles may have evolved after the "classical" events Crombie recounts. More importantly we could use a style that results from the merging of two or more styles. I don't mean that we commonly use more than one style in any modern inquiry, but that there may have evolved a style which is essentially composed of two classical styles.

Now I turn to two examples. As a philosopher of mathematics, I see proof where Crombie sees postulation. His first style emphasizes the Greek search for first principles. It is there that he brings in Greek medicine, with its battle between empirics and dogmatists. It is in discussing (a) that we meet Aristotle – even when he is canonizing what later becomes the taxonomic style (d). All that is right history, putting (a) and its contemporary correlatives first, in their place. Yet there is no doubt that what individuates ancient mathematics now is that we recognize proof and to a limited extent calculation. Indeed Wibur Knorr speculatively orders segments of actual and lost texts by the development of proof procedures. Mathematics has the astonishing power to establish truths about the world independently of experience. That is the phenomenon that so astounded Socrates in the *Meno*, and has so vexed every serious epistemologist of mathematical science ever since – a point particularly present to David Fowler. I will want my account of the mathematical style to help understand that phenomenon. Hence my emphases will be different from Crombie's. Moreover, I would probably add a distinct "mathematical" style (a *), not postulational but algorithmic, not Greek but Arabic and Indian in origin, and I would regard the chief problems of the philosophy of mathematics as arising from the interactions between the Greek and Indians styles.

For another example, the historical distinction between styles (b) and (c) is profoundly important and has to do with the familiar tensions between today's experimenter and theoretician. The former is surely heir to the empiricists in medicine, who insisted that we should never go beyond observables in our descriptions of the course of disease and its cure, while it was the dogmatists who introduced what we would now call theoretical entities that play so major a part in hypothetical modelling (c). Crombie speaks of "controlling postulation" in his summary description of (b), but the postulation is at the level of observables and measurable quantities. It is by and large phenomenal science. Something else began just about the end of the period in which Crombie describes (b) and (c). I call it the laboratory style, characterized by the building of apparatus in order to produce phenomena to which hypothetical modelling may be true or false, but using another layer of modelling, namely models of how the apparatus and instruments themselves work. The relationship between the laboratory style, call it (bc), and styles (b) and (c) is complex. And we still have both (b) and (c) in full play on their own. For all the talk about intervening variables and the like, many of the social sciences operate only at the empirical level of (b). On the other hand cosmology and cognitive science – none other than the chief modern instances of the Galilean style so admired by Weinberg and Chomsky – remain at the level of (c), hypothetical modelling. Those sciences answer to observation but experimental manipulation is impossible. They remain sciences that represent; the laboratory style introduced sciences that intervene. I judge that the laboratory style began about the time that Boyle made the air pump in order to investigate the spring of the air.

It is characteristic of styles that they have popular myths of origin. Crombie's list strikes the right note just because it is not innovative but only codifies familiar legend. How could it be otherwise if one is recapitulating European science from within? There was that legendary moment when, as Althusser put it, Thales "discovered the continent of mathematics" (1972, p. 185). Next in the list of continents is "and Galileo discovered the continent of mechanics." Well, Galileo is everybody's favourite Hero; indeed Crombie's talk about styles of scientific thinking that aroused my interest long ago was about – Galileo, Stillman Drake's paper read yesterday told us that Galileo (by the purest use of style (b)) established the very first experimental and quantitative law. But when Chomsky or Weinberg speak to the Galilean styles, they, like Husserl or Spengler or Koyré have (c) in mind. Galileo is the stuff of myth.

Althusser continued, "and Marx discovered the continent of history." Good myth, wrong man; I much prefer Michel Foucault's retelling as Bopp, Cuvier and Ricardo. Cuvier, as many have noticed, is questionable, but Bopp's philosophy is perfect at the start of the historico-genetic style. As for the fifth style, that too has its legends. "A problem about games of chance proposed to an austere Jansenist by a man of the world was the origin of the calculus of probabilities" or so wrote Poisson (1837, p. 1). I take Schaffer and Shapin's (1986), subtitled *Hobbes, Boyle and the Experimental Life* as setting out the myth of origin for (bc), the laboratory style. Their Hero is, importantly, not a person but an instrument, the apparatus, the air pump. And as for my other supplement to Crombie's styles, (a *), the algorithmic style, the myth is surely that of Abu Jafar Mohammed ben Musa al-Khwarizmi. The Latin translation of his ninth century work gave us the very word "algorism" or "algorithm". His opening sentence can be translated, "And thus spake Algorismi."

It is important that styles are mythic. Also, so continue Althusser's metaphor, that they open up new territory as they go. The algebrizing of geometry, the Arabizing of the Greek, was an essential piece of territorial expansion. Every expansion is contested. It may be that I emphasize this more than Crombie, but that is partly because we can experience present contests. For example: are computer generated concepts and proofs really mathematics? When I was a student I went around with some topologists who would talk and draw pictures and tell amazing stories; today, when I have topological house guests, the first thing they do is set up their Macs in my basement, not calculating but generating ideas to which real-time computation is integral. And I know others who say that my friends have stopped doing mathematics. That's how it is, when a style goes into new territory.

For all these differences in emphases, I do not differ significantly from Crombie either in my individuation of styles of reasoning or in how I describe them. Indeed I take some pleasure in his having a three volume vindication of his canonical list. Without that I would be left with dubious anecdotes and fables. I'm not claiming that I'm on solid non-ideological ground when I resort to an historian for an initial individuation of styles. I am claiming only a certain inter-subjectivity, because his motivations are so different from mine, yet the list he presents so admirably suits my purposes. To use yet another obsolete metaphor, it does cover the waterfront, and provide a directory to the main piers, in a readily recognizable and fairly satisfactory way. Of course it could be the wrong waterfront. Maybe he's describing a once wondrous but now gutted Liverpool, or at any rate a dignified San Francisco that has taken up leisure pursuits like high finance and tourism, harbours that history has passed by. Perhaps I should instead be attending to a bustling container port like Felixstowe or Oakland. But I don't think so. The proof of my confidence that Crombie's list remains germane is, however, not a matter of principle but of the success of the resultant philosophical analysis.

Our differences lie not in the identity of styles or their description, but in the use to which we put the idea. Crombie's advance notice of his book begins:

When we speak today of natural science we mean a specific vision created within Western culture, at once of knowledge and of the object of that knowledge, a vision at once of natural science and of nature (1988, p. 1)

A little later he says that,

The whole historical experience of scientific thinking is an invitation to treat the history of science, both in its development in the West and in its complex diffusion through other cultures, as a kind of comparative historical anthropology of thought. The scientific movement offers an invitation to examine the identity of natural science within an intellectual culture, to distinguish that from the identities of other intellectual and practical activities in the arts, scholarship, philosophy, law, government, commerce and so on, and to relate them all in a taxonomy of styles. It is an invitation to analyse the various elements that make up an intellectual style in the study and treatment of nature: conceptions of nature and of science, methods of scientific inquiry and demonstration diversified according to subject matter, evaluations of scientific goals with consequent motivations, and intellectual and moral commitments and expectations generating attitudes to innovation and change (1988, p. 2).

This is history in the grand manner, an invitation to a comparative historical anthropology of thought. Regardless of interest, philosophical or historical, many of us may be glad that at a time of so much wonderfully dense and detailed but nevertheless fragmented studies of the sciences, we are offered such a long-term project. This is especially so for philosophers to whom the most fascinating current historiography of the sciences is of a type not represented at this conference – the work of the "social studies of science" schools of history of science (called sociology but in fact philosophically motivated history). I mean the strong programme, network theory, the doctrine of the construction of scientific facts by negotiating. Increasingly fine-grained analyses of incidents, sometimes tape-recorder in hand, have directed the history of science towards the fleeting. On the other hand many of my philosophical colleagues take it to the quasi-timeless, as when Hilary Putnam writes of "the ideal end of inquiry." Crombie's styles may seem to be edging off towards the excessively long run, in the way in which Braudel turned the history of the Mediterranean into a study of the climate that, incidentally aided by ship-builders, transformed this country from a forested peninsula into a rockheap. But his intentions are plain, to conduct an historical understanding of that specific vision mostly created around the Mediterranean basin and then in more northerly parts of Europe, "a comprehensive historical inquiry into the sciences and arts mediating man's experience of nature as perceiver and knower and agent [that must] include questions at different levels, in part given by nature, in part made by man." Crombie is well aware of the need to establish the historical continuity of styles across periods of latency, of the need to understand "the intellectual and social commitments, dispositions and habits, and of the material conditions, that might make scientific activity and its practical applications intellectually or socially or materially easy for one society but difficult or impossible for another." He will compare those now familiar items, "the numbers, social position, education, occupation, institutions, private and public habits, motives, opportunities, persuasions and means of communications of individuals" ... and so on: "military context," "rhetorical techniques of persuasion." The grand view need not neglect the fashionable topics of the moment, nor on the other hand ignore philosophical chestnuts like the existence of theoretical entities. That conundrum is described in the mandarin manner you will have noticed from my other quotations, "distinguishing the argument giving rational control of subject-manner from an implication of the existence of entities appearing in the language used...."

I have hardly begun to mention Crombie's historiographic aims; you will have to read the rest for yourselves. How can a philosopher make use of so grand and expansive an idea of a style of scientific thinking or reasoning in the European tradition? First I notice the way in which styles become autonomous. Every style comes into being by little microsocial interactions and negotiations. It is a contingent matter, to be described by historians, that some people with disposable time and available servants should value finding something out. Yet each style has become independent of its own history, remembering it chiefly in myths of origin. Each has become a canon of objectivity, a standard or model of what it is to be reasonable in this or that type of subject matter. We do not check to see whether mathematical proof or laboratory investigation or statistical "studies" are the right way to reason: they are what it is to reason rightly, to be reasonable in this or that domain. Indeed every style of reasoning makes possible, introduces, a great many novelties, including

new types of: objects
evidence
sentences, new ways of being a candidate for truth or falsehood
laws, or at any rate modalities
possibilities.

One will also notice, on occasion, new types of classification, and new types of explanations. One should not have the picture of a style and then the novelties. That is one of the many merits of the word "style." We did not first have fauvism and then Matisse and Derain painting fauve pictures in 1905. The style comes into being with the instances although (as the example of the fauves makes plain) the recognition of something as new, even the naming of it, may solidify the style after it has begun (Crombie notes how self-reflection on styles influences their deployment).

Each style, I say, introduces a number of novel types of entities, as just listed. Take objects. With every style of reasoning there is associated an ontological debate about a new type of object. Do the abstract objects of mathematics exist? That is the problem of Platonism in mathematics. Do the unobservable theoretical entities of the laboratory style really exist? That is the problem of scientific realism in the philosophy of the natural sciences. Do the taxa exist in nature, or are they, as Buffon urged, mere artifacts of the human mind? Are there objects, such as languages, to be understood in terms of their historical derivation, or are they just a way of organizing a mess of complexity on top of the only reality, a postulated innate universal grammar (or is *that* the unreal, imagined by hypothetical modelling?). Are coefficients of correlation, the rate of unemployment, real features of populations or are they products of institutional arrangements of classification and measurement?

Each style of reasoning has its own existence debate because the style introduces a new type of object, individuated using the style, and not previously noticeable among the things that exist. One may run down my list of novelties checking that each style introduces these novelties. I then propose a necessary condition for being a style of reasoning, in the intended sense: each style should introduce novelties of most or all of the listed types, and should do so in an open-textured, ongoing and creative way. Mathematicians do not just introduce a few sorts of abstract objects, numbers and shapes, and then stop; the type, "abstract object," is open-ended once one sees how to reason in a certain way. Note that on this criterion, logic, be it deductive, inductive or abductive, does not count as a style of reasoning. This is as it should be. Crombie did not list them, and no wonder. People everywhere make inductions, draw inherences to the best explanation, make deductions; those are not peculiarly scientific styles of thinking nor are they European in origin.

Given such a necessary condition for being a style of reasoning we can address a question posed by a number of special-interest groups. What are the other styles of reasoning? Historical reasoning? Legal reasoning? Mystical reasoning? Magical reasoning?

I use my list of novelties as a criterion, as a necessary condition for being a style of reasoning. I've mentioned objects; now I shall say a little more about sentences, so well-liked by my fellow analytic philosophers. Each new style, and each territorial extension, brings with it new sentences, things that were quite literally never said before. That is hardly unusual. That is what lively people have been doing since the beginning of the human race. What's different about styles is that they introduce new ways of being a candidate for truth or for falsehood. As Comte put it – and there is a lot of Comtianism in my philosophy – they introduce new kinds of "positivity," ways to have a positive truth value, to be up for grabs as true or false. Any reader who fears too much early positivism should know also that I took the word in the first instance from Michel Foucault, whose influence on my idea of styles of reasoning is more profound than that of Comte or Crombie. I should add for philosophers that this idea of positivity falls far short of what Michael Dummett calls bivalence, of being definitely true, or definitely false. Bivalence commonly requires far more to be in place than a style of reasoning. I am concerned with new ways of being investigated as true or false. As Dummett has well taught, even when similarities in the surface grammar and in possible ways on inquiry may make us think that sentences we investigate using them are beyond question bivalent, closer inspection may make us sceptical.

The sentences that acquire positivity through a style of reasoning are not well described by a correspondence theory of truth. I have no instant objection to a correspondence theory for lots of humdrum sentences, what we might call pre-style sentence, including the maligned category of observation sentences, for example. But I reject any uniform all-purpose semantics. The instant objection to correspondence theories, for sentences that have positivity only in the context of a style of reasoning, is that there is no way of individuating the fact to which they correspond, except in terms of the way in which one can investigate its truth, namely by using the appropriate style. As J. L. Austin showed, that objection does not so instantly apply to for example "observation sentences" in subject-predicate or subject-relation-object form. We want to say many different kinds of things, and I reject the first dogma of traditional anglophone philosophy of language, that a uniform "theory of truth" or of "meaning" should apply across the board to an entire "language." (That is one lesson to draw from Wittgenstein's talk of different "language-games").

The truth of a sentence of a kind introduced by a style of reasoning, is what we find out by reasoning using that style. Styles become standards of objectivity because they get at the truth. But a sentence of that kind is a candidate for truth or falsehood only in the context of the style. Thus styles are in a certain sense "self-authenticating." This is not Hamlet's despondent subjectivism, "nothing's either true or false but thinking makes it so." It amounts at most to this: no sentences of a certain large class are true-or-false but a style of reasoning makes them so. Even this statement induces an unsettling feeling of circularity.

I welcome it. For the remarkable thing about styles is that they are stable, enduring, accumulating over the long haul. Moreover, in a shorter time frame, the knowledge that we acquire using them is moderately stable, quasi-stable as Sam Schweber put it. Our knowledges are subject to revolution, to mutation, and to several kinds of oblivion. Nevertheless, despite the important truths about refutation and revolution that we successively learned from Popper and Kuhn, a great deal of what we have found out stays in place. A few years ago when I published a brief paper about the stability of the laboratory sciences, I could refer only to Schweber and to some work on "finality in science" done by a group in Frankfurt (Hacking 1988). Now the topic is positively trendy. Some of you will have noticed that it has just now been the preoccupation of the correspondence pages of the *Times Literary Supplement* (starting from March 15 and continuing until April 19, 1991).

In respect of stability I wholly endorse one conclusion of the strong programme in the sociology of knowledge. The truth of a proposition in no way *explains* our discovery of it, or its acceptance by a scientific community, or its staying in place as a standard item of knowledge. (Nor does being a fact, nor reality, nor the way the world is). My reasons for saying so are different from those of the early work done in Edinburgh, and more reminiscent of very traditional philosophy. I would transfer to truth and reality what Kant said about existence, that it is not a predicate, adding nothing to the subject. This is no occasion to develop that theme, except to say that anyone who endorses the Edinburgh conclusion, that truth is not explanatory, will want an understanding of the stability of what we find out, and not settle for "because that's the way that the world is." Notice that I am not here coming out against truth or reality or objectivity (as if that made sense). Of course some things are true and others false, of course there is a real world, of course there is a way that the world is. But truth, reality, and so forth don't explain anything. They are very useful ideas, but are too often used for a pointless regress. That is what's right about the misnamed "redundancy theory of truth." Properly used, the word "true" is seldom redundant, but it is a mere consoling placebo when used in some kinds of philosophical explanations.

If we want some understanding of the quasi-stability of some of our knowledge we shall not find it in remarks about science or method in general. Each style of reasoning has its own characteristic self-stabilizing techniques. An account of each technique requires detailed analysis, specific to the style, and vivid historical illustration. Each is a long story. I've three papers, each upwards of fifty pages, about the laboratory, the statistical and the mathematical styles. There is no overlap between them, because the techniques and the history of their evolution is so different. Almost the only thing they have in common is that they enable a self-authenticating style to come into being. This bears also on the individuation of styles. We began with an ostensive explanation, Crombie's list. Then we moved to a criterion, a necessary condition: a style must introduce certain novelties, new kinds of objects, laws and so on. But now we get closer to the heart of the matter. Each style persists, in its peculiar and individual way, because it has harnessed its own techniques of self-stabilization.

This talk of techniques sounds very unfamiliar, but my chief innovation lies in organization. Many of the techniques I describe are quite well-known, but, I claim, inadequately understood. For example, Duhem's famous thesis about how to save theories by adjusting auxiliary hypotheses is a small part of the stabilizing techniques that I distinguish in the laboratory sciences. It has to be augmented by ideas that I owe to recent work by Andy Pickering. He showed that we are concerned with a material self-stabilization involving what I call ideas (which include theories of different types), materiel (which we revise as much as theories) and marks (including data and data analysis) (Hacking 1992a). All three are what Pickering calls plastic resources that we mould into semi-rigid structures. I should emphasize that although I use Duhem, this account does not go in the direction of the underdetermination of theory by data (Quine's generalization of Duhem's remarks). On the contrary, we come to understand why theories are so determinate, almost inescapable.

Likewise my account of the stability of the mathematical style owes much to two unhappy bedfellows, Lakatos and Wittgenstein (Hacking, forthcoming). But we no more arrive at the radical conventionalism or constructionalism sometimes read into the latter's *Remarks on the*

*Foundations of Mathematics* than we arrive at the underdetermination of theory by data.

A happy by-product of this analysis is not only that each style has its own self-stabilizing techniques, but also that some are more effective than others. The taxonomic and the historic-genetic styles have produced nothing like the stability of the laboratory or the mathematical style, and I claim to be able to show why. On the other hand, although Mark Twain or whomever could, in the earlier days of the statistical style, utter the splendid canard about lies, damn lies, and statistics, the statistical style is so stable that it even has its own word that gives a hint about its most persistent techniques: "robust." In the case of statistics there is an almost too evident version of self-authentication (the use of probabilities to assess probabilities). But that is only part of the story, for I emphasize the material, institutional requirements for the stability of statistical reasoning (Hacking 1992b). Indeed if my accounts deserve to be pegged by any one familiar philosophical "ism," then it is materialism. That is most notably true of my account of the laboratory style, so strongly contrasting with the idealism of the "Quine–Duhem thesis" that I incorporate.

This proposed account of self-stabilizing techniques cannot end the story of how a style becomes autonomous of the local microsocial incidents that brought it into being. Its persistence demands some brute conditions about people and their place in nature. These conditions are not topics of the sciences, to be investigated by one or more styles, but conditions for the possibility of styles. An account of them has to be brief and banal because there is not much to say. What we have to supply are, to quote Wittgenstein (1981, p. 47), "really remarks on the natural history of man: not curiosities, however, but rather observations on facts which no one has doubted and which have only gone unremarked because they are always there before our eyes." Wittgenstein also called this philosophical anthropology. The resonance is with Kant's *Anthropologie* rather than the ethnography or ethnology so commonly studied in departments of anthropology or sociology. Crombie's "comparative historical anthropology of thought" is by and large historical ethnology, a comparative study of one profoundly influential aspect of Western culture. Wittgenstein's philosophical anthropology is about the "natural history of man," or, as I prefer to put it, about human beings and their place in nature. It concerns facts about all people, facts that make it possible for any community to deploy the self-stabilizing techniques of styles of reasoning. It is in philosophical anthropology that we slough off the Eurocentrism with which our study began. And, to continue this list of -logies for a moment, a fitting name for the study of self-stabilizing techniques would be philosophical technology. This label does not now carry its meaning on its face, for I am not talking about what we usually mean by "technology," namely the development, application and exploitation of the arts, crafts and sciences. What I mean by philosophical techno-logy is the philosophical study of certain techniques, just as philosophical anthropo-logy is the study of certain aspects of man, epidemio-logy of epidemic diseases.

We have reached, then, a foundational difference between the historian's and the philosopher's use of the idea of a scientific style of thinking or reasoning. Crombie leads us to a comparative historical anthropology (moved, he has also told us, by the experiences of teaching in Japan, and of crossing parts of Asia and its oceans when visiting his native Australia). I invite you instead to consider what I call philosophical technology, a study of the ways in which the styles of reasoning provide stable knowledge and become not the uncoverers of objective truth but rather the standards of objectivity. And when asked how those techniques could be possible at all, I fall back on a few and very obvious remarks about people, of the sort to which Wittgenstein has already directed us. Less all-encompassing histories will provide the social conditions within which a style emerged and those in which it flourished; less ambitious essays in philosophical technology will describe, in a more fine grained way, the ways in which a style took on new stabilizing techniques as it pursued its seeming destiny in new territories.

I began by saying that the philosopher requires the historian. If Crombie's three volumes do not present a coherent ordering and analysis of European scientific practice and vision, then my talk of self-authenticating styles and of philosophical technology would be empty. That is why I called the relation between the history and the philosophy of the sciences asymmetric. The philosopher who conceives of the sciences as a human production and even invention requires the historian to show that analytic concepts have application. After learning from the historian's analysis I turn to a different agenda, which, you will have noticed, summons all the old gang; truth, reality, existence. But also, as is always the case in philosophy, we are directed to a complementary range of entirely new topics.

For all the manifest differences of endeavour between the historian and the philosopher we have this in common: we share a curiosity about our

Western "scientific" vision of objectivity. That is as central a philosophical concern as could be, the core question of Kant's first critique. Crombie's volumes will, I hope, be read in part as an account of how conceptions of objective knowledge have come into being, while the philosopher can describe the techniques, which become autonomous of their historical origins, and which enable styles of reasoning to persist at all. Yet I would not push this division of labour too far. Some of the best new, detailed, work on the idea of objectivity, with which I am acquainted, is by young but established historians of the sciences, such as Lorraine Daston, Peter Galison and Ted Porter. However much the historian may abjure the philosophical topics, old or new, every sound history is imbued with philosophical concepts about human knowledge, nature and our visions of it, an involvement eloquently expressed yesterday by Erwin Hiebert. But aside from central shared concerns, there is a more general predicament that the historian and the philosopher share. Crombie himself is powerfully aware of the reflexive elements of his volumes: he knows that he who describes a certain vision of ourselves and our ecology has that vision himself. And more constraining, although more difficult to come to coherent terms with, philosopher and historian alike are part of the ecosystem that has been transformed by bearers of that vision in their interactions with nature as they saw it.

*University of Toronto*

## REFERENCES

Althusser, L., 1972, *Politics and History: Montesquieu, Rousseau, Hegel, Marx*, London.

Chomsky, N., 1980, *Rules and Representations*, Oxford.

Cohen, I. B., 1982, "The *Principia*, Universal Gravitation, and the 'Newtonian Style', in Relation to the Newtonian Revolution in Science: Notes on the Occasion of the 250th Anniversary of Newton's Death," in *Contemporary Newtonian Research*, Zev Bechler, ed., Dordrecht, pp. 21–108.

Crombie, A. C., 1981, "Philosophical Perspectives and Shifting Interpretations of Galileo," in J. Hintikka *et al.*, eds., *Theory Change, Ancient Axiomatics and Galileo's Methodology. Proceedings of the 1978 Pisa Conference on the History and Philosophy of Science*, Dordrecht, pp. 271–286.

Crombie, A. C., 1988, "Designed in the Mind: Western Visions of Science, Nature and Humankind," *History of Science* **24**, pp. 1–12.

Crombie, A. C., forthcoming, *Styles of Scientific Thinking in the European Tradition: The History of Argument and Explanation Especially in the Mathematical and Biomedical Sciences and Arts*, 3 vols, London.

Fleck, L., 1979 (1935), *Genesis and Development of a Scientific Fact*, T. J. Trenn and R. K. Merton, trans., Chicago.

Gavroglu, K., 1990, "Differences in Style as a Way of Probing the Context of Discovery," *Philosophica* **45**, pp. 53–75.

Hacking, I., 1982, "Language Truth and Reason," in M. Hollis and S. Lukes, eds., *Rationality and Relativism*, Oxford, pp. 48–66.

Hacking, I., 1988, "On the Stability of the Laboratory Sciences," *The Journal of Philosophy* **85**, pp. 507–14.

Hacking, I., 1992a, "The Self-Vindication of Laboratory Science," A. Pickering, ed., *Science as Practice and Culture*, Chicago, pp. 29–63.

Hacking, I., 1992b, "Statistical Language, Statistical Truth and Statistical Reason: The Self-Authentication of a Style of Reasoning," E. McMullin, ed., *Social Dimensions of Science*, Notre Dame, Ind., pp. 130–157.

Hacking, I., forthcoming, "Radically Constructivist Theories of Mathematical Progres," *Iride*.

Husserl, E., 1954, *Die Krisis der Europäischen Wissenschaften und die Transzendentale Phänomenologie: Eine Einleitung in die Phänomenologische Philosophie*, Den Haag.

Husserl, E., 1970, *The Crisis of European Sciences and Transcendental Phenomenology: An Introduction to Phenomenological Philosophy*, D. Carr, trans., Evanston, III.

Knorr, W., 1975, *The Evolution of the Euclidean Elements: A Study of the Theory of Incommensurable Magnitudes*, Dordrecht.

Pickering, A., 1989, "Living in the Material World," in D. Gooding *et al.*, eds., *The Uses of Experiment: Studies in the Natural Sciences*, Cambridge, pp. 275–98.

Poisson, S. D., 1837, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilités*, Paris.

Schaffer S. and Shapin, S., 1986, *Leviathan and the Air Pump: Hobbes, Boyle and the Experimental Life*, Princeton.

Spengler, O., 1918, *Der Untergang des Abendlandes: Umriss einer Morphologie der Weltgeschichte*, 2 vols, Munich.

Spengler, O., 1926, *The Decline of the West, Form and Actuality*, C. F. Atkinson, trans., London.

Weinberg, S., 1976, "The Forces of Nature," *Bulletin of the American Academy of Arts and Sciences* **29**, pp. 26–33.

Wittgenstein, L., 1981, *Remarks on the Foundations of Mathematics*, I-142, 3rd edition, Oxford, p. 47.

JED Z. BUCHWALD

# KINDS AND (IN)COMMENSURABILITY[1]

## I. KINDS, INSTRUMENTS AND THE HISTORY OF SCIENCE[2]

For several decades many historians of science have not felt comfortable with philosophers of science, because contemporary philosophy has not often seemed to provide much that would be useful in historical practice. History wants pragmatic value from philosophy. Philosophy has until recently been unable to provide much of it. "Until recently" seems to imply that philosophy of science is or is about to become useful. To the extent that *normative* concerns persist, philosophy remains without much importance to historians.[3] It must instead try to penetrate what characterizes science in a way that captures something historically essential about it, something that can for that reason be put to practical use by historians in their work. I believe that something like this may soon come into being; it is, moreover, something that in a vastly less formal way many historians have long used.

Over the last decade or so Tom Kuhn has developed a new approach, stimulated originally by the concept of kinds, to resolve issues that emerged from his *Structure of Scientific Revolutions*, in particular those that orbited about the concept of incommensurability.[4] Caught initially by John Stuart Mill's theory of real kinds in the *System of Logic*, and which other philosophers had developed into a theory of natural kinds, Kuhn eventually found these to be unsatisfactory and prefers now to write of unqualified kinds. Ian Hacking, on reading in particular Kuhn's Shearman Lectures, has proposed a philosophical explication of Kuhn's still-evolving theory.[5] On reading, first, a draft of Hacking's discussion and then Kuhn's recent work, I realized that one had here the possibility of formalizing and thereby clarifying something that has long formed a part of my own work on the history of electromagnetism and optics, namely the frequent failure of groups of scientists to understand the work of other groups: the problem, that is, of incommensurability. Although there are broad and important differences of opinion between Kuhn and Hacking about kinds, these do not matter for my purposes here, which will rely only on the core idea. In order to avoid these controversial issues I will



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 49-63.

follow Hacking in writing of "scientific kinds", without however intending anything more than that these are the kinds that are deployed by scientists. I will here remain neutral concerning the important and difficult question between Kuhn and Hacking as to whether all kinds are taxonomic.

The essential concept, which I will not elaborate in any detail, is this. Scientific practice (at least) is characterized by the separation of whatever scientists working in a particular area investigate into special groups, or *scientific kinds*. Physicists might examine kinds of light, or kinds of electric conductors or, in general, any set of kinds that are together thought to constitute a related group. The kinds that form such a scientific group differ from categories in general in (at least) one essential respect: namely, that kinds can completely contain other kinds, but partial overlap among kinds is forbidden.

To take a deceptively simple example, consider kinds of electric conductors as they were conceived in, say, the 1850s. Two large classes seemed to exhaust the universe of conductors: metals and electrolytic solutions. They differed in a number of respects, but the essential one concerned their chemical behaviour while carrying a current. Metals did not decompose into their chemical constituents; electrolytes did. Metals as a class did not have known sub-classes, since they were distinguished among one another solely by the value of a single parameter, their conductivity. Electrolytes had many sub-classes, reflecting their particular chemical structures.[6] As a kind, then, 'conductors' contained two sub-kinds, metals and electrolytes, one of which in turn contained other sub-kinds. Containment is a natural essential characteristic of this structure. But partial overlap among the classes must not occur, because otherwise they would lose their meaning as scientifically meaningful kinds. No conductor can be both a metal and an electrolyte; no electrolyte can decompose chemically when carrying a current into more than one set of constituents.

In other words, nothing which is embraced by a given class within a particular group of scientific kinds can be both an *a-thing* and a *b-thing*, where *a* and *b* are group kinds, unless all *a-things* are *b-things* or vice versa. All decomposable, current-carrying liquids are electrolytes, but there is no such thing as an electrolytic liquid that can decompose into one set of constituents when, say, carrying a large current, and into another set of constituents when carrying a small current. If there were such a thing then the entire kind-structure for electrolytes, and possibly for conductors in general (even more radically, perhaps for chemistry itself), would have to be thoroughly re-worked. In more formal terms, if kinds are thought of as sets of objects,[7] then an object can be a member of more than one set in a given group of sets (call this group the object's *family* in respect to a particular area of investigation)[8] only if every set that it is a member of either subsumes, or is subsumed by, another member of the family – scientific sets have no mutual connection at all or else they nest within one another like Russian *Babushka* dolls (better, like a strange kind of *Babushka* in which each doll can contain many other dolls that are immediately visible when it is opened, and so on until dolls that cannot be opened are reached).

Scientific categories can accordingly be thought of as forming a taxonomic tree. The tree-trunk constitutes the major, all-encompassing category that distinguishes the group (e.g. "electric conductors"). Distinct branches emerge from the trunk, defining its immediate sub-kinds (e.g. "metals" and "electrolytes"). Each branch may have its own sub-kinds (e.g. kinds of electrolytes), and these must eventually bottom-out, which is to say that they end up in classes with no sub-kinds (e.g. "water" as a sub-class of electrolytes: nothing further distinguishes water as a conductor from other conductors, and so it has no presently relevant sub-kinds). This way of representing scientific kinds nicely embodies the no-overlap condition. Like the limbs of a tree, every scientific kind therefore emerges directly from a single immediately preceding branch or else from the trunk itself. Since no kind descends from more than one ancestor there is no possibility of partial overlap among kinds (whereas if some kind $A$ had, say, two otherwise-distinct immediate ancestors then these two latter classes would have $A$ in common and so would partially-overlap one another).

Any additions to the tree of kinds must accordingly be grafted onto its structure without violating its integrity: additions can be made, but multiple connections between existing kinds cannot be forged, nor can new kinds be added unless they emerge directly from a preceding kind or from the trunk. If, to take a second example, a new kind of metal is discovered, then it must not also reflect light like, say, glass since optical properties, among others, had already been used to distinguish metals as kinds from glasses. If someone did fabricate such a thing – a sort of optically glassy metal – then 'glass' and 'metal' might be called into question as true kinds. Or perhaps optical characteristics would be called into question as essential for class membership. Even more radically, perhaps the structure for some distance up the tree from both metals and glasses might be reconstituted.

If scientific kinds do properly represent the sorts of things that scientific schemes deploy, at least since *circa* 1800, then the notion of *incommensurability* takes on a thoroughly new, and (we shall presently see) pragmatically significant character. But instead of specifying what *incommensurable* corresponds to we must instead specify what *commensurable* corresponds to. If two schemes are *commensurable* then their taxonomic trees can be fit together in one of the following two ways: (1) every kind in the one can be directly translated into a kind in the other, which means that the whole of one tree is topologically equivalent to some portion of the other, or (2) one tree can be grafted directly onto a limb of the other without otherwise disturbing the latter's existing structure. In the first case one scheme is subsumed by the other. In the second, a new scheme is formed out of the previous two, but one that preserves intact all of the earlier relations among kinds. If neither case holds then you are in the previously-fuzzy realm of the *incommensurable.*

Now, however, many puzzling aspects of *incommensurability* disappears. Take for example the perennial question of whether schemes as wholes, or only certain terms in them, are 'untranslatable' into other schemes and their terms. From the viewpoint of scientific kinds the division that is implicit in the question cannot be sustained. A schema as a whole can only mean (from our new perspective) its complete taxonomic tree. If this scheme can be superimposed onto a segment of another scheme then not only is the former, as a whole, translatable into the latter, but so necessarily are its individual terms, because the tree consists of relationships between the terms. Conversely, if some term in one scheme overlaps more than one term in the other scheme, then it cannot be translated into any term in the latter, and it follows necessarily that some parts at least of the taxonomic trees must themselves have different topologies. It has now become literally meaningless to divorce schemes as wholes from the terms whose mutual relationships they represent.

Why is this practical? It sounds after all rather abstract. It is practical because it provides a good account of what historians taking apart long-dead schemes often do, and because it may even provide guidelines for how to do it. Many historians regularly resort to categories that distinguish the kinds of things that scientists deploy. More than that, they also contrast schemes with one another by pointing out that categories in the one do not correspond to categories in the other because of different relations to other categories. An awful lot of interminable discussion about such questions as whether dephlogisticated air is just another name for oxygen, or not, revolves about precisely this sort of thing. Scientific kinds cut the Gordian knot. They suggest that one thing to do is to attempt the very difficult task of reconstructing the taxonomic tree in its full historical flower, both at a given moment and as it changes over time.[9] There, in that process, you might find new ways to understand many things, such as (for just one example) the birth of sub-disciplines,[10] or of revolutions.[11] And because the construction of a taxonomy consists in the mutual association of whatever elements the practising scientist has available, is forced to use, wishes to use, or thinks it good to use, there is no reason for separating absolutely, say, a belief in energy conservation from a laboratory director's insistence on the investigator's using the laboratory's big, expensive, new equipment. These are two elements that may, contingently, have to be brought together by the investigator in constructing a taxonomy. Which is however not to suggest that all taxonomies are equal among one another (because, it might be asserted, all are based on the contingent association of available elements). They are not, and one aspect of their difference concerns how closely a taxonomy is tied to a particular piece of equipment. The difference I have in mind is this. A novel taxonomy may emerge as someone attempts to grapple with a particular device, but the taxonomy's strength – its ability to assimilate and to fabricate new apparatus – depends to a very large extent on its device-independence, the ease with which it can be separated from the device. In other words the strength of the taxonomy depends on the degree to which the device-taxonomy relation is asymmetrical: the taxonomy inevitably entails a special understanding of the device, but a robust taxonomy is also compatible with many other devices that do what the taxonomy considers to be the same thing that the first one does but in entirely different ways. If you know the taxonomy, you always know what sort of an effect to expect, and for robust taxonomies there must be many conceivable (though not necessarily practical) ways to get that effect. Weak taxonomies tie effects directly to particular devices.

Most important for us here, experiment takes its place as a, perhaps the, central element in the construction of a taxonomy in at least two ways. First, experimental work divides the elements of the tree from one another: sitting at the nodes or branch-points of the tree, experimental devices assign something to this or to that category. Second, experimental work may generate new kinds that can either be assimilated by, or that may disrupt, the existing structure. Moreover, experimental work (as we shall see in Section III through a brief example concerning Anderson's discovery of the positron) may have its own taxonomic structure that to a very large extent exists apart from that of trees with which it is in other respects associated – provided that experimental relations do not violate otherwise-accepted taxonomies, or at least that incommensurable taxonomies are not brought into contact with one another.

All sorts of interesting historical issues take on new aspects in this context. Suppose one group of people fabricates a novel device that produces effects which they can at once assimilate to their tree. Suppose further that these effects cannot easily be assimilated to another group's tree, indeed that the success of the first group's assimilation produces an apparent violation of the integrity of the other's tree.[12] What is to be done, supposing that the second group admits that the first has generated an effect and not an artefact? The answer is clear: the second group must attempt to build a new set of kinds whose relations among one another can accommodate those aspects of the novel effect that they accept, and that can be grafted without disruption onto the existing tree. This will inevitably produce a tremendous amount of verbiage concerning just what had been found, and possibly also new experimental work on the second group's part to manipulate the effect into a more tractable form.

## II. THE GENERALITY OF KINDS

(In)commensurability is on this account bound to the sorting of objects or effects into this or that category, which in turn depends on the critical role of experimental apparatus. Devices act at the nodes of the tree to assign objects to the appropriate categories. Absent the apparatus there would be no sorting, and the apparatus proper often constitutes an embodiment of the relevant kind-structure. One may very reasonably ask, therefore, whether (in)commensurability, and the doctrine of kinds discussed here, are highly limited in historical application, to, say, science after the late seventeenth century, or perhaps even to science post-1800. What, for example, do kinds have to say about the sort of astronomy practised by Kepler, in which the apparatus can scarcely be thought of as embodying kinds in the way that, e.g., Fresnel's rhomb did in wave optics?

This is not an easy question to answer, and I am not certain that the doctrine of kinds can in fact embrace all forms of scientific behaviour. It may just be that it is particularly well-adapted to apparatus-based science, and that it was brought into being along with experimental science. If the doctrine of kinds must be linked to laboratory equipment then their history belongs also to it. I think, however, that a somewhat broader notion of apparatus than I have used to this point may extend the utility of the doctrine beyond these boundaries.

'Apparatus' naturally suggests – and is so defined by the *Oxford English Dictionary* – material devices, machines, entities that make things happen to objects or that react to happenings. A signal characteristic of such devices is one's ability to change them in essential ways, and, in so doing, to make different things happen or to elicit different reactions to the same event. Keplerean astronomy used no such devices, because the telescope cannot work the (celestial) object that is being investigated, nor can it do more than one thing with the object's (optical) effects.[13] Kepler, in working with the observations of Mars bequeathed to him by Tycho, might nevertheless be said to have worked with apparatus of a kind, though not apparatus that did anything to celestial objects or with their light. His "apparatus" consisted of the rules and the mathematical methods that he was prepared to deploy in accommodating Tycho's observations. That apparatus – mathematical devices developed in antiquity – resisted application to some of the effects (the positions of light smudges on the celestial sphere) that Kepler brought it to bear on so long as those effects were also assimilated to Copernican motions. Changing the latter opened a new path, but it also generated a great deal of unresolved tension in the apparatus (antique mathematics). One might be inclined to say that this is just theory-work, rather than laboratory-work, and that writing in this context of "apparatus" is otiose, but it seems to me that these two kinds of labour share at least one basic characteristic which links them to the doctrine of kinds: that of working on something to see what can be made to happen – either through paper "apparatus", or through material devices. Some scientific activity, such as astronomy or astrophysics, works only in the former way; laboratory science usually works in both ways (but see Note 15). Learning standard problems is a kind of training in paper demonstration that is analogous to learning standard demonstration experiments; solving new paper problems bears a similar relation to performing new experiments.

It may accordingly be possible, and useful, to consider a set of rules, procedures and beliefs to constitute a kind of "apparatus". In conformity with usage that is becoming increasingly common, one might want to call this sort of apparatus a theoretical "technology", whereas laboratory devices constitute a material "technology".[14] From the standpoint of kinds, both forms of apparatus can act as sorters. A slice of crystal in a polarimeter does things to light that assign it to a particular category. One may know almost nothing at all about the crystal's likely behaviour beforehand. Worked properly, the polarimeter produces novel information about the crystal. Theoretical devices can do something similar. Succeeding observations of the loci of a strange heavenly object can be subjected to astronomical theory, and it may as a result become possible to assign it to known categories, e.g. to comets. There is an evident difference between the two cases. The polarimeter acts on the object and sorts it. Astronomical theory acts on something other than the object, something that is itself produced by an instrument that engages an effect of the object. Whereas optical theory does not have to intervene in the polarimeter's sorting (once the device has been properly built and worked), astronomical theory itself does the sorting work.[15]

Many historical situations exhibit both types of "technologies". A slice of some transparent stuff may produce coloured rings in a polarimeter, thereby assigning it to the class of ring-producing-things. But the rings may not look like ones previously seen, at which point "theoretical technology" comes to bear, yielding in this case a novel class of objects in respect to their optical behaviour, namely the class of biaxial crystals. This might even occur without the intervention of much "theoretical technology" through the construction of novel material devices that produce new sortings without violating old connections. If these material and paper attempts at sorting fail, then radical new technologies may be produced, or perhaps the effect may be relegated to the sidelines as something inconsequential. The point is that sorting "technologies" do not have to be physical devices,[16] and this may make it possible fruitfully to use the doctrine of kinds for pre-laboratory science.

The critical role of devices in configuring the taxonomic tree for laboratory science means that taxonomies may be distinguished from one another in two very important ways: first, as to their comparative freedom from device-induced category violations, and second, as to their robustness in respect to novel devices. This is, furthermore, not solely an abstract, philosophical point because scientists do just this all the time. They are continually using different kinds of existing apparatus to be certain they have properly understood something, and they generally try to produce new apparatus to get at a process in different ways. A taxonomy that is weak in the first respect and that is not robust in the second will almost certainly not gain adherents over time because it does not work well with or is not fruitful in producing (or both) scientific devices. To the extent that a premium is placed on building a world with apparatus, and on generating new apparatus from that world, such a taxonomy is objectively weak in comparison with one that fits well with existing devices and generates new ones. Nothing in this description requires invoking an absolute, eternal world of entities that apparatus-based science uncovers over time. It does require that, as a matter of fact, devices can be made to work and that new devices can be fabricated as scientific practice grafts, buds and restructures taxonomic trees.

Some taxonomic schemes can work to sort things consistently with existing devices, and can produce new ones; other schemes may have difficulty with the first and find the second nearly impossible. This is objectivity, though certainly not the kind that speaks to knowledge of an abstract, noumenal world.[17] It entails a number of things, among others that taxonomies to which new processes are merely grafted on are not likely to be actively pursued over time because they will not birth devices, which are simultaneously the binding glue and the sorting mechanism of the taxonomic tree. Without devices the tree simply falls apart because its categories are vacuous; but with devices the stable tree can sort things into distinct kinds. Devices have accordingly two apparently conflicting faces, and this is perhaps why they occasion a great deal of trouble for many philosophers of science. Do devices tell you what things are or what they are not? From our perspective here it is clear that they do both simultaneously, because to be a scientific kind of something is to be sorted by a device at some node in a taxonomic tree.

## III. DISCOVERY AND THE AUTONOMY OF TRADITIONS

I will not canvass the objections that might be raised to scientific kinds, at least as we have examined them, but one deserves immediate consideration.[18] If, it might be said, the essential character of scientific practice resides in the tree, then how can discoveries of new kinds be made without, as it were, moving outside science? The taxonomy cannot, by its very nature, produce new kinds, because it is a relation among existing kinds. On the other hand it can certainly produce new devices, even new effects. The answer to our question lies here, in the difference between the operating principles of a device and the things that the device works on.

Let us take an example. Among the apparatus that were critical in Augustin Jean Fresnel's construction of wave optics were doubly-refracting crystals. For the optically well-known kinds of crystals he constructed a novel mathematical surface that linked together light's polarisation and refraction within the crystals, and that could be used to obtain relations that were accepted by the wider optical community (and that accordingly could be understood in several ways). But he went further, and conjectured that an optically less well-known kind of crystal would have a more general surface than the novel one that he had built for ordinary crystals. He himself never built a device that took advantage of this premise, but the Irishman Humphrey Lloyd (at William Rowan Hamilton's suggestion) did, thereby producing an entirely new effect, namely conical refraction. Here we have a situation in which Fresnel as it were created a new optical class of crystals that was subsequently embodied in a nicely-working piece of equipment. This sort of event, which happens very often in the history of science, seems to stand outside the taxonomic structure because the novel crystal kind was not present beforehand; it seems simply to have been grafted on as an otherwise-arbitrary conjecture.

The problem here consists in confusing what the taxonomy's categories are about. They are not about crystals, or metals, or glass. They are about light – polarised light, unpolarised light and their subkinds. And here, in this optical taxonomy, Fresnel *did not graft on anything at all new*. Rather, he used the taxonomy as it stood it to envision a new class of optical stuff in respect to polarisation and refraction – stuff that dealt with light according to his novel surface. New devices might be, and eventually were, constructed using this stuff, but not because new kinds of optical kinds had been conjectured. On the contrary, the optical kinds remained utterly and *essentially* inviolate.

Which is why there can be, and usually is, experimental activity that has nothing much to do with overt theory. Indeed, it seems to me that the manifest existence of this sort of thing – perhaps even its substantial dominance of scientific practice since 1800 at the latest – has driven a great deal of recent historiographical insistence on the independence of experimental tradition. Take, as a rather old example, Anderson's discovery of the positron.[19] That discovery had nothing at all to do with contemporary overt theory, namely with quantum mechanics. It had however a lot to do with Anderson's understanding of the way his cloud chambers and photographs sorted the kinds of things that he had built them to deal with. When Anderson discovered something that curved in a strange way he sifted through the kinds of things it might be and concluded that he had come up with something that simply had to be grafted onto the existing taxonomy: he had discovered a new kind of thing.

Take this a bit further. The taxonomy that Anderson deployed in the laboratory related such kinds as charge, mass, electric and magnetic fields, as well as the elements that constituted his device, the cloud chamber. Much of this was unarticulated, including a great deal of craft knowledge that went into building and operating the cloud chamber; some of it was not. More to the point, the scheme that Anderson was working with contained kinds of particles only secondarily. That is, as far as he was concerned kinds of particles were not essential elements in his laboratory taxonomy, in the sense that his device, electromagnetic fields, charge, mass, etc. were not essentially affected by the kinds of particles that he was detecting. The particle taxonomy was not, for Anderson, am important working scheme. When, consequently, he found a track that he could not assimilate to the behaviour of particles that he knew about, then he just concluded that he had found a new particle with properties that came entirely from his working taxonomy: he had found an object with the mass of the electron and an equal but opposite charge. The essential point to grasp about this is that nowhere did Anderson have to envision a new kind of property for his working taxonomy; he simply used what he already had, pursuing a substantially autonomous tradition of investigation. He didn't have to change anything in his understanding of cloud chambers, fields and so on; he had just found a new kind of stuff, to wit a new particle, which behaved strangely only in the sense that it didn't behave like the particles that he knew about. As far as Anderson's device was concerned they posed no problems whatsoever. Of course he might have found something that he couldn't understand in this way, say a particle that orbits magnetic field lines in ellipses, which would have violated electromagnetic principles. Then he might have called into question his laboratory taxonomy.[20]

*University of Toronto*

## NOTES

[1] This is a portion of a longer article entitled "Kinds and the wave theory of light" to appear in 1992 in *Studies in the History and Philosophy of Science.*

[2] Both Tom Kuhn and Ian Hacking have been more than generous in providing me with their recent thoughts on the subject of kinds. Their work on it motivates and underlies my considerations here, which do not claim to provide philosophical novelty. My own interests

here are more practical than theirs, in that I am concerned with the doctrine of kinds almost entirely for its usefulness in understanding the behaviour of groups of scientists, and in particular their creation and use of instruments.

[3] Although I am going to be discussing something that does look as though it might be used to make normative decisions about past science, I do not think that it can be because it concerns only one aspect of science history – an important, indeed central, one to be sure, but not one that provides a Royal Road, as it were, to good science. See Note 17 for a brief discussion.

[4] See Kuhn's "The Presence of Past Science", "The Shearman Memorial Lecture", University College, London, 1987; "Possible Worlds in History of Science", Proceedings of Nobel Symposium 65, ed. S. Allén, Berlin: de Gruyter, 1989; "An Historian's Theory of Meaning", talk to Cognitive Science Colloquium, UCLA, April 26, 1990; "The Road Since Structure", Presidential Address to the Philosophy of Science Association, 1990.

[5] I. Hacking, "Working in a New World: the Taxonomic Solution", for a volume containing essays in honour of T. S. Kuhn read at the Massachusetts Institute of Technology in May, 1990, forthcoming from MIT, ed. P. Horwich.

[6] Chemical structure was not important for metals because it was precisely their failure to decompose when carrying a current that distinguished them as a class from electrolytes.

[7] For present purposes no distinction will be made between objects such as chairs or even tachyons, and objects like waves of light, even though one commonly says that objects of the former kind may produce objects of the latter kind as effects. Such a relationship would presumably takes its place in an appropriate taxonomy, in which objects of the latter kind sort objects of the former kind.

[8] A particular object might fall into kinds that belong to completely different areas of inquiry. A liquid might for example be considered solely in respect to its viscosity for determining what kind of hydrodynamic entity it is; the same liquid might be considered in respect to its effect on the phase of reflected light in considering what kind of optical object it is. These two groupings of kinds – hydrodynamic and optical – might very well have nothing at all to do with one another, in which case one might have a hydrodynamic object of type $h$ that has reflection properties of type $r_a$ and another hydrodynamic object of the same type $h$ that nevertheless has optical properties of type $r_b$. (In fact, during the late nineteenth century one did have something like this, since aniline dyes are rather viscous, like most solutions, but they are nearly unique in exhibiting marked anomalous dispersion in the visible spectrum, with accompanying oddities in reflection spectra.) What one cannot have is hydrodynamic object that falls simultaneously into *hydrodynamic* kinds that are not nested. Of course it is certainly possible that optical and hydrodynamic behaviour might eventually be brought together (by, e.g., linking molecular structure to both hydrodynamic and to optical properties), but if this does happen then the kinds will have to be reconstructed to prevent overlap.

Hacking pointed out to me an important, and I think related, objection to the non-overlap condition [see Hacking, "Working in a New World"]. Consider arsenic and hemlock as, respectively, kinds of minerals and kinds of vegetables. Both are also kinds of poison – and there is a forensic science of poisons – which accordingly overlaps as a category both minerals and vegetables. Hacking resolves this mundane impasse by distinguishing poisons as a category relative to us from minerals and vegetables, which are not relative to us. This problem, as well as the one concerning hydrodynamic and optical kinds, refers back to one of the central issues underlying the entire doctrine, namely to distinguish properties that are essential to an object's being this or that kind of thing from those that are not. It seems to me that, among practising scientists, instruments are the incarnated repositories of these sorts of questions, which are for the most part rarely brought directly to the surface and addressed forthrightly. I will accordingly rely quite strongly in my detailed example on the function that instruments serve in putting objects into this or that category.

[9] The tree has a deeply historical character because its topology is fabricated by practice, and its distinctions are activated by specific devices that sit at the nodes. The tree is formed by the pragmatic activity of scientists as they try to make things cohere on paper and in the laboratory – which, it is critical to say, certainly is not at all the same thing as asserting that scientists are free to do anything they want to do at any time. I will return to this point through examples below.

[10] Which in taxonomic terms corresponds to the grafting of one scheme onto another, followed by the separate and perhaps autonomous development of the graft – so long as it does not disrupt the rest of the structure).

[11] Which might be taxonomic changes that occur very far up the tree.

[12] I distinguish here between assimilation and grafting for reasons that will be made clear through example below. The distinction corresponds to the difference between being able to understand something new in existing terms (assimilation), which leaves the taxonomy unaffected, and having to add new terms to the taxonomy (grafting).

[13] Increasing the power of a telescope may reveal things not seen before, but it does not do anything qualitatively different – the kind of effect that is being examined (light used to produce an image) remains the same. Using that light in a spectroscope is indeed doing something essentially different, as is using a radio-telescope, because the effects involved are entirely novel (absorption and emission spectra or long-wave emanations). Then the several effects can even be played off against one another in a sort of romp of devices.

[14] See S. Shapin and S. Schaffer, *Leviathan and the Air-Pump*, Princeton: Princeton University Press, 1985 for this unusual way of deploying the word "technology". I thank Andrew Warwick for discussion about "technologies".

[15] One can envision a device that would automatically sort smudges of celestial light into the appropriate objects. Such a thing, it seems to me, would be rather like a polarimeter despite the obvious difference that crystal slices sit in the polarimeter whereas, e.g. comets travel through the heavens. Ian Hacking ("Extragalactic reality: the case of gravitational lensing", *Philosophy of Science* **56** (1989), pp. 555–81) argues that the absence of the object from the laboratory, with one's concomitant inability to manipulate it, constitutes a fundamental distinction (though Hacking's argument aims at grounds for scientific realism, with which I am not concerned).

No doubt the ability to do something to something and see what happens as a result may rapidly produce confidence in what the thing is (i.e. can sort it); being able only to examine what it does as a result of humanly-uncontrollable influences is not so felicitous a situation for the investigator. This is obvious: if you must find an appropriate natural object-stimulator rather than make one yourself then you cannot try to force the type of responses that are interesting when you want them. You must look around for an appropriate natural stimulus. But such stimuli often do exist, and if there are enough of them then you may still feel confidence in saying that the object is such-and-such a thing. Control of the object lies at the heart of laboratory science, but it is not perhaps essential for sorting-activity. For the latter, the issue is rather *what* the object is than *whether* it is an object at all.

[16] There is an obvious caveat that the stuff which feeds into a theoretical sorting mechanism must be produced by something else, often a device, whereas material technologies may both produce and sort. This distinction is important, but it seems to me that the issue revolves rather about historical substance than philosophical absolutes, because successful theoretical technologies often become embedded in physical devices. A related difficulty concerns the kind of device that produces the stuff that a "theoretical technology" may work on. There are intricate devices that may take years of training to work; other devices may require an hour to become skilled with. Some devices may be extremely complicated in construction and yet simple in operation; others may be the reverse. It is for example much easier to push a button on a radio-rangefinder than to make a careful triangulation, even though the former device is much more intricate than the latter. If what you're interested in is the range, then it may not make much difference how you find it. You might even pace it out. But if something funny happens when the range is subjected to your "theoretical technology" then you may start to wonder about the device that gave it to you.

[17] Moreover it does not provide normative criteria that can generally be used in retrospect to assert that *this* rather than *that* scheme should (or should not) have been pursued; it depends upon the context. If the major contemporary *desiderata* revolved about the behaviour of certain kinds of apparatus, and about the production of new kinds, then it may indeed be possible to say that, in this single respect, taxonomy $x$ is weak and taxonomy $y$ is strong. It is possible to do so in the case of early nineteenth-century optics. But there are usually many other factors at work as well, and it would be deeply misleading to ignore them.

Two other sorts of factors are worth mentioning because of their pertinence for our example. First, scheme $x$ may be able to produce all sorts of clever new processes, but it may have trouble dealing at all with some older ones that scheme $y$ could at least account for qualitatively. Indeed, just this will almost always be a major element in the critiques produced by $y$-adherents, such as that the universe should not be filled with stuff. These beliefs can be just as important to scientific work as success with devices because they often underpin the reasoning, covert as well as overt, that produces a novel taxonomy. Divorcing the taxonomy from its belief-structure may very well rob it of something essential to its subsequent vitality.

Belief-structures and arguments over whether this or that process must be taken into account cannot be evaluated normatively, and yet they are *always irremediably* present in the development of science, which makes it otiose to set up comparative evaluations of schemes during periods of intense controversy – that is, during the only times when it is philosophically interesting to do so. Over time this can change, though it is usually difficult to mark a single point at which one can say that $y$-adherents have ceased being rational, primarily because the devices that $x$-adherents claim for their own have by then formed an entire universe. When a remaining $y$-adherent spends all of his time adapting to $x$-devices and generating nothing new then most community members will conclude that the time for dissent has irreversibly passed.

[18] Another question that I have encountered is not so much an objection as puzzlement over what the taxonomic tree is built out of. Where in it, someone asked me, are say Maxwell's equations? The answer is I think reasonably simple: to the extent that Maxwell's equations are considered to specify the essential properties of fields, then to that extent they sit essentially in the devices that sort fields into this or that category. Is there a rapidly-growing

but small magnetic field here? Bring to bear a device that can respond to electromotive forces and you will find out, say Maxwell's equations. No response? Perhaps it is a small but rapidly-growing electric field. Bring in a device that is sensitive to induced magnetomotive forces, say Maxwell's equations. The quantitative structure of a subject generally becomes part of the devices with which it sorts things as the subject stabilizes. New mathematics might destabilize relations between categories, or perhaps even the categories themselves, by calling into question the behaviour of previously-closed devices.

[19] See N. R. Hanson, *The Concept of the Positron*, Cambridge: Cambridge University Press, 1963 and P. Galison, *How Experiments End*, Chicago: The University of Chicago Press, 1987, pp. 90–93.

[20] Note that under these circumstances lots of things are suddenly up for grabs since changes in the laboratory taxonomy have the potential to destabilise previously-secure results. To avoid that possibility requires creating a very high-order kind that, by virtue of its height in the scheme, has no effect on equal-level categories and their subkinds. In general, one might say that a new kind always has the potential for establishing at least the kinds that contain it.

KOSTAS GAVROGLU

# TYPES OF DISCOURSE AND THE READING OF THE HISTORY OF THE PHYSICAL SCIENCES

## INTRODUCTION

It is always more attractive to examine scientific debates in terms of differences in types of discourse rather than examining theories. Independent of anything else, the following is the most significant feature of a type of discourse: It provides a framework where it becomes legitimate to pose certain kinds of questions and to discuss a particular class of phenomena. This legitimizing framework provides the possibility to discuss a whole new set of issues and, *at the same time*, it creates all those problems that make it difficult for the scientific community to have a consensus *both* about the formulation of the new questions as well as about the proposed answers. The success of a new theoretical approach has always been the result of a curious mix of persuasion and proof. Though absolutely essential, proofs were by no means sufficient. Persuasion, however, became indispensable whenever the novelties introduced by a new type of discourse were at stake. "To persuade" meant two things at the same time: consensual activities and legitimizing procedures.

## SOME REMARKS ABOUT STYLES OF SCIENTIFIC DISCOURSE

To attempt a reading of the history of science in terms of types of discourse, it is necessary to understand the differences among such types. In order for the differences to have a practical effect for the reading of the history of science, it becomes necessary to articulate the criteria which will substantiate the particularities of each type. These criteria are:

1 The ontological status of theoretical entities
2. Contextualization of deviations
3. Affinity of a set of propositions

Ian Hacking has argued that (grand) styles of scientific reasoning can be understood without necessarily getting into problems related to relativism. To try and specify the notion of a type of discourse, I shall



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 65-86.

follow Hacking's proposals which assert that the style of reasoning associated with a particular proposition $p$ determines the way in which $p$ points to truth or falsehood. "We cannot criticize that style of reasoning, as a way of getting to $p$ or to not-$p$, because $p$ simply is that proposition whose truth value is determined in this way."[1] One then introduces a range of propositions that are either true or false. Propositions have a positivity in consequence of the styles of reasoning in which they occur. A discourse, in other words, brings into being candidates for truth. The types of discourse are introduced as categories of possibilities. "If positivity is consequent upon a style of reasoning, then a range of possibilities depends upon that style. They would not be possibilities, candidates for truth or falsehood, unless that style were in existence. The existence of that style arises from historical events."[2] Summarizing his views on the styles of scientific reasoning, Hacking's conclusion is

> There are different styles of reasoning. Many of these are discernible in our own history. They emerge at definite points and have distinct trajectories of maturation. Some die out others are still going strong... Propositions of the sort that necessarily require reasoning to be substantiated have a positivity of being true-or-false, only in consequence of the styles of reasoning in which they occur... Many categories of possibility, of what may be true or false, are contingent upon historical events, namely the development of certain styles of reasoning.[3]

Hacking, of course, talks about styles of reasoning that have a global character, such as the geometrical approach, the statistical discourse etc. I feel, however, that it is possible to talk about more "limited" styles in basically the same way. What I would like to do is to examine the extent to which Hacking's proposals may provide a way of reading history of science. But, let me emphasize the point that a type of discourse is characterized by bringing into being a proposition that is true or false. Therefore, it is necessary to comprehend the sense in which $p$ is true *as well as* the sense in which it can become false. A type of discourse is a "package deal:" It is *both* the successes and the failures. Failures will have to be regarded as a necessary ingredient of a type of discourse as successes are.

A discourse is really a network of constraints and the kind of reasoning imposed by these constraints. A discourse can be considered as delineating the conceptual boundaries which determine the types of problems that are posed as well as the types of their solutions. Thinking in such terms emphasizes the possibility to reason towards certain kinds of propositions, but does not – of itself – determine their truth value.

A discourse possesses a peculiarly self-referential character about the criteria it sets and against which it assesses its own coherence. It is a conceptual coherence characteristic of a set of propositions when they become the allowable possibilities of a particular type of discourse. These propositions can, in fact, be accommodated within another type of discourse, and there are obviously ways for understanding their meaning as well as deciding their truth value within this second type of discourse. But, as a whole, they will not seem to have a coherence within this second type of discourse. It is the case that, again as a whole, these propositions do not appear to establish an affinity with the latter discourse. This discourse is "indifferent" towards them, exactly because these propositions, as a whole, do not offer any clues for tracing out the categories of possibilities of the second discourse – even though they were decisive in doing just that in the original.

Therefore, it is possible to define a notion of coherence or affinity for a set of propositions with respect to a particular type of discourse, the assessment of which can be achieved only through self-referential means. This self-referential character, then, can become a way for dealing with the question of the identity of a particular type of discourse. But there is still a problem concerning the practical means for examining this identity and explicating its characteristic signature. One, and by no means the only, such probe for understanding the specificity of each discourse, is to explore questions about model building, analogy and, above all, the ontological status of theoretical entities in each discourse. Bringing into being new candidates for truth or falsehood is neither a matter of using language effectively nor of just being imaginative. Rather, it is a process of realizing the possibilities allowed by the network of constraints.

Let me examine the different types of discourse concerning the problem of chemical bond.

## LONDON AND HEITLER AND HOMOPOLAR BONDING

Undoubtedly the simultaneous presence of Heitler and London in Zurich in the spring of 1927 was one of those unplanned happy coincidences. Heitler and London decided to calculate the "van der Waals" forces between two hydrogen atoms, considering the problem to be "just a small, 'by the way', problem". Nothing suggests that London and Heitler were either given the problem of the hydrogen molecule by Schrodinger, under whom they both had come to work, or that they had detailed talks with him about the problem.

Their initial aim was to calculate the interaction of the charges of two atoms "without thinking even of the exchange" – the term which eventually explained the attraction and whose origin was purely quantum mechanical. They were not particularly encouraged by their result since the "Coulomb integral," despite the small attraction it implied, was too large to account for the van der Waals forces. "So we were really stuck and were stuck for quite a while; we did not know what it meant and did not know what to do with it" Heitler remembered.

Then one day was a very disagreeable day in Zurich; [there was the] Fohn. It's a very hot south wind, and it takes people different ways. Some are very cross... and some people just fall asleep.... I had slept till very late in the morning, found I couldn't do any work at all... went to sleep again in the afternoon. When I woke up at five o'clock I had clearly – I still remember it as if it were yesterday – the picture before me of the two wave functions of two hydrogen molecules joined together with a plus and minus and with the exchange in it. So I was very excited, and I got up and thought it out. As soon as I was clear that the exchange did play a role, I called London up; and he came as quickly as possible. Meanwhile I had already started developing a sort of perturbation theory. We worked together until rather late at night, and then by that time most of the paper was clear.... Well, I am not quite sure if we knew it in the same evening, but at least it was not later than the following day that we knew we had the formation of the hydrogen molecule in our hands. And we also knew that there was a second mode of interaction which meant repulsion between two hydrogen atoms – also new at the time – new to the chemists too. Well the rest was then rather quick work and very easy, except, of course, that we had to struggle *with the proper formulation of the Pauli principle, which was not at that time available, and also the connection with spin... There was a great deal of discussion about the Pauli principle and how it could be interpreted.*[4] (Emphasis added)

## THE APPLICATION OF GROUP THEORY TO PROBLEMS OF CHEMICAL VALENCE

The first indications that their common work can be continued by using group theory are found in a letter to London by Heitler in late 1927. By September 1927 Heitler was back in Germany, having become an assistant to Born in the place of Hund. He was very excited about the physics at Göttingen and especially about Born's course in quantum mechanics where everything was treated with the matrix formulation and then one derived "God knows how, Schrodinger's equation".[5] Heitler felt that the only way the many-body problem could be dealt with was with group theory, and he outlined his program to London in two long letters.

His first aim was to clarify the meaning of the line chemists draw between two atoms. His basic assumption is that every bond line means exchange of two electrons of opposite spin between two atoms.

The general proof for something like this cannot be given, except group theoretically... [If we achieve this] we can, then, eat Chemistry with a spoon.[6]

This overarching program to explain all of chemistry got Heitler into trouble more than once. "I often teased Heitler because he sort of felt that he had explained the whole of chemistry, and I was sceptical of that. I asked him, well now, 'what chemical compounds would you predict between nitrogen and hydrogen?' And of course, since he did not know any chemistry he couldn't tell me."[7] Heitler confesses as much in his interview: "The general program was to continue on the lines of the joint paper with London, and the problem was to understand chemistry. This is perhaps a bit too much to ask, but it was to understand what the chemists mean when they say an atom has a valence of two or three or four... Both London and I believed that all this must be now within the reach of quantum mechanics."

London was in agreement with Heitler that group theory may provide many clues for the generalization of the results derived by perturbation methods. The aim is quite obvious: to prove that quantum mechanics stipulates that from all the possibilities resulting from the various combinations of spins between atoms, only one term provides the necessary attraction for molecule formation. Nevertheless, London was not carried away by the spell of the new techniques – as Heitler was in the company of Wigner and Weyl in Göttingen.[8]

By the middle of 1928, London drew a program to tackle "the most urgent and attractive problem of atomic theory: the mysterious order of clear lawfulness, which is the basis for the immense factual knowledge of chemistry and has been expressed symbolically in the language of chemical formulas". It was a three-pronged program. First, he intended to deal with the problem of the mutual force interaction between the atoms. Second, he wished to examine whether it was possible to decipher the meaning of the rules that the chemists had found in semi-empirical ways and to place those on a "sound" theoretical basis. Third, he attempted to determine the limits of these rules and, if possible, to initiate a quantitative treatment of them.

The recurring issue of the ontological status of theoretical entities should be emphasized in order to fully assess London's novel theoretical approach. To conceptualize a notion of valence from the quantum mechanical and group theoretical considerations and, then, to persuade that this is formally equivalent to the actual chemical valence. The fact that the attraction of the two hydrogen atoms is, in the last analysis, due to symmetry considerations points towards the use of group theoretical methods. One examines all the possible combinations between two or more atomic wave functions, and where the Pauli principle becomes the syntactic rule that points to all the situations that have physical/chemical meaning.

## TWO CONFERENCES

Questions related to chemical bonding and valence were exhaustively discussed in two important meetings. The first was a "Symposium on atomic structure and valence" organized by the American Chemical Society in 1928[9] in St. Louis. The second was organized by the Faraday Society in 1929 in Bristol and its theme was "Molecular spectra and molecular structure."

The level of sophistication in the chemists' talks in the American Chemical Society meeting was impressive. The speakers appeared to be fluent in the ways of the new physics.

G. L. Clark's opening was quite remarkable in that respect

> It may be asserted, in spite of discrepancies and disagreements, that in the first quarter of the twentieth century, the actual existence of an atomic world became an established fact. Part of the difficulties in details may be ascribed to failure of chemists to test their well-founded conceptions with the facts of physical experimentation, and far too few physicists inquired critically into the facts of chemical combination. So, firmly enthrenched each in his own domain, a certain long-range firing of static cubical atoms against infinitesimal solar atoms has ensued, with few casualties and few peace conferences. The position of the Bohr conception has seemed so convincing that perhaps the majority of thinking chemists were coming to accept the dynamic atom, which is fully capable of visualization.[10]

Clark was not alone in attempting to specify the newly acquired consciousness about this strange relationship between the physicists and the chemists. Worth Rodebush went a step further than Clark. He claimed that the divergent paths of physicists and chemists had started being drawn together after the advent of quantum theory and specially after Bohr's original papers. But in this process

> the physicist seems to have yielded more ground than the chemist. The physicist appears to have learned more from the chemist than the chemist from the physicist. The physicist now tells the chemist that his way of looking at things are really quite right because the new

theories of the atom justify that interpretation, but, of course, the chemist has known all the time that his theories had at least the justification of correspondence with a great number and variety of experimental facts.[11]

He gracefully remarks that it is to the credit of the physicist that he can now calculate the energy of formation of the hydrogen molecule by using the Schrodinger equation.[12] But the difficulty in a theory of valence is not to account for the forces which bind the atoms into molecules, suggesting that one can do this by using electrochemical theory. The outstanding tasks for such a theory is to predict the existence and absence of various compounds, and the unitary nature of valence which can be expressed by a series of small whole numbers leading to the law of multiple proportions.

Perhaps the most cogent manifestation of the characteristic approach of the American chemists is Harry Fry's contribution in this Symposium.[13] He attempts to articulate what he calls the "pragmatic outlook". He starts by posing a single question that should be dealt with by the (organic) chemists. What kind of modifications to the structural formulas would conform with the current concepts of electronic valency? This should by no means lead to a confusion of the fundamental purpose of a structural formula: to present the number, kind and arrangement of atoms in a molecule, as well as correlate the manifold chemical reactions displayed by the molecule.

It should here be noted that no theory in any science has been so marvelously fruitful as the structure theory of organic chemistry.... When we are considering methods of modifying this structure theory of organic chemistry, by imposing upon its structural formulas an electronic valence symbolism, are we not, as practical chemists, obligated to see to it that such system be one that is calculated to elucidate our formulas rather than render them obscure through the application of metaphysically involved implications on atomic structure which are extraneous to the real chemical significance of the structural formulas, *per se*... The opinion is now growing that the structural formula of the organic chemist is not the canvas on which the cubist artist should impose his drawings which he alone can interpret... Many chemists believe that the employment of a simple plus and minus polar valence notation if all that is necessary, at the present stage of our knowledge, to effect the further elucidation of structural formulas. On the grounds that practical results are the sole test of truth, such simple system of electronic valence notation may be termed 'pragmatic'.[14] (Emphasis mine)

'Chemical pragmatism' resists the attempts to embody, in the structural formulas, what Fry considers to be metaphysical hypotheses: Questions relative to the constitution of the atom and the disposition of its valence electrons, It is the actual chemical behavior of molecules that is the primary concern of the pragmatic chemist, rather than the imposition on these formulas of an electronic system of notation which is further complicated by the metaphysical speculations involving the unsolved problems of the constitution of the atom. Fry has to admit the obvious fact that the more chemists know about the constitution of the atom, the more fully they will be able to explain chemical properties. But he invokes Kant's assistance, warning that premises lying outside the territory of sensation experience are bound to lead to contradictory conclusions. His position is that the formula which should be assigned to a molecule should be in conformity with its particular behavior at that particular time.

A particularly interesting aspect of the Faraday Society meeting was the systematic articulation of the molecular orbital approach as a way of providing a quantitative dimension to the possibilities made explicit by the group theoretical considerations for valence by London (and to a lesser extent by Heitler). This does not mean that these considerations were the sole reason for the formulation of the molecular orbital approach. Nor that all its adherents had the same starting point. The widely held view regarding this approach at the antipodes of the valence bond method, though methodologically justified, is historically untenable. London's main result in his group theoretical considerations was to enumerate the possibilities which could be realized based on the fundamental principles of quantum mechanics. Hund's contribution[15] was an attempt to alleviate a weakness of the group theoretical approach where chemical *binding* could not be understood in terms of energetics and that, mainly, only the saturation of the valences was explainable in terms of spin. He suggested a series of criteria for molecule formation to account for the characteristic difference between $H_2$ and HeH. One may be the fact that in the hydrogen molecule, as in He, the two electrons are in equivalent levels, whereas this is not possible in HeH since two of its electrons are in such a level already. A second criterion may be the splitting of the H+H term and the absence of such a splitting in He+H. Lennard-Jones, in the same meeting[16], proposed a set of rules for the assignment of electrons to molecules and which are consistent with the implications of the group theoretical considerations of both London and Heitler.

## PAULING'S RESONANCE STRUCTURES

Linus Pauling spent three months in Zurich where he met Heitler and London before going to Copenhagen in late 1927. Right after the appearance of the Heitler–London paper, Pauling published a short note to bring attention to an unforgivable omission: Lewis was nowhere mentioned in the paper. Pauling emphasized that London's extension of the Heitler–London approach "is in simple cases entirely equivalent to G. N. Lewis's successful theory of the shared electron pair, advanced in 1916 on the basis of purely chemical evidence",[17] acknowledging at the same time that the quantum mechanical explanation of valence is more powerful than the old picture. Pauling suggested the direction along which he moved to derive some new results and he explicitly stated his methodological commitments:

It is to be especially emphasized that problems relating to choice among various alternative structures are usually not solved directly by the application of the rules resulting from the quantum mechanics; nevertheless, the interpretation of valence in terms of quantities derived from the consideration of simpler phenomena and susceptible to accurate mathematical investigation by known methods now makes it possible to attack them with fair assurance of success in many cases.[18]

In this paper he mentioned for the first time that the changes in quantization may play a dominant role in the production of stable bonds in the chemical compounds. That was the first hint as to the hybridization of orbitals. He perceived that perturbations to the quantized electronic levels may produce directed atomic orbitals whose overlapping would be better suited for the study of chemical bonds.[19]

At the same time as the publication of the paper by Heitler and London, Pauling was hard at work to find an alternative approach. After stressing that London's treatment for the simple cases is fully equivalent to G. N. Lewis's ideas of the non-polar bond, Pauling suggested that refinements and extensions of London's simple theory were also possible, involving the quantitative consideration of spectral and thermochemical data. These would lead to a number of important conclusions regarding the hydrogen bond, the nature and occurrence of the double and triple bond, the stability of various valences, the structure of graphite, of the benzene ring and so on.[20]

The first paper of a series where he developed his thoughts about the nature of the chemical bond was published in 1931.[21] He tried to "work out hybridization of bond orbitals in a simple enough way so that I could get somewhere in a finite length of time in making calculations".[22] Concerning the electron-pair bond, Pauling proposed a set of rules, not all of which were rigorously derived, but mostly inferred from the rigorous treatments of the hydrogen molecule, and the helium and lithium atoms.

Linus Pauling exploited maximally the quantum mechanical phenomenon of resonance and was eventually in a position to formulate a rather comprehensive theory of chemical bonding. The success of the theory of resonance in structural chemistry consists of finding the actual structures of various molecules as a result of resonance among other "more basic" structures. In the same manner that the Heitler–London approach provided a quantum mechanical explanation of the Lewis electron pair mechanism, the quantum mechanical theory of resonance provided a more sound theoretical basis for the ideas of tautomerism, mesomerism and the theory of intermediate state. Within these theories it is considered possible for the actual state of a molecule to be unidentical with that represented by any single classical valence-bond structure; it may be intermediate between those represented by two or more valence bond structures. The quantum mechanical resonance approach leads to an understanding of *the conditions under which a molecule can be expected to exist in an intermediate stage or mesomeric state as well as an accounting for the greater stability of those molecules that are the result of resonance.*

Wheland, who was one of the closest associates of Pauling, has expressed all this in the simplest manner:

> Resonance is a man made concept in a more fundamental sense than most other physical theories. It does not correspond to any intrinsic property of the molecule itself, but instead it is only a mathematical device, deliberately invented by the physicist or chemist for his own convenience.[23]

What Pauling greatly emphasized was not the arbitrariness of the concept of resonance, but its immense usefulness and convenience which "make the disadvantage of the element of arbitrariness of little significance".[24] This, as he repeatedly said, was his constitutive criterion for theory building in chemistry – the way he particularized Bridgman's operationalism in chemistry.

## MULLIKEN AND HIS MOLECULAR ORBITALS

Though the method of molecular orbitals was first introduced by Hund, it was Mulliken who provided both the most thorough treatment of the different kinds of molecules as well as the theoretical and methodological justifications for "legitimizing" the molecular-orbital approach. When a molecule is formed from two atoms approaching each other there should be an increase of the *n* value of some of the electrons so that the Pauling principle would be satisfied for the molecule as well as the corresponding "united-atom". Any electron whose *n* value is increased in such a manner is called a promoted electron. Rather than accepting the division of electrons into two classes (the "bonding" electrons that come in pairs and the "non-bonding" electrons that do not), Mulliken introduces the notion of a varying "bonding power" for various orbit types. An electron possesses energy-bonding-power which can be understood by the effect that the removal of a particular electron has on the dissociation energy of a molecule. An electron also possesses distance-bonding-power which can be understood by the effect that removal of a particular electron has on the equilibrium internuclear distance between the atoms that make up a molecule. The energy-bonding-power of an electron, precisely because it is more intimately connected with the concept of a promoted electron, is more convenient to use in formulating the molecular orbital theory.

In a paper in 1928,[25] Mulliken attempted to determine the extent to which the electronic states of the atoms that are produced by dissociation form the molecules whose electrons have been assigned specific quantum numbers. His results are quite impressive. He believed that the vast amount of spectroscopic data appeared to be quite sufficient to become the basis of his heuristic arguments for a phenomenological treatment without having to "resort" to the details of the "new" quantum mechanics. If anything, this would be the most characteristic aspect of Mulliken's overall methodology. He was, of course, aware of the issues raised by such a strict dependence on experimental data and the analogy with atomic physics.

The assignments are based mainly on band spectrum, and to a lesser extent on ionization potential and positive ray data. The methods used involve the application of Hund's theoretical work on the electronic states of molecules. Although the actual state of the electrons in a molecule, as contrasted with an atom, cannot ordinarily be expected to be described accurately by quantum numbers corresponding to simple mechanical quantities, such quantum numbers can nevertheless be assigned formally, with the understanding that their mechanical interpretation in the real molecule ... may differ markedly from that corresponding to a literal interpretation. With this understanding, a suitable choice of quantum numbers for a diatomic molecule appears to be one corresponding to an atom in a strong electric field" (emphasis added).[26]

After his work on band spectra and the assignment of quantum numbers to electrons in molecules, Mulliken proceeded to formulate a theory of valence in a series of papers in 1932.[27] The theory is, in a way, the "natural" outcome of a program whose aim was to describe and understand molecules in terms of (one-electron) orbital wave functions of distinctly molecular character. The attempt was to articulate the relative (not relatively) autonomous character of molecules through a process that depended on analogies with atoms and the extensive data concerning band spectra. In fact, his theory became an alternative mode to the treatment of the problem of valence by Heitler, London, Pauling and Slater. Mulliken would claim that by his theory it could be shown that the notion of electron sharing is not a necessary component of the more successful theory of valence and that the chemical data that had led Lewis to propose such a concept could be explained quantum mechanically. However, he started building his theory by exploiting the possibilities provided by the very concepts that he would eventually discard.

Unshared electrons are described in terms of atomic orbitals and the notion of molecular orbitals is introduced to describe shared electrons. Electrons are divided into three categories according to their role in the binding process: shared electrons (at least for diatomic molecules) are either bonding or anti-bonding electrons, unshared electrons are the non-bonding electrons. The latter occur in diatomic molecules only when accompanied by a larger number of bonding electrons. Mulliken then examined the situation with three different, yet partially overlapping modes of describing a molecule. In the one-nucleus viewpoint the electrons immediately surrounding each nucleus of the molecule are considered in terms of atomic orbitals from the "viewpoint of that nucleus".[28] Such an approach is best for unshared electrons whose orbits, compared to the orbits in a free atom, would be deformed in a molecule. Though it is possible to describe the orbits of the shared electrons as well, a more convenient mode for this is either the united-atom viewpoint or the view making use of shared orbitals belonging to radicals.

> London and Heitler, generalizing results obtained from a quantum theoretical study of the formation of $H_2$ from H+H, attempted to construct a valence theory which has often been supposed to be the quantum-mechanical equivalent of Lewis's ideas... This so-called spin theory of valence emphasizes the pairing of electrons and their spins, but deals primarily with the interaction of the atoms as wholes. It has, however, not proved very successful... One should distinguish between Heitler and London's valence theory and their valuable perturbation-method for calculating energies of molecule formation.[29]

Mulliken, time and again, emphasizes that the concept of the bonding molecular orbitals is more general, more flexible and certainly more "natural" than the Heitler–London electron-pair bonding – even though the latter may turn out to be more convenient for quantitative results for a number of problems.

In general it may be said that there are many phenomena which can be interpreted in terms of electron-pair bonds only if after setting up these bonds, various linear combinations are formed, while the molecular orbital concept goes more directly at the solution, although often seemingly neglecting certain features expressed by the electron-pair concept. It appears probable that in practice on can expect most of the phenomena expressible by the special concept of electron-pair bonds to drop out of the application of the less specialized method using molecular orbitals.[30]

These considerations lead to the formulation of the valence rule: Every nucleus in a molecule tends to be surrounded, by means of sharing or transfer of electrons, by an electron distribution corresponding to some stable configuration, having a total charge approximately equal to or somewhat exceeding the charge of the nucleus. This rule is formulated for molecules that are not "united-atom" molecules, with respect to each particular nucleus. The Pauling principle is used for, at least a partial, justification of this rule.

The assignment of the various quantum numbers to the molecular orbitals leads to an alternative explanation of homopolar valence that does not depend on resonance, but rather on the redefinition of the notion of the promotion – termed premotion – to be used for the one-nucleus viewpoint of the nuclei in the molecule. Then, bonding-electrons are, in effect, unpremoted electrons whereas anti-bonding electrons are strongly premoted electrons. Therefore, "chemical combination of the homopolar type is a result of the shrinkage and consequent energy-decrease of atomic orbitals in the fields of the neighboring nuclei, when such orbitals are shared with little or no premotion".[31] It is shown that the role exchange integrals of Heitler and London correspond to the electron density of the molecular orbitals: bonding orbitals have a higher electron density, anti-bonding orbitals have a lower density in the regions between the nuclei than the densities that would have resulted by the overlapping of the electron densities of the orbitals of isolated atoms.

## THE CORRESPONDENCE BETWEEN HEITLER AND LONDON

The correspondence between Heitler and London is quite revealing on various levels: it shows the attitude of each about the possible development of the approach laid down in their common paper, the tension between them, and the search for a means to consolidate their theory at a time when the Americans appeared to be taking over the field of quantum chemistry. The correspondence reflects the different styles of their respective environments. Faithful to the Göttingen spirit, Heitler is "more mathematical," London continues in the Berlin tradition of theoretical physics with its aversion to the semi-empirical approaches, but inclined to examine intuitive proposals.

There was no correspondence between the two until 1935. By then, London was in Oxford and Heitler in Bristol. Both were working on different problems and far removed from the problem of chemical bonding. Heitler was working on the theory of radiation, London had started work in low temperature physics. The publication, nevertheless, of the papers by the Americans and especially Slater, Pauling, Van Vleck and Mulliken prompted a rather desperate exchange of letters between the two of them at the end of 1935. They discussed the possibility of writing an article in *Nature* where they planned to present their old results and include some new aspects which had not been emphasized properly. These were the activation of spin valence and the possibility of a bond that would not depend on spin saturation. "That is what I meant in past note – vaguely and wrong – with the term orbital valence."[32]

They felt that among the missed opportunities was their lack of insistence about the oxygen molecule: "It is only due to our negligence that now comes Van Vleck (after the publication of the matter!) and writes that $O_2$ is a triumph of Mulliken–Hund, because our theory is 'less elementary"'.[33] London insisted that the "essence of a discovery is to know what one is doing."[34]

Heitler's attitude concerning the approach by Slater and Pauling was that they were correct about the principles they adopted and he was quite sympathetic about the direction of their research, even though their work was not unnecessary in order to derive a series of results. He thought a polemic against them was quite unjustified. "I simply find that the importance of this theory has been monstrously overrated in America".[35]

Doubts were expressed for the first time about the character of the attractive forces. Perhaps they were not only due to spin. Attractive forces of the same order of magnitude as the usual ones did not follow from their original theory of spin valence.

> The next question is whether one should call these forces, that are added to our original ones, as *valence* forces. Well, the chemists undoubtedly do it, since they name, or, rather, they named in this way whatever gives molecules (in contrast to the v.d. Waals forces and the pure ionic molecules). This is exactly our job. To say *that* there are also other forces of molecule formation except our old ones and *which* phenomena of chemical valence depend on those,

and that especially that our old scheme can be extended.[36]

Heitler's feeling was that there had been no attack by the Americans except for the case of the oxygen molecule whose diamagnetism they could not explain.

The nucleus of our theory is the spin valence and that our theory is the only one that explains the mechanism of repulsion in a qualitatively exact manner. It is needless to write this since we surely agree on that. You could perhaps include the above discussion under the title: *Delineation of completeness* (so much of theory as well as of the chemical notion of valence that corresponds to theory). In any case, we should stress that the extension could be realized on the basis of our theory and, substantially, – it is ridiculous – it includes whatever one could wish (this last thing only as a footnote for us). It is ridiculous even from a quantitative point of view.[37]

London answers by expressing his not–too–kind feelings about the chemists:

The word "valence" means for the chemist *something more than simply forces of molecular formation*. For him it means a substitute(?) for these forces whose aim is to free him from the necessity to proceed, in complicated cases, by calculations deep into the model. It is clear that this remains wishful thinking. Also the fact that it has certain heuristic successes. We can, also, show the quantum mechanical framework of this success ... the chemist is made out of hard wood and he needs to have rules even if they are incomprehensible.[38]

Their realization that, in no uncertain terms, they were becoming increasingly isolated became apparent to them. The fact that they had not been attacked was no indication of the acceptance of their theory. They felt their theory may have been forgotten or that it "can be combatted much more effectively by the conscious failure to appreciate and avoid mentioning it".[39]

London suggested that they clarify the situation with the returned oxygen bond and they planned to meet in London. Both decided to study chemistry.

I was looking for ways to devour the so-called theory of Sl[ater]–P[auling]. These types are so proud about something which is not so bad, but which, under no circumstances, is so distinguished. It gives a general formula for the bond, that corresponds to the pair bonding and the repulsion of the valence lines.[40]

There are not many places where we can read the opinions of either Heitler or London concerning the molecular orbital approach. Heitler thought that their basic objection to "Hund's people" – who they both agreed were not their biggest and most unpleasant enemies – was not related so much to the actual results derived by this method. Sufficient patience with the calculations and a lot of semi-empirical considerations give, in fact, correct results. "Nevertheless, no one can name this a general theory – much less a valence theory – since all the *general and substantive* points are forever lost".[41]

At long last they realized that "in the last analysis the pressure to do what is necessary falls on us. What is needed to keep the more dangerous of our colleagues, those, in other words, who work with our method falsifying history (Eyring, Pauling, etc.) in their place, is a good standard book. Would you not want to write it?".[42] Oxford University Press suggested that Heitler write another book, especially after the success of the book on radiation, and he toyed with the idea of writing one with London about quantum mechanics and chemistry. There was no more talk about these issues in the few remaining letters. They were both working on totally different topics – Heitler on cosmic rays and heavy electrons, London on superconductivity – and the question of finding a permanent job was, once more, seriously preoccupying them.

## THE NECESSARY SYMBIOSIS OF DIFFERENT STYLES

The time during and after the publication of the Heitler–London paper coincides with the formation, in the U.S.A., of a community of physicists and chemists with particularly significant contributions, who were being taken seriously by their European colleagues. This group consisted mainly of Linus Pauling, John Slater, Richard Mulliken and John van Vleck. The first three worked almost exclusively on problems related to chemical bonding and it was something that interested Van Vleck as well. Furthermore, the treatment of the problems by Pauling and Slater was never fully approved by London, whereas Mulliken's molecular orbitals was a distinctly different approach to the problem of the formation of molecules.

It is often the case that when referring to the different approaches to the question of atomic bonding, to consider that there are two methods: The Heitler–London–Pauling–Slater method and the Hund–Mulliken method. I would like to argue for the following: drawing up a program for investigating the nature of the chemical bond presupposes a particular attitude on how to construct a theory in chemistry, on how much one "borrows" from physics and what the methodological status of empirical observations is. Concerning the work of the people we have mentioned above there were basically two different "styles". Heitler and London insisted on an approach which – while it was not as reductionist as Dirac's pronouncement of 1929 – followed this path of orthodoxy. Pauling and Mulliken had a strong attraction to semi-empirical methods whose only criterion for acceptability was their practical success. To suppose that the question of a stronger command over the mathematical details is the sole differentiating criterion between the two styles is, I think, quite misleading. The difference could only be understood in terms of two different cultures for "doing" (physical) chemistry. In the last analysis, it is a matter of explicating the internal, theoretical and methodological coherence of the proposed schemata, and realizing that they constitute two diverging programs.

Pauling felt more at ease with the Schrodinger approach rather than with matrix mechanics and did not worry about questions of interpretation of quantum mechanics. "I tend not to be interested in the more abstruse aspects of quantum mechanics. I take a sort of Bridgmanian attitude toward them. Bridgman with his ideas about operational significance of everything would say that a question does not have operational significance ... is meaningless".[43]

Almost everything in the series of Pauling's papers, starting in 1931 and titled *The Nature of the Chemical Bond*, are included in his book of the same title. There are, however, some details of significance. In the opening paragraph of the first paper in the series Pauling states his assessment of the situation concerning work on the chemical bond as well as the method he will follow.

> During the last four years the problem of the nature of the chemical bond has been attacked by theoretical physicists, especially Heitler and London, by the application of quantum mechanics. This work has led to an approximate theoretical calculation of the energy of formation and of other properties of simple molecules ... and has also provided a formal justification of the rules set up in 1916 by G. N. Lewis for his electron bond. In [this] paper it will be shown that many more results of chemical significance can be obtained from the quantum mechanical equations, permitting the formulation of an extensive and powerful set of rules for the electron-pair bond supplementing those of Lewis.[44]

Texts of this sort are, in a way, pacesetting texts; they are texts influencing chemists and contributing to the formation of the "chemists' culture". It is the theoretical physicists who applied quantum mechanics to a chemical problem, but at the same time, he considers his own work as an extension of their program. His applications provide "many more" results which can be obtained in the form of rules supplementing other rules. And, in fact, the rules which are formulated later in the paper are provided with a kind of quantum mechanical justification and they are by no means rules derived from first principles. Pauling's papers are mathematically sophisticated and from the calculations he had published it is evident that he was at home with the details of quantum theory. Nevertheless, it is impressive that he was able to present a coherent and convincing argument about the unfolding of the nature of the chemical bond with such little mathematics in his book. In this manner Pauling was able to inaugurate the language of quantum chemistry which could be used by chemists in a practical manner.

In his analysis of resonance, Pauling expresses in the most explicit manner his views about theory building in chemistry. He asserts that the theory of resonance is a chemical theory, and, in this respect, it has very little in common with the valence-bond method of making approximate quantum mechanical calculations of molecular wave functions and properties. Such a theory – which is an empirical theory – is "obtained largely by induction from the results of chemical experiments".[45] The development of the theory of molecular structure and the nature of the chemical bond, Pauling asserts in his Nobel speech in 1954, "is in considerable part empirical – based upon the facts of chemistry – but with the interpretation of these facts greatly influenced by quantum mechanical principles and concepts".[46]

Pauling himself is very clear about the features of his own style.

> It soon became the custom to say that there are two alternative methods of discussing the structure of a molecule. One method based on molecular orbitals was called the Hund–Mulliken method ... the other method came to be called the Heitler–London–Slater–Mulliken method.
>
> It is my opinion that it might be worthwhile to distinguish between the Heitler–London method and the Slater–Pauling method. They are identical for the simplest molecules, such as the hydrogen molecule, but different for more complicated molecules. The difference depends on the acceptance or rejection of the assumption that the atoms in molecules retain essentially the electronic structure of their individual normal states. For a molecule or other aggregate of noble-gas atoms or of ions with the noble-gas structure, this assumption is justified, but for other molecules and crystals it is a poor assumption.
>
> My later work on the nature of the chemical bond has been based almost entirely on the Slater–Pauling method. In fact, since 1930 I have made very few precise quantum mechanical calculations. My work on the nature of the chemical bond and its application to the structure of molecules and crystals has been largely empirical, but for the most part guided by quantum mechanical principles. I might even contend that there are four ways of discussing the nature of the chemical bond: the Hund–Mulliken way, the Heitler–London way, the Slater–Pauling way, and the Pauling semi-empirical way.[47]

Apart from its quantitative aspects, the molecular-orbitals approach is essentially an alternative theoretical framework to the one formulated after the pioneering work of Heitler and London. In a remarkable paper titled "On the method of molecular orbitals" published in 1935, Mulliken expresses his views on what he considers to be the most characteristic and differentiating aspects of his theory. The Heitler–London method "follows the ideology of chemistry and treats every molecule, so far as possible, as composed of definite atoms... It has had the notable success as a qualitative conceptual scheme for interpreting and explaining empirical rules of valence and in semiquantitative, mostly semi-empirical calculations of energies of formation".[48] The method of molecular orbitals departs from "chemical ideology ... and treats each molecule, so far as possible, as a unit". This seemingly terminological – or shall we say procedural? – difference highlights the more theoretical issues involved in the study of molecular physics.

> It is the writer's belief that, of the various possible methods, the present one may be the best adapted to the construction of an exploratory conceptual scheme within whose framework may be fitted both chemical data and data on electron levels from electron spectra. A procedure adapted to a broad survey and interpretation of observed relations is here aimed at, rather than (at first) one for quantitative calculation, which logically would follow later. Given an observed molecule or ion of known shape and size, what is its electronic structure in terms of an electron configuration using, in general, non-localized orbitals for shared electrons? What is the relation of this structure to the molecule's spectrum, to its ions, and to the structures of other similar and dissimilar molecules?... To a considerable extent the present method and objective are very analogous to those used by Bohr in developing the theory of atomic structure.[49]

Mulliken recognizes the legitimacy of the criticism that one of the reasons for the poor quantitative agreement using the molecular orbital approach, is because of the inability of this theory to include the details of the interactions between the electrons. But even though their quantitative inclusion would make a theoretical calculation from first principles a quite impossible job, "their qualitative inclusion has always formed a vital part of the method of molecular orbitals used as a conceptual scheme for the interpretation of empirical data on electronic states of molecules".[50] Such considerations, in fact, led to the qualitative explanation of the paramagnetism of oxygen – one of the main weaknesses of the valence bond approach.

In two of his interviews, Thomas Kuhn tries to discuss the issue of difference in styles concerning the different approaches to chemical bonding. He asks Mulliken whether there were a series of schools – a Heitler–London school, a molecular orbital school – with some geographical localization and whether in certain places one approach was being used and in other places another approach was being used "so that people were somewhat past each other?".[51] Mulliken is noncommital in his answer. "I do not know. The way I was thinking was not in such terms as to notice things quite in that framework. I would say there were some people who were stronger for one thing than for another, but whether they were more abundant in one particular place I do not know".[52] Wigner, on the other hand, states that he never felt this opposition, since it was very clear to him from the very beginning that there were different objectives of these approaches. For example, the molecular orbital method does not speak about the bond, but rather it has molecular orbitals which extend over the whole molecule. "This is too far away from the very useful and very fruitful chemical concepts".[53]

A very characteristic reaction – apart from that of Heitler and London in their correspondence – was by Hund who was the cofounder of the molecular orbital approach. He had completed a paper where he discussed various points concerning the molecular orbitals. Just before sending his paper for publication, he was sent a preprint by Mulliken who had essentially done the same calculations. But Hund decided to go ahead and publish his paper since "Mulliken's paper is rather American, e.g., he proceeds by groping in an uncertain manner, where one can say theoretically the cases for which a particular claim is valid".[54]

Could it be the case that it is possible to discern different national styles? I do not think one can give a definitively negative answer, but neither can one exclude a positive answer.

*National Technical University, Athens, Greece*

## NOTES

[1] I. Hacking "Styles of Scientific Reasoning," in *Post-Analytic Philosophy*, eds. J. Rajchman and Cornel West, Cornell University Press, 1985; p. 146.
[2] *Ibid.*, p. 155.
[3] *Ibid.*, p. 162.
[4] Interview with Walter Heitler at the Archive for the History of Quantum Physics (A.H.Q.P.), American Institute of Physics.
[5] Heitler to London, September (?) 1927.

[6] Heitler to London, September (?) 1927.
[7] Interview with E. Wigner by T. S. Kuhn, December 4, 1963, p. 14.
[8] Heitler to London, December 7, 1927.
[9] Division of Physical and Inorganic Chemistry.
[10] G. L. Clark, "Introductory remarks in the Symposium on Atomic Structure and Valence," *Chemical Reviews* 5, 1928, p. 362.
[11] W. H. Rodebush "The electron theory of valence," *Chemical Reviews* 5, 1928, p. 511.
[12] Rodebush who was a student of Lewis does not mention Heitler and London by name when he comes to the quantum mechanical treatment of the hydrogen molecule.
[13] H. S. Fry, "A pragmatic system of notation for electronic valence conceptions in chemical formulas," *Chemical Reviews* 5, 1928, pp. 557–568.
[14] *Ibid.*, pp. 558–559.
[15] F. Hund, "Chemical binding", *Trans. Far. Soc.* **25**, 1929 pp. 645–647.
[16] J. E. Lennard-Jones, "The electronic structure of some diatomic molecules", *Ibid.*, pp. 665–686.
[17] L. Pauling, "The shared-electron chemical bond," *Proc. Nat. Acad. Sc.* **14**, 1928, p. 359.
[18] *Ibid.*, p. 361.
[19] For a comprehensive analysis of Pauling's work and a full list of his papers up to 1968 compiled by G. Albrecht, see *Structural Chemistry and Molecular Biology*, edited by A. Rich and N. Davidson, Freeman and Co. 1968. In the same volume see J. H. Sturdivant, "The scientific work of Linus Pauling," pp. 4–11.
[20] Pauling's unpublished Notebooks. Deposited at the American Institute of Physics, Archive for the History of Quantum Theory. *Notes titled "1928 London's paper. General ideas on bonds". There follows a note by Pauling: Here we have the first discussion of hybridization (pp. 14, 18)*, p. 22. Pauling's unpublished notebooks provide us with an insight about some of the early developments of the theory of the chemical bond. We notice extensive notes on the papers of both Heitler and London on group theory.
[21] L. Pauling, "The nature of the chemical bond I", *J. Am. Chem. Soc.* **53**, 1931, p. 1367; *Ibid.*, p. 3225; *Ibid.*, **54**, 1932, p. 988.
[22] Pauling interview, p. 16.
[23] G. Wheland, *The Theory of Resonance and its Applications to Organic Chemistry*, New York, John Wiley, 1944, p. 31.
[24] L. Pauling "Modern Structural Chemistry," in *Les Prix Nobel*, Stockholm 1955, p. 95.
[25] R. S. Mulliken, "The assignment of quantum numbers for electrons in moleculs. II. The correlation of molecular and atomic states," *Phys. rev.* **32**, 1928, pp. 761–772.
[26] R. S. Mulliken, "The assignment of quantum numbers for electrons in moleculs. I," *Phys. Rev.* **32**, 1928, pp. 186–222. p. 186.
[27] R. S. Mulliken, "Interpretation of band spectra. III. Electron quantum numbers and states of molecules and their atoms," *Rev. Mod. Phys.* **4**, 1932, pp. 1–86; "Electronic structures of polyatomic molecules and valence. I," *Phys. Rev.* **40**, 1932, pp. 55–62; "Electronic structures of polyatomic molecules and valence. II. General Considerations," *Phys. Rev.* **41**, 1932, pp. 49–71.
[28] *Ibid.*, p. 52.
[29] *Ibid.*, pp. 54–55.
[30] *Ibid.*, p. 60.
[31] *Ibid.*, p. 64.

[32] W. Heitler to F. London, November 4, 1935.
[33] W. Heitler to F. London, November 7, 1935.
[34] F. London to W. Heitler, November 6, 1935.
[35] W. Heitler to F. London, November 12, 1935.
[36] W. Heitler to F. London, November 12, 1935.
[37] *Ibid.*
[38] F. London to W. Heitler, October or November 1935.
[39] *Ibid.*
[40] W. Heitler to F. London, February 6, 1936.
[41] Heitler to London, October 7, 1936.
[42] *Ibid.*
[43] Interview with Linus Pauling, A.H.Q.P.
[44] L. Pauling, "The Nature of the chemical bond. Application obtained from the quantum mechanics and from a theory of paramagnetic susceptibility to the structure of molecules," *Journal of Am. Chem. Soc.* **53**, 1931, pp. 1367–1400. Quote on p. 1367.
[45] L. Pauling, *The Nature of the Chemical Bond*, p. 219.
[46] Nobel Prix, p. 92.
[47] Pauling, Private communication.
[48] R. S. Mulliken, "Electronic structures of polyatomic molecules and valence. VI. On the method of molecular orbits," *Jour. Chem. Phys.* **3**, 1935, pp. 375–378; p. 376.
[49] *Ibid.*, p. 525.
[50] *Ibid.*, p. 378.
[51] Interview with R. S. Mulliken by T. S. Kuhn, February 1, 1964. p. 17.
[52] *Ibid.*, p. 17.
[53] Interview with E. Wigner by T. S. Kuhn, December 4, 1963, p. 13.
[54] Hund to London, July 13, 1928.

ERWIN N. HIEBERT

# ON DEMARCATIONS BETWEEN SCIENCE IN CONTEXT AND THE CONTEXT OF SCIENCE

## PROLOGUE

As far as I can remember most historians of science of the post-World War II generation acknowledged the significance of examining scientific documents in conjunction with the contextual circumstances in which science developed. Historians, it seems, simply took the pertinence of contextual circumstances for granted. That perspective was endemic to the times and inherent in the questions being raised from within the discipline. The generation of historians of science and technology mentioned here were lured into history almost immediately after the end of World War II. For most of them, who were trained scientists, history became the vocation; the subject matter of these new endeavors nevertheless remained firmly rooted in the sciences.

Surprising as it may seem, the history of science as a profession is much younger than the history of humanistic branches of knowledge such as philosophy, religion, languages, literature, music, and art. For a discipline as recent and unorthodox as history of science – in which from the scientist's point of view the subject matter is perennially obsolete, even if it is said to be "interesting" (an historically muddy word) – one question looms large: Why should anyone, especially someone knowledgeable about science, want to become an historian of science, given that the life of science lies almost totally in the present and future?

It has been suggested, and plausibly so, that the motivations for studying the history of science as a distinct historical discipline in the late 1940s had a great deal to do with reflections about the war, the events leading up to it, its outcome, and prospects for the future. In scientific circles there were heated discussions concerning the military use of the atomic bomb. Within months of the Japanese surrender in August 1945 Chicago scientists were meeting on a regular basis to discuss the political and social responsibility of the scientist, civilian control of atomic energy, the economics of atomic power, the freedom of scientific information, relations with the Soviet Union, etc. These issues were aired and brought to the public in the *Bulletin of the Atomic Scientists of*



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 87-105.

*Chicago* – launched in February 1946 and distributed free of charge.

It was in this environment that the American history of science discipline was generated. Fortunately the strongest encouragement came from scientists who believed that the history of science was a very proper study to be engaged in and therefore worthy of moral and financial support. When in 1950 the National Science Foundation was established as an independent agency in the executive branch of the United States Federal government the history of science and the philosophy of science, alongside physics, chemistry, and biology, became eligible for financial support to conduct research and launch new teaching ventures. The resulting post-war expansion that took place in the history and philosophy of science enterprise in America undoubtedly contributed to the recognition of the high caliber of American scholarship in these domains at the international level.

Young scholars were enticed to take up graduate studies in the history of science. Others decided simply to abandon what they were doing in science and switch to the more challenging study of history. Most of these students ended up in academia. They often became involved en route in teaching introductory science courses or survey courses that covered the history of the natural sciences from "Babylon to Einstein". Their own scholarly research, by contrast, was given over to well-defined problem studies laid out more specifically by time, place, and scientific discipline. A substantial number opted to work in the period of modern history of science and because of their scholarly investment and involvement in science were attracted to the history of recent science.

## THE ARGUMENT

The objective in this paper is to examine the meaning and significance of "content" and "context" in relation to the acquisition and production of scientific knowledge. Observations and interpretations initially are offered to clarify the way in which historians of science of the immediate post-war generation attempted to come to grips with the conceptual components of their discipline. How did they deal with "science in context"? Expressed in this way the focus falls on the *content* of science as a study in its own right. Indeed there was a fairly sharp and tacitly understood line of demarcation between the *content* of science and the *context* of science, between the cognitive component of science and extra-scientific contingencies.

The fundamental distinctions between content and context nevertheless were not to be harmonized away as if they did not exist. The context was not to be treated as an incidental or trivial component of the historical analysis of the content of science. In fact, the historian's mandate was to weave content and context – admittedly essentially a matter of choice – into a meaningful, seamless, and more or less dense account of what really had happened in science. At another more philosophical and analytical level attention also was directed toward *understanding* what really had happened; and this invariably meant examining what had happened by undertaking a close examination of the context.

The point of view at that time may be characterized as one in which the need for firm mastery of some domain within the sciences was presupposed as a prerequisite to engaging in the history of the subject. This perspective undoubtedly derives from the fact that the earliest group of American scholars to enter the discipline, including myself, came to the subject out of strong background training in one of the sciences. This suggested that the history of science, like the pursuit of science, was to be understood as a no-nonsense enterprise. For example it would have been utterly foreign to the discipline to get mixed up with the trendy jargon that today sports with structuralism, literary criticism, the rhetoric of rationalization, social reconstruction, and post-modernism.

The historian of science was committed to assessment and criticism – always – but what that entailed mostly was a dense analysis of science in context. It was based on clearly articulated internal, science-oriented premises rather than on a too easily confounded juxtaposition of history and criticisms external to the discipline. Properly conceived there was no "outside", only an enormous wealth of contexts from among which it was mandatory to make some meaningful but limiting choices in order to get on with the task at hand. Examination of contexts never became an end in itself. Contexts were means to achieving ends, *viz.* accurate, critical, rich, and thick descriptions/analyses of science in context. In retrospect such an agenda may well be judged to be too narrow, too simplistic, not interdisciplinary enough, or too exclusively focused on the so-called "cognitive" aspects of science.

As the papers and discussions at the Corfu symposium have revealed, an analysis of the current configuration of the history of science, in the light of its own very recent genesis, merits a far more serious, informed, ventilated, and perhaps more focused examination than yet has been undertaken. It would be profitable to approach such an enterprise from a number of diverse perspectives on the discipline. In doing so it is conceivable that efforts to establish more meaningful demarcations between what is meant by content, and what by context, would serve to clarify some of the vital issues that have been hotly discussed and debated over the past two and a half decades. Historians, philosophers, and sociologists of science and technology all have become progressively alert to the impact of social, political, and other contextual/environmental factors on the *mode* of creating and shaping knowledge. Rightly so. To go a step further, it has been maintained not merely that the mode of creating knowledge but the actual creation or production of knowledge is the crucial issue. It is obligatory for historians of science to come to terms with these matters. They are not inconsequential.

The central problem that is given the greatest visibility in this paper revolves around a genre of arguments in which it is asserted by some scholars that the contextual factors of science, sometimes called the "social dimensions of science", enter not only into the scientific enterprise as process or mode of pursuit, but come to be embedded in the final product of science, i.e. as a part of the cognitive content. That position is put to test and challenged in this paper. It also is maintained, incidentally, that the expression "social dimensions of science" is a much too restrictive term for designating contextual factors if taken in any way to exclude the intellectual, philosophical, ideological, political, military, institutional, and economic dimensions.

It is worth pointing out at the start that virtually no historians of science deny that contextual factors play a major role in guiding the *process* and *mode* of acquiring scientific knowledge. By contrast, in the "strong program" context takes on a contradistinctive image and becomes part and parcel of the content of science. The Lakatosian hard core cognitive notion, with its traditional realist façade, then falls by the way. According to the strong program, scientific knowledge takes its rightful place alongside other forms of knowledge like literature, religion, and baseball lore – all being culture conditioned essentially in the way that science is said to be culture conditioned.

In general, "context" traditionally has been taken to be a consequential component of the scientific enterprise – a necessary part of telling how and why scientists have gone about doing science – even if that context was not seen to comprise any intrinsic component of the cognitive content of what scientists accomplish in doing science in a given context. The case for building context into the *historian's account* of science, without claiming that such context possesses the same degree of durability within the hierarchy of scientific knowledge as the content, did not strike us, as it does now, as being a position in need of apologetics. Historians of science plainly assumed that the analysis of context and its bearing on the content was an authentic part of the normal task of an historian. We were not at all partisan to constructing a reference manual of scientific accomplishments, or providing an abbreviated and more readily accessible and cleaned up account of the documents than those written by scientists. We were concerned about furnishing genuine explanations and elucidations of what scientists were up to in their work. How might one conceivably have been witless enough, we supposed, to ignore the circumstances in which thought experiments, scientific ideas, and experimental investigations were generated, tested, corroborated, announced, challenged, rejected, defended, transformed, and exploited? Who would have been unimaginative enough to assert outright that the contextual conditions for the growth of thermodynamics or bacteriology were unimportant in an historical study?

## ANCILLARY DISCIPLINES

It was recognized early on that the social, political, economic, and institutional accounts of science were domains of history closely akin to and therefore consequential for the history of science. Basically they were ancillary or auxiliary to the discipline, being pursued professionally on their own merits by social, political, and economic historians. The sociology of science as a closely related neighboring discipline generally was held in high esteem by historians of science. There was much to learn from sociologists of science. They often provided the historian of science with important components of the contextual matrix of science that helped elucidate what was nowhere else accessible in the scientific texts or in conventional history literature.

I well remember, for example, how students in our group stood in awe of the expertise with which Robert Merton documented and made use of the primary and secondary literature of science in order to support a specific sociological thesis. I specifically may mention a paper about "priorities in scientific discovery" written by Merton sometime in the 1950s. It was about how the so-called great men of science had argued and quarreled not about which discovery or phenomenon was true or how it had been shown to be confirmed, but about who had achieved it or announced it first. It struck a sympathetic note with an observation made a few years earlier, in 1946, by the American baseball player Leo Durocher: "Nice guys don't win ball games, they finish last". Competitions, jealousies, and ambition hold sway in the sciences as in baseball.

There were other disciplines ancillary to the history of science that perhaps were even more intimately related to the history of science craft than was sociology. We mention here primarily the philosophy of science and especially the history of the philosophy of science. It was reasonable to recognize the kinship of the history of science to the history of philosophy because history of science was assumed to be a sub-branch of the larger field of intellectual history that encompassed the history of philosophy, religion, linguistics, art, music, and so on. Philosophy of science was a considerable concern to some of us and its study accordingly was taken seriously. We mention a number of issues of mutual interest: questions related to realism and subjective idealism (what sociologists of science commonly have designated as relativism), the various strands of positivism (Comtian, evolutionary, critical), the scientific conception of the world of Vienna Circle philosophers, and the role of experiment in testing and establishing criteria for the enunciation of scientific claims and theories. A more sensitive issue would be the ostensible gulf that separates philosophers from scientists, and historians of philosophy from historians of science when it comes to interpreting and evaluating the merits and accomplishments of science at a given time or the directions that science ought profitably to pursue in the future. Finally, in those areas where philosophy, politics, and economics overlap I would like to call attention to the relevance, for the history of science discipline, of determinist philosophies and ideologies of history.

We dare not expand on these thorny and treacherous philosophical topics here except to offer a few comments about the corner of philosophical thinking that perhaps comes closest to impinging on the role of contextual analysis in the history of science, *viz.* determinism in history and the closely related political, economic and materialist interpretations of history. This is warranted because a goodly number of historians of science in our generation became quite involved in evaluating the merits and demerits of determinist conceptions of scientific discovery, technological invention, the essence and make-up of so-called scientific revolutions, and the great man theory of history.

## SCIENCE IN CONTEXT

An attempt has been made to draw a number of historically meaningful lines of demarcation between "science in context" and "the context of science". It is claimed here that the immediate post-war generation of historians of science generally pursued the discipline with a focus, *ab initio*, on the examination and analysis of scientific texts and other scientific records such as laboratory notebooks and scientific artifacts. Sometimes retrospective reflections by scientists on their own works and that of others were helpful for clarifying obscurities and omissions in the texts. Secondary works were consulted when accessible and pertinent.

Ancillary or supplementary to the *content* was the *context*; it was crucial for the examination of "what happened" and "why". This, as already mentioned, includes serious excursions into other branches of history, sociology, philosophy, religion. To be more candid, one acknowledged that anything contextual and any conceivable stratagem was seen to be fair game, provided only that it contribute to the clarification of *what* happened in science, and *why* it happened the way it did at that time, in a certain locus, and under a given set of contingent circumstances.

The success of such an approach necessarily rests upon a division of labor of sorts. The historian of science may well claim to be the expert with respect to the content of science. Unable – no matter how proficient – to master all of the contexts within which the content conceivably acquires new significance, the historian of science will be compelled to search for, borrow, or pilfer ideas from ancillary disciplines on the contextual periphery to his or her expertise. It normally was opportune to become sensitively involved in as much context as would turn out to be germane to a meaningful examination of the content.

And how, one might ask, does one proceed to select specific contexts to explore? The *modus operandi* being advanced here is to forge ahead on the assumption that a thorough mastery of the content – namely an intimate technical and intellectual grasp of the texts – when coupled with a good measure of historical perspèctive and intuition, will go a long way toward defining the direction in which to proceed contextually. There is, assuredly, the *proviso* that the alert historian of science will make every effort to keep an open mind concerning the merits of choosing to examine alternative contexts – since not all possible contexts can be examined.

Still, it is unwieldy, if not impossible, to venture out in all conceivable directions and undertake an $n$-dimensional contextual analysis all at once.

Carl Sandburg strikes the right note in his 6-volume biography of Abraham Lincoln, a work that was 30 years in the planning and took 13 years to complete. He says:

> The chronicles are abstracted from a record so stupendous, so changing and tumultuous, that anyone dealing with the vast actual evidence cannot use the whole of it, nor tell all of the story. Supposing all could be told, it would take a far longer time to tell it than was taken to act it in life. Therefore the teller does the best he can and picks what is to him plain, moving and important – though sometimes what is important may be tough reading, tangled, involved, sometimes gradually taking on interest, even mystery, because of the gaps and discrepancies.

Choices always are mandatory; and because of the great mass of potentially relevant contextual information normally available, the choices stand in need of rigorous selection criteria. Optimally not all choices can be intellectually compelling, for not all unexplored historical projects are worth pursuing. When formulating a problem in the history of science, for example, the examination of a scientific text or a group of texts, a discovery, or the integration of a segment of science, it is true that one seldom reaches satisfactory insights merely from an examination of the texts or papers.

The austere and sometimes messy message from the text invariably takes on new meaning and life when examined against the background of a potentially relevant context. Then one may recognize that the heart of the matter in need of explanation may lie at the periphery. Similarly the meaning seldom falls into place simply examining ambient scientific developments of the times. On occasion, essential insights are acquired in devious and sundry ways: by deliberately searching out a scientist's religious or ideological commitments; by adopting an intuitively plausible assumption about how a scientist is expected to act in a given situation; by identifying specific economic incentives that serve as motivation for a scientist to engage in some activities and not others; by recognizing political pressures on a scientist to conform to established norms; or from personality traits in a scientist that contribute to or interfere with the investigations. The essential point is that the text itself invariably provides some clues and a starting point for engaging in meaningful analysis and elucidation of the texts.

On occasion one is blessed with a well-formulated and manageable problem and research strategy because of intimate familiarity and mastery of the surrounding terrain. Ordinarily, however, the historian of science is compelled, as the scientist is in science, to build the ship not in dry dock but while at sea. Questions are raised. Speculations abound. Answers are mostly fragmentary. Texts are singled out for examination. Few internal criteria exist. Reading the texts may lead to other more puzzling questions, partially understood meanings, statements that seem not to add up or make contextual sense. Perhaps the apparent confusion in the text and/or lack of comprehension of the text by the reader should provide the incentive not to abandon the search for meaning but to dig in. Conceivably an anachronistic reading, or a myopic or too-modern ideational mental framework, and not the text, can be shown to be the locus of misreckonings.

Understanding may accompany interpretation but further understanding also may interfere with or modify what one originally was trying to understand. Clarification must then be sought outside the text somewhere in the contextual wilderness. Immersion alternately in the text and in the ostensibly relevant contexts eventually may show the way to genuine historical comprehension. But the search for a final rock-bottom understanding remains elusive. History is after all "very chancy" (Samuel Eliot Morrison) and looks quite different from the perspective of different historians.

In order to explore the methodological issues stirred up in the above comments we may compare two dissimilar but prototypical entrées to fundamental research in the history of science. One feasible directive would be to probe an approach aimed at investigating the contextual circumstances in which some aspect of science *X* developed. Examine, for instance, the social or political history of the genesis of nuclear physics in the Soviet Union in the 1930s and 1940s. The contextual issues loom large; the explorable facets of the subject essentially are unlimited. How much information need be mastered in order to construct a meaningful social history of the subject? Will the reconstruction lead to insights about the mode of evolution of nuclear physics as such, or show which aspects of the Soviet experience were novel and which were acquired by transfer from other countries? Is there a methodology, a sociological systematic, that will guide the investigator into fruitful choices about which contexts to consult? More crucial, how would one advance from the context to the content of nuclear physics?

Take as an alternative approach the examination of a specific scientific problem with the focus on a study of science in context. To cite an example: How might one explore the way in which experimental investigations at low temperatures contributed support for the quantum theory during the second decade of this century? Examine the contents of the classic scientific papers on the subject by Born, von Kármán, Einstein, Nernst, Debye, Sommerfeld. Are there additional primary sources? Are they accessible, significant, readily manageable? Unquestionably a sympathetic immersion in the science and history of specific heats and quantum theory will furnish illumination. Avoid being excessively whiggish about it; yet it cannot be evaded totally because the decision to consider the worthiness of the topic initially was tempered and tainted by the fact that the quantum theory of specific heats today is bound up with widely accepted concepts of basic modern physics.

In any case, examining and searching out the content of science has become the inescapable starting point for the research. Perhaps the focus, content-wise and contextually too, is unduly narrow or prescribed and the decision seems appropriate to move toward broadening. Inevitably other questions concerning the newly established parameters of content and context will surface. As already mentioned, texts will not on their own be self-explanatory. Contextual issues will enter the analysis, shift the content focus, and be exploited wherever their pursuit leads to illumination of the original problem. More likely the original problem and the questions raised will go through a metamorphosis that results in a fertile investigation of the scenes of scientific inquiry at variance with the original point of departure.

One thing is certain. In our generation of the 1940s and 1950s the initial focus was on science in context. Indeed, it was in the analysis and rigorous examination of scientific texts that the history of science captured its intrinsic excitement. Nothing at that time could have surpassed the inspiration, expertise, and intellectual motivation with which Marshall Clagett, for example, enticed students to enter into a meaningful appraisal of the natural philosophy of the Middle Ages as exhibited, for example, in treatises devoted to the study of motion, the infinitely large and small in magnitude and enumeration, and analysis of the intension and remission of forms. In seminars the students read, puzzled over, argued about, and sought out the historic message of works by John of Dumbleton, William Heytesbury, Richard Swineshead and Nicole Oresme – all of whom flourished between 1320 and 1350. What's more, the seminars gave some of us, whose career objectives were anchored in 19th and 20th century history, an extraordinary exposure to sound and clearminded skills and historically vigilant attitudes toward the analysis of scientific texts. It became evident that basic techniques and methods for study of the medieval period were not as far removed from the study of more recent periods than one might have assumed.

Equally valuable was the opportunity to get rid of the myopic vision that "science" always had been cast more or less in the mold of contemporary, modern, Baconian or Galilean–Newtonian science. Our studies led to an appreciation for the immense differences that have served to identify so-called "science" and views on "nature" in disparate times and places: alternative time-bound customs for formulating and resolving seminal questions; the establishment of acceptable criteria for presenting logically unassailable arguments; and the degree of importance given at different times in history to the role of experimental verification and theoretical reasoning.

To recapitulate: An approach that accentuates the examination of "science in context" rather than the "context of science" has here been given pride of place. Such a bias has been earmarked for consideration if for no other reason than that the content of science is more readily and more clearly identifiable than the essentially open-ended context of science. To proceed from context to content, if possible at all, has been thought to represent a shapeless *von oben bis unten* stratagem that proceeds – dare we say it – from a sociological presupposition or from awkwardly formulated philosophical premises and generalities to specifics rather than vice versa. Such an approach if stretched to its limit becomes a precarious act of probing, choosing, and maneuvering the scientific content into an "interesting" narrative that, however, cannot claim to be history.

For fear of giving the impression that it is a simple matter to set up a clear-cut division of labor between scholars who find their home-base in the content of science and others who locate themselves more comfortably in the study of societal components of science, or even identify science more singularly with social constructivism, a word of caution is necessary. In both of these sharply drawn perspectives – which at one time were labelled the internal and the external – there are hidden agendas and unspoken or inadequately formulated metaphysical underpinnings. A huge literature is available on these issues and the historical problems that they generate. We shall not take the time to raise any of them here now. Rather, it would seem appropriate to examine the question of the content of the context.

What shall be taken to count as contextual? To simply refer in generalities to scientific context is not satisfactory; the aim is contextual illumination of the content. Unfortunately a multi-dimensional analysis of contexts is unmanageable, for anything and everything might be said to be relevant. Everything is subject to analysis, but everything cannot be subjected to analysis at the same time. Indeed, views on what may count as significant context for shaping the history of science have changed enormously during the last half-century.

## THE HISTORY OF IDEAS

In order to present additional reflections from first hand experience about how historians of science of my generation dealt in the main with contextual issues, it is imperative to offer some comments about the history of science as a branch of the history of ideas. The single person who most understandably championed the "history of ideas" in our generation was Alexandre Koyré (1892–1964). This Russian-born, Göttingen- and Paris-trained dual career historian-philosopher, who became directeur d'études à l'École pratique des Hautes Études in Paris in 1930, emigrated to various countries during the German occupation of France. After 1956 Koyré spent half of his time in the United States (mostly Princeton) and the other half in Paris.

Koyré's scholarship, antipositivist in tone, encompassed a wide range of disciplines from ancient, medieval, and modern times. They included mathematics, physics, philosophy, religion, mysticism, and alchemy. The explication of the history of ideas and the analysis of concepts – scientific and philosophical – was commanding in all his studies, and exerted a profound influence especially on American students of the history of science after 1945.

Koyré was averse to compartmentalization. His studies were animated by the search for a unity that links science, philosophy, religious thought, and mysticism at the conceptual level. The evolution of scientific thought and the important transformations and revolutions taking place in the sciences since antiquity – sometimes stretching over very long periods of time – were seen by Koyré to be basically intellectual, spiritual, and trans-scientific; they therefore, to some extent, were dislocated from empirically acquired evidence.

It is apparent from all of Koyré's writings that *texts*, as seen within their historical, scientific, and spiritual contexts, were the all-important *terminus a quo* for scholarly research. His métier was scholarly research in which the goal was intellectual history or the history of ideas: *Histoire de* *la Pensée*. A comparable approach essentially characterized an entire school of late 19th-early 20th century French scientist-historian-philosophers such as Paul Tannery, Emile Meyerson, Arthur Hannequin, Pierre Duhem, Léon Brunschvig and Pierre Boutroux.

Koyré's perspective on the history of science, as "idea history", has been brought into this paper to help disentangle and clarify for post-Koyré scholars some of the knotty issues raised in connection with content/context demarcations. Besides, Koyré epitomized for me personally a first encounter with the history of science at a serious, scholarly, and non-popular level that extended beyond the reflective and historical accounts customarily written by scientists. His approach, nevertheless, did not represent the only or even the dominant influence on the circumstances connected with the launching of a history of science career. Additional embryonic, gravitational pulls toward history and philosophy came from other historians and philosophers of science and most directly from queries and puzzles encountered and formulated in previous and later studies in science itself.

We may first demonstrate in specifics where the contextual analysis of history of science was positioned for those of us who were baptized into the new discipline at the end of the Second World War. For the duration of the war my own graduate studies in physical chemistry and chemical physics were interrupted by work in the Chicago area as a physical chemist engaged in experimental research on the separation of boron isotopes 10 and 11 – a small but crucial component of the objectives of the Manhattan Project. The assignment was a nice mix of experimental challenge and risky pilot plant stage testing of unconfirmed theoretical predictions.

There were fascinating frontier agendas with military urgency, but the work, although morally disquieting, was "technically sweet" as Oppenheimer has remarked. At the University of Chicago, after the war and release from the Project, I became immersed in experimental investigations at the Institute for the Study of Metals. Research on entropy of orientation in the surface of liquid copper by measurement of the thermal coefficients of surface tension, seductive as it was, did not entirely fulfill the spiritual space of vision or thirst for new scenes of intellectual exploration.

Fortunately the University of Chicago at that time was a stopping-off place and haven for an international constellation of scientists, mathematicians, composers, philosophers, *Alleswissers*, humanistically-minded foreign visiting professors, and loquacious political emigrés. The post-war graduate students became very much engrossed not only in formally attending University classes in the sciences and mathematics but in taking advantage of a great wealth of heterogeneous intellectual erudition, mostly of European vintage. Above all it was an exposure to large doses of cultural history sandwiched in between expansive schemes for world government and global internationalism.

The love affair with science and an unambiguous identification with the world of science remained in place, but there were other awesome universes beyond science with their own magnetizing force fields. The lure was irresistible. There were courses on Plato, Descartes, Leibniz, Whitehead, logic, contemporary German philosophy, economic history in relation to religious thought (John U. Nef), analysis of methods in the physical sciences (Rudolf Carnap), history of mathematics and astronomy in Mesopotamia (Willy Hartner), and less pertinent to science but not entirely uninteresting, colloquia on sociology (Edward Shils, Yves Simon).

The most lasting impressions in my own training in Chicago came through a lecture course on "Scientific Thought in the Age of Newton" by Alexandre Koyré. I began to read everything Koyré had written. Shortly thereafter a switch in outlook led, with the help of financial aid, to graduate studies pursued at the University of Wisconsin simultaneously in the history of science, physical chemistry and chemical physics. In 1961/2 at the Institute for Advanced Study in Princeton, the opportunity arose to engage in almost daily discussions with Koyré.

What, we may ask, was distinctive about Koyré's approach to the history of science? First and foremost, as noted above, his scholarship was rooted in the conceptual analysis of scientific texts and commentaries by contemporaries on those texts. That approach was most compatible with my scientific background. There were as well very strong links with the history of philosophy in general and philosophical idealism in particular. They drew heavily on the works of French, German, and Russian intellectuals. For Koyré, whose philosophical horizons were immense, there also were cautiously drawn conceptual sympathies with Plato (they could not possibly have attracted my fancies after having studied with Marshall Clagett), Jacob Boehme (the anti-authoritarian and critical outlook of the Anabaptist visionaries was uniquely appealing), Descartes (I never could tolerate the mind/body dualism), and more weakly stressed by Koyré, Husserl's phenomenology (not at all my cup of tea). Altogether it was a mixed bag of heavy intellectual dimensions – provocative and not to be taken dispassionately – but also not entirely attuned to an historical and philosophical perspective that had been preoccupied with, and was moving ever closer to, the work of scientists and philosophers of science whose halcyon days came after 1850.

Koyré's philosophical emphasis was strong on epistemology and on a metaphysics that had to be out in the open. There was no closet-Hegelianism (although Koyré had mastered Hegel chapter and verse) and certainly no explicit reference to an all-encompassing model as such unless it served to underline the historical consistency within an argument under consideration. His anti-reductivist counter-constructivist outlook with its openness to a possible pluralism of interpretations appealed to many of us who had come out of the sciences. Above all, Koyré was intent on understanding *in extenso* what natural philosophers were arguing about in their texts. Accordingly, he made it clear to his students that to comprehend what was being said and meant in the texts carried an obligation to fight one's way intellectually to an understanding of texts, commentaries on the texts, the variant texts, the disputes about the texts, and the intellectual environment in which the texts were presented, argued about, and defended.

What then did Koyré take to be the content of science? The essential content and most permanent, precious, and historically pertinent part of the entire scientific enterprise was for him the theoretical component. That is not to say that he underestimated or neglected the historical dimensions of experiment and observation when they impinged on the validation of theory. Besides, his conception of theory was expansive enough to reach far beyond the high points of scientific conceptualization and discovery on the part of the major figures. The whole scientific story was to be told, or at least the whole accessible intellectual story, but always with the emphasis on how theory emerged from out of and had moved beyond the empirical investigations.

The Koyré model for pursuing the history of science was uniquely significant in America for its role in launching history of science as a professional discipline that received the moral and financial support of scientists and historians. It also was philosophically provocative and contributed to a revival of appreciation for philosophy of science on the part of historians of science. By insisting on the mastery of the technicalities of science as a prerequisite to doing history of science Koyré was right on target. In time many historians of science came to sense, however, that his view of the discipline was cast too rigidly in a nature/reason framework in which the *context* of science was taken to be no more than an extension of the intellectual aspects of science. As for sociology of science or the social context of the conceptual dimensions of the history of science, Koyré spoke of them as being "interesting" but not very important. The disposition, with its center of gravity in the intellectual and theoretical component of science, would suggest that too many significant aspects of science – essential motors of scientific advancement and seminal sources for explaining and elucidating the scientific content – had been overlooked in the development of science when viewed as a culture-conditioned enterprise.

Koyré held firm beliefs about man's ability to uncover underlying truths about the world. The world was taken to be real, ordered, unitary, and intelligible to the human mind. There was always in what he wrote a robust undercurrent of aversion to positivism. He believed, nevertheless, that the world of nature becomes conceptually intelligible to the human mind through the doorway of concepts invented and fashioned by the mind – a position that Einstein would have endorsed. The history of science ultimately was a true "history of ideas". The truths of science understandably were discovered and intellectually generated within various contexts but for Koyré these mainly were intellectual: philosophy, religion, and higher learning.

## BEYOND THE HISTORY OF IDEAS

Koyré captivated many in America with his French charm, the boundless depths of his learning, and the scientific, historiographic, and philosophical expertise with which he approached and executed the history of science. He carefully, logically, had thought out scientific matters where scientists were accustomed to be on the alert. Scientists valued and could understand his writings even if not the depth and finesse of his historical and philosophical erudition.

But there were problems. Three criticisms may be underlined: First, the *content* of science is not just and perhaps not always mainly ideas and theories. There are vast domains in the sciences, and especially in the life sciences, where scientists, so-to-speak allow the surface of things – the events that occur and those that are fabricated – to speak out loudly on their own before premeditated rationalizations and theoretical machinations take center stage. To neglect or even to play down experiment and observation in an historical account would seem to create a somewhat distorted image of what happens in science. Most of the time working scientists do not tackle or answer really big questions – at least not directly. Most advances, discoveries, and revelations come into focus by chopping away at small tractable problems. The big questions of course are theoretical, and are further removed from the empirical level; they are closer to mental constructions, but invariably emerge from out of the mire and blood (as Bas van Fraassen says) of empirical research.

In second place the *context* of science, and therefore the historical examination of science in context cannot, as Koyré seems to suggest, consist primarily of ideas that envelop only the intellectual constituents of other disciplines such as philosophy and religion. The context is many things; the social, political, and economic contexts, for example, cannot reasonably be said to be "interesting" but not very relevant to how science grows and advances.

In third place, history of science cannot reasonably be a matter of discourse that oscillates back and forth between information about the nature of things and the reasoning machinery of man that creates and invents concepts and theories. If anything, the business of doing science, exploring "the nature of things", is a delicate balancing act that seesaws back and forth between experiment (actual experiments and thought experiments), observations, the acquisition of empirical information, theoretical constructs, speculations, and ratiocination. On this score, *viz.* where the potentially imaginative, but also troublesome and tricky problems and the give-and-take between the theoretical and empirical take place, the critical positivists, in my opinion, were on epistemologically sounder ground. I would suggest that scientists and philosophers like Ernst Mach, Niels Bohr, Ludwig Boltzmann, Philipp Frank, and Rudolf Carnap had historically and philosophically more pertinent things to say to historians of science about the reciprocity between experiment and theory. The construction of bridges to link experiment and observation with acceptable conceptual and theoretical ideas and frameworks has been no trivial matter in the life of the scientist. Nor has it been inconsequential for the historian of science to search out the complex maneuvers that scientists engage in to do so.

Much has happened in our discipline since the high tide of Koyré's influence in the 1960s. It would be out of place and also far too ambitious an undertaking to include such matters in this paper. However, before leaving Koyré behind it should be acknowledged that his strategy of buttressing scientific *content* in order to feature the primacy of scientific texts does live on in the scholarship of those historians of science who believe that proceeding from content to context, and not the reverse, furnishes the most reliable and efficient way in which to study "science in context". We may note that such an approach was normal fare for virtually all historians of science who launched their careers in the discipline at the end of the war. They advocated the central significance of serious, in-depth, scholarly examination of scientific texts and documents. Koyré was unique for having combined insistence on studying the texts with a breadth of philosophical erudition somewhat foreign to American scholars.

## FINAL REFLECTIONS

In this paper an attempt has been made to present a position that champions the primacy of the scientific texts – but not, as should be evident, at the expense of slighting or underrating the importance of the context of the texts or by ignoring the activity, behavior, and beliefs of the scientists who constructed the texts. In the process of spelling out this argument the emphasis has fallen more on my own career than I would have wished. To be honest and open about it, so I think, is desirable and decent. For I tend to believe that views, opinions, and passions are shaped powerfully by the events and circumstances that persons are drawn into and that engulf them during formative years. An acquired historical consciousness tends to intermesh with axes of bias and methodological preferences for engaging in history.

In our generation, scientific texts represented the *sine qua non* and consequential backbone for pursuing the history of science. The same insistence on mastery of the primary scientific documents held for the philosophy, sociology, and sociology of science. The history of technology had its own agenda of texts and artifacts but rested so securely on contextual matters that content/context demarcations at times got washed out. One need not lay rigid claim to the existence of a content that is resting securely somewhere in society apart from and outside of context. Content and context are "out there" somewhere together – sufficiently together, that is, to make it risky business to establish inflexible lines of demarcation between them.

Nevertheless, some demarcations can be made in a clearcut way. If, for example, the task is to examine some aspect of Maxwellian electromagnetic theory one cannot very well avoid studying what Maxwell said about the subject. It also would be unwise to ignore what Poincaré, Lorentz, Boltzmann and a few others wrote on these matters. What needs to be said to avoid getting trapped by inconsistencies is merely that the *context* of Maxwellian theory is not the place to begin examining his theory. We belabor the obvious: master the content and sooner or later one will be motivated to pull contextual components of the theory into the analysis. The content will help to show the way. The content is always contextual.

"Science in context", it has been suggested, is where the hard core of history of science auspiciously may be anchored. The "context" of science, the social history of science, the political, economic, institutional contexts are more than "interesting". The contexts are crucial components in the pursuit of "science in context". With reference to what Professor Mamchur said yesterday in her lecture on the cognitivist and the sociological reconstruction of science I see that I have come down rather decisively in this paper on the cognitivist side. One might on the strength of this appraisal anticipate, but not arrogantly so, that sociological reconstructions of science become a resource for the historian of science to take advantage of when the reconstructions have something of significance to contribute to the analysis of the content.

Let us rather put the issue more directly and crudely by capitalizing on a phrase that Jed Buchwald whispered into my ear during our sessions yesterday: "He who lies down with dogs gets up with fleas". "True", I responded, "but it might be worth it. One can always resort to the use of flea powder."

*Harvard University*

ARISTIDES BALTAS

# ON THE HARMFUL EFFECTS OF EXCESSIVE ANTI-WHIGGISM[1]

## INTRODUCTION

You don't need me to remind you that the coming of age of the history of science has relegated to the discipline's prehistory all the enlightened amateurs' attempts to collect past curiosities which, under some unspecified criteria, appeared as 'interesting'. By the same token, "history of science" has become a legitimate academic discipline in its own right, one possessing a well-delimited subject matter. History of science is now the methodologically principled study of past scientific achievements, in light of all factors which determined their production, their acceptance, and the diffusion and which gave them meaning and significance in the first place.

This is to say, that in ceasing to be just the "repository of anecdote and chronology" – to borrow the opening phrase of Kuhn's *The Structure of Scientific Revolutions* –, history of science has set for itself some well-defined methodological principles. For various reasons which are not difficult to locate, perhaps the most fundamental of these principles is that which stipulates the avoidance of "whiggist" interpretations of past science: To approach past science *correctly* – both from the cognitive and moral or political point of view – it is indispensable to bracket currently dominant views on the relevant issues and treat past science in its own setting and in its own right. Practically everyone present at our conference would agree that at least since Butterfield, professional historians of science have scrupulously followed this principle to the best of their ability and ingenuity. Perhaps it is not even excessive to maintain that the discipline has come of age *because* it conformed to this methodological requirement. Accordingly, if today, at a gathering uniting some of the most prominent historians of science, an "outsider" presumes to raise a question about the possible "harmful effects" of an anti-whiggist position, he risks appearing not simply as coming from another field but from another age. Even if he is cautious enough to qualify the anti-whiggism he objects to as "excessive" at the outset, he will not be saved.

However, be this as it may, I will ignore such risks and proceed. In so doing I am not, of course, presuming to teach historians of science their



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 107-119.

job. The only justification I can appeal to for my stubbornness is that I believe that 'pure', unqualified, anti-whiggism is *impossible* to achieve. And that if this is indeed the case, then explicitly defending it cannot fail to misrepresent the related methodological principle which is *effectively* at work within the discipline of history of science. It follows that such a misrepresentation may induce "harmful consequences" to the very effectiveness of that discipline taken as a whole. In other words, historiography of science cannot do without a certain, minimal, whiggism – which I will try to define in what follows – and that recognizing this may prove methodologically helpful. This is all I want to argue for in the present paper.

A preliminary warning and a corresponding plea for excuses: Given the well-known incontrovertible limitations of space and time, I can only try to defend my position in the most crude and schematic manner. Accordingly, I am obliged to ask you to excuse my oversimplification of issues which certainly deserve better treatment. Having said this, and without waiting for your reply, that is, taking your assent for granted, let me start by offering some quite elementary distinctions and 'definitions'.

## 1. WHIGGISM AND ANTI-WHIGGISM IN GENERAL HISTORY, AND IN GENERAL

The most simple, not to say simplistic, definition of the discipline of general history is to say that it constitutes the study of the past. For the purposes of our discussion, we can distinguish two broad classes of methodological (but not merely methodological) positions regarding such study. On the one hand, there is the classical whiggist position which conceives the past as, somehow, the preparation for the present. In its most extreme formulations, it considers explicitly the corresponding transition as amounting to the upward 'evolution' or 'progress' of what is 'less developed' toward what is 'more developed'. As the relevant literature amply attests, this position has always brought with it the political, not to say moral, connotations and overtones celebrating contemporary Western culture as the culmination of the progress of humanity at large. 'Developed' versus 'underdeveloped' countries or economies, 'primitive' versus 'modern' (not to mention 'postmodern') societies, and 'savage' versus 'civilized' modes of thought and behaviour have been characterizations which are directly or indirectly related to whiggism. These characterizations have constituted and continue to constitute battlefields not only within the discipline of history, but also within the disciplines of economics, ethnology, and social anthropology. These characterizations also stretch outside the boundaries of established academic fields to all kinds of issues from feminism to 'political correctness'. I must add, however, that for the past two or three decades this position – at least in its crudest forms – has not enjoyed the support of most of the intellectuals implicated in the relevant battles.

Countering it, stands the standard anti-whiggist position. Within the discipline of general history, while trying to establish causal links between the past and present (something it usually shares with its opponent), anti-whiggism refuses to celebrate the present over the past. In the related disciplines of ethnology and social anthropology, the 'here' is not privileged over the 'there', the 'familiar' over the 'foreign'. In all concerned disciplines, the anti-whiggist scientist (in this quasi-ecumenical sense) considers the cultures or subcultures she is studying – whether of the past or the present – as self-contained entities which by themselves mould their proper modes of life and behaviour, and by themselves define their proper epistemic as well as moral norms, standards, etc.

Having said this, it follows that the debate over whiggism and anti-whiggism (in the same general sense) cannot fail to arise – in one or another form – in all disciplines where an investigator is obliged to build a bridge to something alien, or at least significantly different from her own culture. In respect to all these disciplines, anti-whiggism presents itself as a *specifically methodological principle* which forbids the uncontrolled importation of the investigator's prejudices, presumptions, and presuppositions to what she is studying. At this level, anti-whiggism presents itself as blocking the corresponding biases, thus sustaining objectivity, both from the moral and epistemic points of view. However, things are not that simple. On the one hand, whiggism and anti-whiggism cannot be restricted to the level of methodology for they cannot escape, as we implied above, a marked political or moral dimension. In particular, and within the present conjuncture, the political dimension of the anti-whiggist position is 'democratic', upholding fairness in respect to the past or to the different. On the other hand, the objectivity that anti-whiggism is designed to defend it at best ambiguous. Even if its presumed blocking of all biases does leave open some access to – and thence allows assessing – what is alien, such assessment risks to amount to unqualified relativism – again both moral and epistemic – and thus, in the final analysis, to no assessment at all.

It is evident that at this general level of formulation, the issue is extremely difficult to handle if it is not simply untractable. Happily enough, nothing obliges me to confront it as such. My only concern here is history of science and it is to this that I now turn. However, in order to make the transition, let me just state the position of principle that I am inclined to take: 'Perfect' or 'absolute', methodological anti-whiggism, although it *is*, with some qualifications, *highly desirable* from the political and perhaps also from the specifically moral vantage point, is *unattainable* within general history as well as within all disciplines studying what, in some sense, is significantly different from or foreign to the culture of the corresponding investigators. The only methodological guarantee we can hope for in these disciplines can only arise from our efforts to assess, as precisely as possible, the nature of this impossibility and thereby the limits this sets upon such desirability.

In a few words, the argument supporting this inclination can be stated as follows: Whether we like it or not (and, while largely unconscious of at least the finer details of its workings), we are constrained to reason from within the 'framework' determined by the culture we are in. This is to say that our way of looking at things is determined by a set of 'hidden' presuppositions, prejudices, presumptions, etc., which determine in a certain sense both our perception of facts and our investments of value in such a manner that they cannot be eradicated by even the most thorough specialized training. And the big methodological question that arises at this point – which I will not presume to confront here in its generality – is what can we and what ought we to do with this.

## 2. ANTI-WHIGGISM IN HISTORY OF SCIENCE

In terms analogous to those we employed above, that is, terms which are so general as to be quasi-empty, we can say that history of science is the disciplined study of past science. But even this overly simple if not simplistic 'definition' is beset' with formidable problems. For example, it already presupposes that 'science' is or can be a well-defined category while, at the same time, it leaves open the question whether such a definition is indispensable to the discipline of history of science so that the latter can carry on with its tasks. And we know, of course, through the teachings of philosophy of science, that such questions are far from having received definitive or at least passably uncontroversial answers. I let this pass for the moment, but I will return to it in what follows.

Speaking very schematically, if we take for granted momentarily that, *grosso modo*, we know what 'past science' refers to, the study of 'past science' involves, on the main, two things: First, the study of various parts or chunks of it in themselves, i.e. the study in its own right – independently of how historians of science may wish to spell this out (e.g., among other specifications, from an internalist or an externalist perspective) – of various past constellations of views, of various past theories, etc which are considered as 'scientific' (in most cases, as a matter of course). Second, the study of changes and transitions in 'past science', leading from one such particular theory or constellation of views to another.

If what precedes is correct, I am entitled to ask the following question: What *particular* parts or chunks of 'past science', and which *particular* changes or transitions thereof does (or can) the historian of science *effectively* study?

I maintain that there can be only one possible answer to this. Given her training, background knowledge, particular field of study, specific disciplinary requirements, current disciplinary focuses, etc, the historian of science, *out of what is or can be made accessible to her*, concentrates on what she finds *most interesting*. And if we inquire further both about the conditions determining what is or can be made accessible and about the criteria assessing what is interesting, there is again, I hold, only one possible answer: What is or can be made presently accessible is what *the present* has preserved, in one form or another, of the past. And what, out of this past, can be assessed as interesting are only the items which have a direct or indirect, immediate or mediate, positive or negative bearing on *present* science or, more generally, on *present day* concerns. Even if the historian of science's object of study is something extremely foreign, as a striking Paracelsian curiosity, this object can be precisely assessed as a *curiosity* only in respect to *present-day* ideas and values, and only by the explicit or implicit employment of *present-day* criteria.

Our first conclusion thus looks inevitable: The historian of science cannot escape a certain whiggism even from the outset, that is, in the very process of picking out her object of study.

However, this conclusion does not exhaust the issue. The inherent limits of methodological anti-whiggism in historiography of science do not concern only the moment the historian sets to work, but pervade all aspects of such work. It is to the examination of these limits that I now turn.

## 3. ON THE LIMITS OF ANTI-WHIGGISM IN HISTORIOGRAPHY OF SCIENCE

To state directly my conclusion, I maintain not only that there are inherent limits to methodological anti-whiggism but, also, that these limits are set by the fact – irrespective of whether historians of science are conscious of it and irregardless of how anti-whiggist they proclaim themselves to be – that present science acts as a certain kind of *norm* regarding the overall study of past science. The following argument tries to explicate the meaning of this thesis and to present an outline for its defense.

1 The culture we are embedded in – in the widest possible sense – constitutes a certain kind of 'framework' that constrains, in a largely unconscious manner, our ways of perceiving facts and investing values. This framework provides the set of explicit or implicit categories, ideas, representations, prejudices, presumptions, etc, that accompany our socially and historically determined practices, endow them with meaning and significance, and thus, allow them to be effectively carried out.

2. Our generic biological, psychological, and social constitution forces all our reasoning to rely heavily on unconsciously held 'premises' and 'lemmas'[2] which constrain it – in the sense of forming the horizon of our perception of facts and investments of value – while remaining hidden from view.

3. We can say that the set of all these 'premises' and 'lemmas' makes up the 'bedrock' on which our cultural framework rests. This implies that this bedrock provides us with our sense of the 'obvious', the 'self-evident', and what we take for granted as a matter of course. This bedrock is, in a sense, multi-layered. At or near its 'surface' lie the presumptions and prejudices that new experiences, struggle, debate, ideological turnovers, etc., can shake, unearth, modify and/or replace without changing the corresponding culture as a whole. On the other hand, it is the 'core' of this bedrock that gives a given culture its identity.

4. As, in the last analysis, all reasoning is based on what remains beyond justification or dispute,[3] the effective functioning of this bedrock makes it impossible (as will be qualified in what follows) to radically escape this framework – almost in the same way it is impossible to escape one's skin or flee one's brain.[4]

5. The impossibility in question amounts to saying that the 'premises' and 'lemmas' which make up at least the core of the bedrock of our cultural framework are effectively *unassailable* by the means which this framework itself can provide. Conversely, we can make sense of things only on the basis of this bedrock, that is, only in the terms of this framework itself.

6. It follows that we can make sense of a *different* framework – that is, of texts and behaviours lying beyond the scope of our *ordinary* understanding – *only* in the terms of our *own* framework.

7. That some text (or behaviour) belongs to a different framework means that it rests on different bedrock 'premises' and 'lemmas'. It is these that place it beyond our ordinary understanding. Accordingly, in order to make sense of it, we have to locate – beyond the text or behaviour's 'surface', because the text or the behaviour itself does not render them explicit – these 'premises' and 'lemmas' and expressly formulate them.

8. This, however, is not as simple as it sounds.
   a. First, the scientist's acceptance that a text (or behaviour) lies beyond her ordinary understanding signifies that she has educated reasons not to dismiss it as pure nonsense. And this, in turn, implies that she is willing to undergo a process that will reveal some of the presumptions and prejudices that *she* is unwittingly harbouring, precisely those that prevented her understanding in the first place.
   b. Second, as these presumptions and prejudices lie on the surface of the bedrock on which her own framework rests, trying to unearth the hidden 'premises' and 'lemmas' of the text or the behaviour under study amounts to digging at *that* bedrock. This is the self-critical (anti-whiggist) practice required by any scientist who is studying what is alien or fundamentally unfamiliar.
   c. Third, success in this work amounts to two things: 1) to the uncovering and explicit formulation of the hidden 'premises' and 'lemmas' of the text or behaviour under study which, after being uncovered, endow it with an internal consistency and coherence; 2) to the concomitant unearthing of the investigator's presumptions and prejudices which prevented her from understanding this text or behaviour, and recognizing that consistency and coherence before. Together these two open up a *new*, legitimate, way of looking at

things which had been previously closed, thus pinpointing some of the *significant differences* between the framework of the investigator and that of the text or behaviour under study.[5]

d. Finally, it is essential to underscore the fact that the hidden 'premises' and 'lemmas' thus unearthed are, and can be, formulated *only* in terms of the investigator's own framework and against the background formed by the bedrock of this framework, i.e. as they appear *from the perspective defined by it*. It is essential to emphasize this because the *internal* distance within the investigator's framework, created by the new way of looking at things, may cause the investigator to 'forget' that what she has uncovered is *necessarily* expressed in the terms of her own framework. This internal distance may create the illusion that she has relativized this framework itself, that is, that the 'premises' and 'lemmas' she has unearthed are on a par with those *constitutive* of her own framework, those that have allowed her to understand her object of study in the first place. It is, I believe, at this precise point that the roots of all relativism can be located.[6]

As specifically concerns the discipline of history of science, the crux of the preceding argument is the following:

1 Points 6 and 7 above imply that for the understanding of a past framework to be at all possible, the 'premises' and 'lemmas' of that framework should be *expressible* (translatable) in terms of the current framework. Otherwise the object of study of the historian remains totally meaningless.

2. If this happens to be the case,[7] then the process of understanding a past framework amounts to opening *a new way of access* to the part or aspect of the world which the object of study of the historian talks about *in the interior of the current framework*. This amounts to the opening of a new *logical* (or rather, grammatical, in Wittgenstein's sense) *possibility* in that interior (Point 8c).[8]

3. For the 'premises' and 'lemmas' of the old framework to be expressible in terms of the current, the current framework should, somehow, have the terms available that render this possible. And for a new logical possibility to appear in the interior of the current framework, this framework should be 'wide' enough or 'rich' enough to hold and

accommodate it. It is only under this condition that history of science (always considered different from "a repository of anecdote and chronology" and as an enterprise that aspires to treat past theories in their own terms) *is itself possible.*

To see how these points apply specifically to historiography of science and its methodology, we have to take into account the fact that an important part of any cultural framework is made up of the fundamental ingredients of its corresponding knowledge-producing practices (basic 'premises' and 'lemmas', procedural norms, the 'hard core' – in Lakatos's sense – of the relevant 'products', etc.). Conversely, these knowledge-producing practices crucially involve the bedrock of the cultural framework in question. In short, any cultural framework includes what we could anachronistically call the fundamentals of the 'science of the day' while any 'science of the day' is constitutively dependent upon the corresponding framework.

Given this, the third of the above points leads us to an inescapable dilemma: Either all cultural frameworks are 'wide' enough to accommodate any other framework, or some are inherently 'wider' than others. I don't presume to possess an answer at this general level. But if we set aside cultural frameworks in general and concentrate on the parts of them involved in science – with the current conception of the term, say, as the term applies from Galileo onwards – then, I maintain, we can choose between the dilemma's two horns.

On the one hand, the first horn amounts to an unqualified relativism and forces us to maintain for example, that the framework of Aristotelian physics is 'wide' enough to accommodate Einstein's theory of relativity. Clearly this is untenable. The second horn, on the other hand, may not lead to relativism but it implies that current science is inherently 'wider' of 'richer' than past science. And this conclusion, if it is supplied with the appropriate moral dimension, is not very far from quintessential whiggism. Thus, if we do not espouse relativism, it looks as if we are not only forced to turn whiggist but, in addition, that we have landed on a paradox: The whole argument begins with the constitutive anti-whiggist requirement to assess past science in its own terms and concludes with quasi-quintessential whiggism!

I don't believe that there is any real paradox involved here. (And what precedes is intended to constitute the relevant argument.) Rather, we have finally arrived at the destination we set ourselves and we meet at this point the inherent limits of unqualified anti-whiggism. Accordingly, I don't hesitate to espouse the dilemma's second horn. Of course, it goes without saying that to show that this choice is not arbitrary needs justification. However, an adequate justification would require a full-blown conception of the structure of science and of the modes and forms of development that this structure allows. Obviously, this cannot be undertaken here (see, however, Baltas 1987 and 1988). What I can do to somewhat justify my choice within the space and time allotted is to focus on physics, the science *par excellence.*

Philosophically unsophisticated physicists take for granted that a new theory of a certain domain of phenomena cannot be accepted unless it somehow accommodates the successes of the old theory in respect to the same domain. It is mainly for this reason that practically all physicists take for granted that their science progresses. Now, this intuition can be made philosophically acceptable without jettisoning what philosophy of science has recently taught us, if we adopt (along with a conception of science that spells it out and justifies it) the following thesis: A radical change in science takes place through the uncovering of what we can call 'ideological assumptions' – which are not very different from what we called 'premises' and 'lemmas' above (Baltas 1987). This is to say that the physical theory, coming after the radical change, frees itself from part of what its predecessor took for granted, left unquestioned as a matter of course, and unwittingly based its own development on.

If this is indeed the case, then, *pace* Kuhn's initial presentation of the corresponding gestalt switch, the relation between the two successive theories, is *not symmetrical.* This implies that we can have 'local' (Kuhn 1983) incommensurability and a 'local' communication breakdown without endangering rationality in any way and salvaging, in an important sense, the idea of progress in science (Baltas 1992). This happens because the theory, coming after the radical change, not only is unconstrained by the disclosed 'ideological assumption' but also organizes itself around this disclosure in the sense that it strives to incorporate (after the appropriate translation) all, or most, of what the previous theory had achieved despite its being blindly constrained by that 'assumption'. For this reason, the two theories, although locally incommensurable, are locally comparable. This allows us to say without contradicting the incommensurability thesis that the new theory is inherently 'wider' of 'richer' than the old (always only locally, for the new theory inevitably harbours its own 'ideological assumptions' which only a new radical theoretical change can bring into light) and hence that progress, in an absolute (although only local) sense, has been accomplished. This additional 'width' can be visualized as the 'measure' of the historical 'distance' separating the framework of the two theories. To give just one example, the framework of Einstein's theory of relativity is rich enough to possess a 'low velocity' limit which can be considered as an (imperfect) translation of Newtonian mechanics, while the framework of the theory of Newton cannot accommodate any (adequate translation of) Einsteinian relativity. This asymmetry makes Einstein's theory inherently 'richer' than that of Newton and the transition between the two objectively progressive.

If what precedes is defensible, then the following two conclusions follow: First, science constitutes the *particular* and perhaps unique intellectual endeavour which progresses in some absolute sense (although, as we said, always only locally). Second, the framework of current science is the widest framework that we can currently 'possess'[9] for it is the outcome of all such past disclosures. It is for this precise reason that we can come to understand past scientific theories, which is to say, that it is this very fact that renders historiography of science itself possible. Irrespective of whether historians of science are conscious of it or not, current science thus indeed acts as the norm that mutely regiments the overall study of past science, for its framework forms the horizon of all possible questions of the past and all possible translations of past theories. Risking total misunderstanding, I would even go so far as to say, following Bachelard and Foucault, that historiography of science is always and only the study of how *current* science came about and of what *current* science allows to be studied.

## CONCLUSION

The "harmful effects" of unqualified, or "excessive", anti-whiggism are now easy to locate. They boil down to the harmful effects that a misconstrual of the methodology that is actually involved in historiography of science may have on the effective exercise of the discipline. Historians of science are infinitely better qualified than myself to pinpoint and eradicate these harmful effects and, thus, I gladly leave the floor to them. However, let me add one more thing. In reading the work of most historians of science, I came to realize that they tend to resist with all their might even considering the possibility that an empirically verifiable *theory* of their discipline can be eventually constructed. Perhaps this resistance itself can be ascribed to the fear of being accused of whiggism which has been created among historians and, thus, it may constitute yet another harmful effect of excessive anti-whiggism. However, be this as it may, I do not agree. I sincerely believe that philosophers, historians and sociologists of science can join in a collaborative enterprise which could eventually lead to the formulation of an empirically adequate theory of the history of a science such as physics. But I must confess, that for reasons which are too lengthy to discuss here, I am not overly optimistic. Independently of this, however, the present conference is on the contemporary trends in historiography of science. Perhaps some day – who knows? – the theoretical trend I am suggesting may indeed become a contemporary trend.

*National Technical University, Athens, Greece*

## NOTES

[1] After the oral presentation of the paper, I had the privilege of a long session of questions and answers. All those who then addressed their queries to me, or expressed their disagreement, helped me, perhaps more than they could realize at the time, in shaping this final version. I thank them all wholeheartedly. In addition, before the Conference, I had the pleasure of sustained conversations on the paper itself as well as on the conception it is based on with Jean Paul van Bendegem, Peter Machamer, Marcello Pera and Wal Suchting, who happen to be not only colleagues but also close friends. All four know very well my debts to them, which extend far beyond their more than substantial help with the paper itself.

[2] The quotation marks intend to convey that the 'premises' and 'lemmas' in question, being largely unconscious, do not enjoy fully the status of their namesakes in explicitly formulated arguments. On the other hand, the distinction between the two is more or less the standard one: 'Premises' are the fundamental foundational blocks of the corresponding 'arguments' (the quotes are placed here for the same reason); 'lemmas' are lateral 'premises' to which the 'argument' appeals in order to be helped out.

[3] Here, as elsewhere in the paper, I am very much indebted to Wittgenstein's ideas, and especially to his *On Certainty*.

[4] After the oral presentation of the paper, Dr. S. Strickland suggested that I change the metaphor and talk instead of getting out of "one's clothes." For reasons that have already started to appear, I cannot agree with this. The cultural framework in which we are embedded is inescapable. This makes it impossible for us to ever get stark naked, to then choose the suit that is more becoming. However, I can compromise: The bedrock of every cultural framework is multi-layered. And this means that, under special conditions, we can – or rather be forced to – change some of our clothes. But, again, in most cases this is not a matter of choice or decision.

[5] See here the exemplary analysis that Kuhn gives in his 1990 on how he came to understand the framework of Aristotle's physics.

[6] According to the present conception, relativism, considered as the thesis that upholds the equivalence of all frameworks, is untenable for two reasons. First, it ignores the impossibility of radically escaping the given cultural framework. Second, at least if it is formulated in such an unqualified manner, it does not pass the well-known self-reference test: It is a thesis which presents itself as being a part of a panoptic viewpoint wherefrom all possible frameworks can be surveyed. Thus it is ultimately self-defeating.

[7] If we consider cultural frameworks in their generality, it is a big question whether this is possible and under what conditions. For example, could the cultural framework of the Incas accommodate all the 'premises' and 'lemmas' of our culture today, for example, those which allow us to distinguish between science, religion and magic?

[8] John Earman, in trying to formulate the various space-time theories that have appeared in the history of physics by using the language of current mathematics notes that: "Incommensurabilities have a way of disappearing when the *initially* seeming incommesurabale set of propositions is fitted into an *appropriately enlarged possibility set*." (1989, p. 27, my emphasis.) Earman's whole approach testifies that "no fear of being labeled Whigs" should prevent us from taking advantage of an "apparatus that can...provide [such a] larger possibility set." (*ibid.*) As will appear more clearly below, the whole point of this paper is to suggest that the kind of approach advocated by Earman is not only convenient, but also, in a sense, accurately displays at least some aspects of the *effective* development of the history of a science, such as physics.

[9] In fact, 'possess' is the wrong verb here for we do not control this framework in any real sense. We are inescapably (up to the next radical scientific change) caught in it and constrained by it in such a way that it would be more appropriate to say that it is this framework which controls and possesses us.

## REFERENCES

Baltas, A. (1987), "Ideological 'Assumptions' in Physics: Social Determinations of Internal Structures", in A. Fine and P. Machamer (eds.), *PSA 1986*, Vol. 2. East Lansing, Michigan: Philosophy of Science Association.

Baltas, A. (1988), "The Structure of Physics as a Science", in D. Batens and J. P. van Bendegem (eds.), *Theory and Experiment*. Dordrecht, Holland: D. Reidel.

Baltas, A. (1992), "Shifts in Scientific Rationality and the Role of Ideology", in M. Assimatopoulos, K. Gavroglu, and P. Nikolakopoulos (eds.), *Historical Types of Rationality*, Proceedings of the First Greek-Soviet Symposium on Science and Society, National Technical University of Athens, 1992.

Earman, J. (1989), *World Enough and Space-Time*. Cambridge, Mass.: The MIT Press.

Kuhn, T. S. (1983). "Commensurability, Comparability, Communicability", in P. Asquith and T. Nickles (eds.), *PSA 1982*. East Lansing, Michigan: Philosophy of Science Association.

Kuhn, T. S. (1990), "What Are Scientific Revolutions?", in L. Kruger, L. J. Daston, and M. Heidelberger (eds.), *The Probabilistic Revolution*. Vol. 1. Cambridge, Mass.: The MIT Press.

Wittgenstein, L. (1972), *On Certainty*. New York, N.Y: Harper and Row Publishers.

DIMITRIOS DIALETIS AND EFTHYMIOS NICOLAIDIS

# ISSUES IN THE HISTORIOGRAPHY OF POST-BYZANTINE SCIENCE

In this paper, we shall present the expressions and trends of the historiography of science in Modern Greece meaning the Greek world after the fall of the Byzantine Empire. There will be two approaches to this subject. One will examine the different kinds of history of science over all these years and the other will examine their relations with respect to the periods of Greek intellectual life.

What we now call "History of science," was present under different forms in the last five centuries. There have been many approaches and uses, depending on the aims of each historiographer from the confusion of history of science and scientific education in the first post-Byzantine centuries, to the independent discipline of today.

The first example of what we can call "History of science" was scientific education. During the 16th and the 17th centuries, some manuscript presented an historical approach to scientific knowledge. For example, astronomy was sometimes taught as the history of ancient Greek astronomy, and the teaching of mathematics meant the teaching of Euclides, Apollonius and other Greek mathematicians. This teaching policy was in fact a critical presentation of ancient Greek science.

Throughout the 18th century, this critical presentation was the consequence of the educational policy of Modern Greek Humanism. This philosophical trend, born in the beginning of the 17th century and expressed by the Greek Orthodox church, consists of the idea of a revival of the ancient Greek spirit in Modern Greece, and there fore of the revival of ancient Greek science in its homeland.

In that spirit we must also consider a purer type of the history of science during the 18th century, that of the historical overviews in prolugues of many scientific manuals of those times. These prologues presented only the history of ancient Greek science, especially the history of mathematics. The splendour and the leading role of ancient Greek science was accentuated, as well as the debt of West European science to Greece.

After the middle of the 18th century, the history of science as examined by these prologues, was extended to the evolution of science after the end of the Ancient World. This trend was greatly accentuated during the



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 121-127.

Modern Greek Enlightenment, that is, after 1780 (Bechrakis and Nicolaidis, 1990). Contrary to the educational policy of Modern Greek Humanism which consisted of the revival of ancient Greek science, the educational policy of the Modern Greek Enlightenment consisted of the popularization of West European science as a means toward the formation of a national conscience different from the conscience which tied Greeks to an Empire, although the fact that this Empire was now under the control of the Ottomans (Apostolopoulos, 1989; Dimaras, 1975, 1977; Kondylis, 1988; Nicolaidis and Dialetis, 1992).

If we had to point out the main expression of the history of science in post-Byzantine Greece, it would be the history of philosophy. For Modern Greek philology, the history of science was part of the history of philosophy for one of two reasons: either the history of science was considered a branch of the "naturfilosophie", or the pure philosophic work of Greek scholars prevailed on their scientific works (Sathas, 1868; Dimaras, 1975; Papanoutsos, 1953; Voumvlinopoulos, 1966).

Now we know the disadvantage of such an approach. The philosophical work of the different Greek authors is not of equal quality or importance as that of their scientific work. Only in very few cases can we say that the work of a Greek scholar *in philosophy* had the same importance on the evolution of the Modern Greek culture than his work in science. More there are many cases, especially during and after the 18th century, where the philosophical and scientific work of some scholars present such different characteristics (Kondylis 1988), that we could speak of cultural schizophrenia for these scholars. We could mention the case of the archbishop Nikiforos Theotokis, a great scholar of the 18th century, who philosophically wrote in the spirit of the Orthodox tradition, while scientifically, he was one of the first to try to synthesize the ancient Greek science with West European knowledge (Theotokis, 1766). This approach led to the elevation of the philosophic work of Greek scholars as the main criterion for the estimation of the culture of whole epochs in Modern Greek history. An example is how historians view the entire Greek 18th century, elevating the Enlightenment as the main cultural event of Modern Greek history. Without an autonomous study of the history of science, we cannot understand the limits and contradictions of the Modern Greek Enlightenment and, also, the importance during this century of Modern Greek Humanism.

It is obvious that these philosophical approaches were made by scholars who had a deep knowledge in that field – we can mention, for example

Papanoutsos or Tatakis, two major scholars of the 20th century – but *ignored all about exact science.* One characteristic example of the influence of these scholars is that until recently, the Greek scholar community appreciated the already mentioned Nikiforos Theotokis by his philosophical work titled "Metaphysics".

After 1922, the year of the defeat of the Greek army in the campaign of Asia Minor and the formation of the Greek socialist current, the history of science also becomes a part of sociology (we can mention, for example Kordatos, 1957). This new approach, interesting by itself, presents the same disadvantages as the previous one. Now, science was viewed only as an educational policy or as closely related to social events. Under this approach, educational manuals of the 18th century that were written by authors without any scientific knowledge, were considered extremely important texts for the Modern Greek history of science. An example is a manual on Physics by Rigas Feraios (Feraios, 1790), a book without any serious scientific educational value, written by a revolutionary without specific knowledge in that field.

In the 18th and 19th century we also had another, indirect type of history of science. It consisted of the implication of the history of science on the main cultural problem of that time, the cultural identity of the Greek nation. This was the problem of the independence of the Orthodox Church from West European influences. In the late 18th century, a major anti-West European current again formed in the Orthodox Church, that of Kolyvades (Apostolopoulos, 1989; Dialetis, 1992). As a main vehicle of European culture, contemporary science was distrusted, and polemics rose about the role and utility of science. These polemics also involved the history of science, used at this time for the arguments of each party.

Another indirect approach to the history of science is the history of the Education made by Greek historians of the 19th and the 20th centuries (Gritsopoulos, 1966, 1971; Evagelidis 1936). As, after the fall of the Byzantine Empire, the original Greek scientific production was almost non-existent, the history of science of that period consists mainly of the history of the transfer of science – transfer either from ancient Greece or Byzantium, during the 15th, 16th, 17th and 18th centuries, or from West Europe during the 18th, 19th and 20th centuries. Education having been the main tool of this transfer, the history of the Greek Colleges has been confounded with the history of science. We must note here that Greek Colleges held the control of science during all centuries after the fall of Byzantium until the beginning of the 19th century and their history contributes mainly to the history of Modern Greek science.

The Greek Colleges began organizing in the 15th century and spread all over the Greek world, even during the Diaspora, in Italy, Central Europe and Russia. Therefore, their role in the transfer of scientific knowledge is obvious. Another fact is that as these Colleges were independent, their educational science policies were varied. Sometimes, even in the same college, many scientific trends existed, as in the case of the most important, the College of the Patriarchate, which was obliged *by the competition* to change its educational policy toward science after the impact of the ideas of the French Revolution (Apostolopoulos 1989, Dialetis 1992, Karas 1977).

The main problem of the confusion of the history of science with the history of the Colleges is similar to those of the other indirect approaches. The historians who tried this approach had neither enough information about scientific education in the Colleges nor the adequate scientific knowledge to evaluate this education.

The fact that history of science was either part of the history of philosophy, or sociology, or history of education created an anti-history of science position in the Greek intellectual milieu. It must be clear that the debate was not on the level of the internal or external approach, but on a previous level, this of the existence or not of an history of science. In the last two decades we have begun to see some samples of "pure" – if we can say so – history of science. The main characteristic of this late trend is that it originates from the actual milieu of science, a milieu which can understand and classify the historical matter by criteria derived by the same subjects they are studying. The fields of interest in this new milieu are varied, from science in ancient or Modern Greece to the evolution of science in Europe.

Now we shall try another historiographic approach, the relation of the different kinds of history of science with respect to the periods of post-Byzantine intellectual life.

The first great historical period, the rule of the Ottoman Empire from 1453 until 1821, presents a major interest for the historiographer because of the main intellectual currents, those of the Orthodox tradition, the Modern Greek Humanism and the pro-Occidental current which was related to the Modern Greek Enlightenment.

The Orthodox tradition, related to the Oriental mysticism of the 14th century, is expressed by an "anti-science" position. For this way of thinking, science corresponds to a spiritual exercise and man cannot approach nature and the world by scientific knowledge alone, but, more directly, through their relation with creation. On the other hand, contrary to Catholicism, the Orthodox Church is liberal concerning scientific theories (Gedeon, 1888). Note that the works of Saint Grigorios Palamas, a Byzantine Ecclesiastic of the 14th century, expresses the relations of Orthodoxy and science. This position toward science influences the history of science. An "anti-scientist" history denies the idea of the scientific evolution of the world.

The relation of science to evolution was made by Modern Greek Humanism, and that is why the history of science itself made its appearance in the prologues of the scientific books written by the followers of that intellectual current. Modern Greek Humanism is a late edition of Byzantine humanism, which appeared in its first form in the 11th century. For the humanism of the 17th and 18th centuries, Orthodoxy is not in contradiction with the ancient Greek spirit but closely related to it. A picture can be seen today in a monastery of Mount Athos that expresses that spirit, where Plato and other ancient Greek philosophers are represented as saints.

History of science could serve perfectly this way of thinking if viewed from a specific point. The Modern Greek Humanists, the scholars who introduced science and especially mathematics in Modern Greece, almost always spent some pages in their scientific books on the history of ancient Greek science, insisting on the point that all sciences were born in Greece and that the major evolutions in the history of science took place in the ancient Greek world. An example of that spirit is the repetition by many authors of the 18th century that algebra was born in ancient Greece and not in the Arab world, or that no European mathematician could present in geometry so perfect a system as the Euclidean (Nicolaidis, 1989).

The other major intellectual current of post-Byzantine Greece, the Modern Greek Enlightenment, used the history of science for a totally different aim. The educational policy of that current was to present contemporary science, meaning West European science, as a key to the understanding of the world. This position was better served by a philosophical approach and a presentation of the achievements of contemporary Western science. When the followers of this intellectual current dealt with the history of science, they presented mainly the history of West European science, viewed as the history of the evolution of the humankind.

After 1821 and the birth of the independent Greek State, a completely new period for Greek intellectual life began. This was the period of the development of philosophy against science, which nearly disappeared from education, especially physics.

Characterized by the incorporation of the history of science into the history of philosophy (Papanoutsos, 1953), this period will mark the historiography of science until today. The few pages on the history of science of those years concern the history of mathematics, the only discipline which preserved a certain vitality.

After 1922 and the advent, of socialist ideas in Greece, the history of science and many other disciplines were incorporated into sociology by the Greek Marxists (Kordatos, 1957). The two major schools of the history of science, the history as part of philosophy and as part of sociology, coexisted and represented two different intellectual currents. But this "political" approach began to fade after 1950 with the advent of what is called the "new generation" of Greek historians, well represented by K. Dimaras (1975, 1977). Even then, history of science was not an independent discipline and was confounded by historians with philosophy, sociology, or the history of education. Even now the current image of post-Byzantine Greek science in the Greek intellectual circles is deeply influenced by that school, which presents the Greek Enlightenment as the major event of the history of Modern Greek science.

We must mention that marginally to those major trends in the historiography of science there existed some exceptions of "pure" historians, who contributed a lot in their respective fields of interest. But those scholars, e.g. E. Stamatis (Stamatis, 1975) who worked on ancient Greek mathematics and M. Stefanidis (Stefanidis, 1926; 1938), did not succeed in creating a community and changing the current trends of the historiography of science.

The actual milieu of the history of science in Greece is a result of post-1974 Greek intellectual life. Greek historians of science now come from the milieu of exact sciences and their fields of research concern mainly the history of ancient Greek science, the history of science in post-Byzantine Greece and, also, the history of science in Europe after the 18th century. This community has a certain variety and is incorporated, with the evident specificity that mainly concerns the history of science of Modern Greece, in the international community.

*National Observatory of Athens, Greece*
*National Hellenic Research Foundation, Athens, Greece*

## REFERENCES

Apostolopoulos, D., 1989, I galliki epanastasi stin tourkokratoumeni elliniki kinonia (The influence of the French Revolution in the Greek society during the Ottoman rule), Athens.

Bechrakis, Th. E. and Nicolaidis, E., 1990, Statistiki analysi lexikon dedomenon (Statistical analysis of lexical data), Athens, Edition EKKE/EIE

Dialetis, D., 1992, Reason and Revelation in the 18th century Greek astronomy, Proceedings of the Greek–Soviet symposium "Science and Society", Athens.

Dimaras, K. Th., 1975, Istoria tis neoellinikis logotechnias (History of the Neohellenic literature), 6th edition, Athens.

Dimaras, K. Th., 1977, Neoellinikos Diafotismos (The Neohellenic enlightenment), 5th edition (1989), Athens.

Evagelidis, T., 1936, I pedia epi Tourkokratias (The education during the Ottoman rule), Athens.

Feraios, R., 1790, Physikis apanthisma ... (elements of Physics), Wien.

Gedeon, M., 1888, I pneumatiki kinisis tou genous kata ton 17o kai 18o aiona (The intellectual activity of the Greek nation during the 17th and 18th centuries), Collection of texts of the period 1888 - 1889 edited by A. Agelou and F. lliou, 1976, Athens.

Gritsopoulos, T., 1966, Patriarchiki Megali tou genous sxoli, Tome A' (1966), Tome B' (1971) (The great patriarchate's school of the Greek nation), Athens.

Karas, G., 1977, I fysikes kai thetikes epistimes ston elliniko 18o aiona (The natural science in the greek 18th century), Athens.

Kondylis, P., 1988, O Neoellinikos diafotismos, I philosophikes idees (The Neohellenic enlightenment, the philosophical ideas), Athens.

Kordatos, G., 1957 - 1958, Istoria tis neoteris elladas (The history of Modern Greece), Tomes I-Iv, Athens.

Nicolaidis, E., 1989, I anagenissi ton archaion mathimaticon stin Ellada tou 18ou aiona (The renaissance of ancient greek mathematics in Greece during the 18th century) in Ancient Greek Mathematics (ed. Anapolitanos - Karasmanis), Athens.

Nicolaidis, E. and Dialetis, D., 1993, L'influence des lumières sur la formation scientifique greque, *Revue d'Histore des Sciences* XLV, 4.

Papanoutsos, E., 1953, Neoelliniki Philosophia (Neohellenic Philosophy), Tomes I, II, Athens.

Sathas, K., 1868, Neoelliniki filologia (Neohellenic Philology), Athens.

Stamatis, E., 1975, Istoria ton ellinikon mathimaticon (History of Greek Mathematics), Athens.

Stefanidis, M., 1938, Isagogi is tin istorian ton thetikon epistimon (Introduction in the history of natural science), Athens.

Stefanidis, M., 1926, Ai fysikai epistimai en elladi pro tis epanastaseos (Natural science in Greece before the revolution), Athens.

Theotokis, N., 1766 - 1767, Stoixeia Physikis ... (Elements of Physics), Breitkopf, Leipzig.

Voumvlinopoulos, G. E., 1966, Bibliographie critique de la philosophie Grecque, Athens.

VYACHESLAV S. STEPIN

# SOCIAL ENVIRONMENT, FOUNDATIONS OF SCIENCE, AND THE POSSIBLE HISTORIES OF SCIENCE

Relatively stable bases may be outlined in the system of our scientific knowledge of physics as a branch of science. These bases determine the strategy of scientific investigation and generalize the experimental results. The most important components of such bases are: (1) ideals and norms of science; and (2) the pattern of the reality under investigation.

The following basic kinds of ideals and norms of science may be singled out: (a) explanations and descriptions; (b) probability and substantiation of knowledge; and (c) arrangement and systematization of knowledge. There are several hierarchically related levels in each type. The first level is represented by normative structures that are invariant for science in any epoch. They define the specificity of scientific cognition, its difference from other forms of cognitive activities (artistic endeavors, every cognition, religious and mythological understanding of the world, etc.). This level is concretized by means of historically transient objectives in cognition (ideas of the norms of explanation, description, probability, arrangement of knowledge, etc.) which specify the style of *thinking* in some epoch of the development of science and form the second level of the ideals and norms of scientific investigation. For example, the ideals and norms of explanation adopted in the science of the Middle Ages radically differ from those which characterize the science of our day; the norms of elucidation and substantiation of knowledge in the epoch of classical natural science differ from modern ones. The third level of the ideals and norms of scientific investigation solidifies the objectives of the second level with respect to the specific features of each branch of science (physics, chemistry, biology, etc.). This layer of normative structure is perceived in the form of the methodological principles of some or other subject (in physics, for example, they are the principles of correspondence, observability, symmetry, etc.).

The system of ideals and norms of science forms a generalized and dually determined scheme of the cognition method: it is determined on the one hand by socio-cultural factors and, on the other, by the kind of objects under investigation. The transformation of ideals and norms changes the



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science,* 129-137.

scheme of the method, thereby making it possible to recognize new types of objects.

The essential and salient features of scientific subjects are expressed by the existing reality. The best studied pattern of such reality is the physical picture of the world. It gives notions of: (1) fundamental objects that underlie all other physical objects; (2) the typology of the objects under study; (3) general regularities of their interaction (regularities and causality of physical processes); and (4) space-time characteristics of the physical world. The concrete forms of all these notions change along with the development of cognition and practice. For example, the mechanical picture of the world was transformed in the last quarter of the 19th century into an electrodynamic one; the latter was superseded in the 20th century by a quantum-relativistic picture of physical reality.

The reality pattern is a unique form of theoretical knowledge that is correlated between concrete scientific theories and empiric facts. It is made explicit by means of a system of ontological postulates (for example, the postulates of a mechanical picture of the world: bodies consist of atoms, bodies interact through instantaneous transmission of forces, etc.). One and the same pattern of reality may form the basis of a multitude of theories, including fundamental theories.

The picture of the investigated reality in each branch of science is always taking shape not only within science itself, but also through interrelation with other fields of culture.

In the course of the formation and development of the special pictures of the world, science makes extensive use of images, analogies and associations rooted in the practical activity of mankind (the images of the corpuscle, the wave, the continuum, the correlation of "part" and "whole" as visual notions of the systemic organisation of objects, etc.). This layer of visual images is incorporated in the picture of reality being investigated and makes it, in many respects, an understandable and natural system of the notions of nature.[1]

Reconstruction of research premises implies a change in the very strategy of scientific inquiry. But any new strategy does not assert itself immediately, but in the prolonged struggle with former sets and traditional visions of reality.

The process of assertion of new premises in science is determined not only by forecasting new facts and generating concrete theoretical patterns, but also by socio-cultural causes. New cognitive sets and the knowledge they generate must be fitted into the culture of a given historical period, consonant with the values and ideological structures which underlie it.

In this respect, reconstruction of scientific premises in a period of scientific revolution represents a choice of particular guidelines in expanding knowledge which should provide both the extension of the scope of research and a certain correlation of the dynamics of knowledge within the values and ideological sets of a given historical period.[2] in the context of such a revolution, knowledge could be expanded in several ways, however, not all of them are realised in the actual history of science. Two aspects can be distinguish in the nonlinear growth of scientific knowledge.

The first aspect involves the competition of research programmes within the framework of a specific branch of science.[3] A case in point are the research programmes characterised by particular research premises. Victory of one programme and degradation of the other orient the development of the branch of science along a definite route, at the same time barring any other routes of possible progress.

Take, for example, the struggle of two trends in classical electrodynamics en Ampère–Weber, on the one hand and Faraday–Maxwell, on the other. Maxwell, while evolving a theory of electromagnetic fields, did not obtain new results for a long time when compared with those provided by the electrodynamics of Ampère–Weber. On the surface everything looked like deduction of the known laws in a new mathematical form. Eventually, by arriving at the fundamental equations of electromagnetism, Maxwell obtained the famous wave solutions and forecasted the existence of electromagnetic waves. Their experimental verification led to the triumph of Maxwell's trend and asserted the notions of short-range interaction and fields of forces as the sole, true basis of the physical picture of the world.

However, in principle, the effects that were interpreted as the proof of electromagnetic waves could have been predicted within Ampère's trend as well. It is known that in a letter to Weber in 1845 Karl Gauss remarked that to further develop Ampère–Weber's theory he should admit, in addition to the known intercharge forces, the existence of other forces propagating with terminal velocity.[2] G. Riemann carried out this programme and derived an equation for potential, analogous to Lorentz's equations for retarded potentials. In principle, this equation could underlie the prediction of the effects which were interpreted in the paradigm of Maxwell's electrodynamics of electrodynamics assumed a physical picture of the world wherein forces with various velocity are propagated in hollow space. This picture of the world lacks ether and electromagnetic fields, giving rise to the questions: What would the electron theory have looked like in that non-realised physics trend and what would have been the road to the theory of relativity?

A physical picture of the world depicting the intercharge action as a transfer of forces with terminal velocity, with no accounting for material fields, is quite possible. It is worth noting that R. Feinman proceeded from this image of electromagnetic interactions to provide a new formulation of classical electrodynamics and, later, to develop quantum electrodynamics in the terms of trajectory integrals.[5] Feinman's reformulation of classical electrodynamics could be regarded, to a certain extent, as a modern reproduction of a non-realised potentiality in the historical development of physics. But we should bear in mind that the current concepts of nature are formed in conformity with another scientific tradition, rather than of the classical period following the new ideals and standards of explanation of physical processes.

While asserting these standard, the progress of quantum-relativistic physics "schooled" physicists to consider multiple the progress of quantum-relativistic physics "schooled" physicists to consider multiple formulations of theory, each of them being capable of expressing the essentials of the investigated subject. A 20th-century theoretical physicist regards various mathematical descriptions of the same processes not as an abnormality, but as something quite normal. He is fully aware that the same objects could be tackled by various linguistic means and the differing formulations of one and the same physical theory is a requisite of research progress.

Modern physics also traditionally appraises the picture of the world as a relatively true system of notions of the physical world which could undergo changes both in part and in its entirety. That is why, when Feinman developed the ideas of intercharge action without "field mediators" he was not troubled by such considerations as the necessity to introduce in the evolving theory, in addition to retarded advanced potentials which led to the emergence, in the physical picture of the world, of notions about the impact of current interactions on the future and even on the past. R. Feinman wrote, for example, that by then he was a physicist enough to realise – as did all physicists since Einstein and Bohr – that sometimes an idea seemingly paradoxical at first glance could turn out to be valid upon considering all the minute details and thereby its verification in experiment is discovered.[6]

But to be a "20th century physicist" is something else than being a 19th century physicist". During the preclassical period a physicist would not introduce "extravagant" ideas of the physical world on the assumption that he had arrived at a new and promising mathematical form of theory which could be verified by future empirical details. In the classical period, before generating new theoretical ideas the physical world had to provide a "visual picture" of reality, verified by experiment. Formation of competing pictures of the investigated reality presupposed their tough confrontation under which each picture was regarded by its adherents as the solely authentic ontology.

We should assess the potential realisation of the Gauss–Riemann programme in 19th century physics under this angle. To introduce, in the physical picture of the world of that period, the concept of forces propagating with different velocities one had to substantiate this concept as a visual image of "the real structure of nature." Physicists of that period traditionally associated force with material carriers. Therefore, its changes in time from point to point (differing velocities of force propagation) assumed the introduction of material substance, the state of which caused the change in propagation velocity. But such concepts were already delineated in the Faraday–Maxwell programme and incompatible with the Ampère–Weber programme. (In this picture the connection between force and matter was treated as an interrelationship of electrical and gravitation forces. On the one hand, a charge-mass relation existed and on the other, charges and masses acted as material carriers of forces, whereas the principle of momentary transfer of forces in space excluded the necessity of introducing a special substance to transfer force from point to point.)

So, the failure of the Gauss–Riemann idea in leaving a substantial imprint on the history of 19th century classical electrodynamics was caused by the style of physical thinking in that period. This style of thinking, with its intention to provide absolutely authentic notions about the essence of the physical world, was a manifestation of the "classical" type of rationality which found its realisation in philosophy, science, and other spiritual phenomena of that time. This rationality contemplates an object from the outside, mastering its true nature in this way.[5]

But the modern style of physical thinking (characterising the above approach with the non-realised but possible line of development of classical electrodynamics) manifests itself in another, non-classical type of rationality governed by a particular attitude of thinking about an object and itself. Here, thinking reproduces an object as interwoven with human activity and builds up the images of an object, correlating them with the notions of the historically established means of its cognition. Thinking gropes for further progress and becomes aware, to a certain degree, of its own being as an aspect of social development and, therefore, of its determination by this development. Under this type of rationality the once-obtained images of existents are not considered as the solely possible ones (in the other linguistic system, in other cognitive situations the image of an object can be visualised in another way, with all these varied concepts conveying the objective truth).

The formation of the modern type of rationality is conditioned by the historical development of society, by the change in "the field of social mechanics" which "supplies things to consciousness'.[8] Study of these processes is another matter. But in general, we could state that the type of scientific thinking taking shape in the culture of a certain historical period always correlates with the character of human communication and activity of the given period, being predetermined by its cultural context. Factors of social determination of cognition affect the rivalry of research programmes, accelerating some routes of discovery and braking others. As a result of the "selective work" of these factors within the framework of each scientific discipline, only some potential routes of scientific development are realised while the rest remain non-realised.

The second aspect of the nonlinear growth of scientific knowledge is related to the interaction of scientific disciplines, preconditioned by specifics of both investigated objects and the socio-cultural environment of scientific development.

The emergence of new branches of knowledge, the replacement of leading sciences, and revolutions arising from the transformation of the pictures of the investigated reality and standards of research activity in separate branches may exert a noticeable effect upon other fields of knowledge, changing their vision of reality, their ideals, and standards of research. All these processes of scientific interaction are mediated by and exert an intensive feedback impact on various cultural phenomena.

With all these involved mediations in view, we outline one more type of potential route in the history of each science which represents a specific aspect of the nonlinearity of scientific progress. This aspect could be illustrated by an analysis of the history of quantum mechanics.

As is known, a key moment in the construction of quantum mechanics was Bohr's new methodological idea, according to which, the concepts of the physical world should be introduced through explication of an operational scheme specifying the investigated objects. In quantum physics this scheme is based on the complementarity principle, in conformity with which the nature of a micro-object is described by two complementary characteristics, correlative to two types of devices. This "operational scheme" was combined with a number of ontological concepts, for instance, the corpuscle-wavelike nature of micro-objects, the existing quantum of action, and the objective interrelationship of dynamic and static laws of physical processes.

However, the quantum picture of the physical world did not present an integral ontology in traditional terms, for it failed to describe natural processes as causal interactions of some objects in space and time. Spatial-temporal and causal descriptions acted as complementary (according to Bohr) behavioural characteristics of micro-objects.

The two types of description were applied to a micro-object only through the explication of an operational scheme which combined various and outwardly incompatible fragments of ontological concepts. This method of constructing a physical picture of the world was philosophically substantiated, on the one hand, by a number of epistemological ideas (about the special place in the world of an observer as a macro being, about the correlation between the methods of explanation and description of an object and cognitive means) and, on the other hand, by the development of a "categorial network" embracing the general specifics of the subject of study (the notion of interactions as transformation of possibility into reality, the approach to causality in a broader context including probability aspects, etc.).

In this way the conceptual interpretation of the mathematical body of quantum mechanics was constructed. During the formation of this theory the described manner was apparently the only possible method of theoretical cognition of the microcosm. But afterwards (namely, at the present-day stage) the vision of quantum objects as complex dynamic systems (large-scale systems) came into being. Analysis of the quantum theory identifies two levels in the description of reality inherent in its conceptual structure: concepts, describing the integrity and stability of the system, and its incidental characteristics.[9] The idea of such division of theoretical description corresponds to the concept of complex systems characterised by the presence of subsystems with stochastic interaction of its elements, and a certain "regulating" level providing for its integrity. This vision of quantum objects is also validated by the findings obtained in the theory of quantum fields which point to the narrowness of the conventional notions about particle localisation (Bell's theorem, Haag's theorem, etc.).[10]

While assessing all these trends in the development of physical knowledge, we should keep in mind that the vision of physical objects as complex dynamic systems was based on the concept that arose from the progress of cybernetics, the theory of systems, and the introduction of large-scale systems into production. During the formation of quantum mechanics, this concept was still missing in science and in ordinary physical thinking objects were not approached as large-scale systems. In this connection the following question may seem appropriate: Could the history of quantum physics unfold in any other way under another scientific environment? In principle (as a thought experiment) we may admit that cybernetics and the corresponding introduction of technological systems could have arisen before quantum physics and promoted a new type of vision of objects in culture. In this context, while drawing a picture of the world, a physicist would have been able to imagine quantum objects as complex dynamic systems could and proceed with constructing the corresponding theory. But in this case the whole subsequent evolution of physics would have been different. Following this route of development, physics was likely to gain as much as lose, since it was not obligatory to explicate the operational scheme of the world vision in this case (consequently, there is no impetus for developing the principle of complementarity). The development of quantum physics on the basis of the complementarity concept, which radically changed the classical standards and ideals of physical cognition, made science evolve along a special route. There appeared a specimen of the new cognitive movement, and even now if physics generated a new systemic ontology (a new picture of reality), this would not signify a return to the previously unrealised route. Ontology should be introduced through the drawing of an operational scheme while a new theory could be evolved by including operational structures in the picture of the world.

The development of science (as any other process of development) means transformation of a potentiality into reality, and not all potentialities are realised in its history. In predicting such processes, a tree of potentialities is always drawn up taking into account different versions and trends of development. The notions of rigidly determined progress of science arise only in retrospect, when we analyse history knowing the final result and restore the logics governing the progress of ideas which led to this result. However, there are possibilities for other trends which could have been realised under other historical developments in human civilisation, but prove to be "closed" in the actual history of science.

During periods of scientific revolution when scientific premises are reconstructed, culture seems to select – out of several potential routes in the future history of science – those which are most consonant with the fundamental values and ideological structures dominating the given culture.

*Institute of Philosophy, Russian Academy of Sciences*

## NOTES

[1] V. S. Stepin, *The Formation of Scientific Theory and Research Programme*, Minsk, 1976, pp. 76–78 (in Russian).
[2] T. S. Kuhn, *The Structure of Scientific Revolutions*, 2nd ed 1970, Chicago 4. Press.
[3] I. Lakatos, *The Methodology of Scientific Research Programmes*, 1978, Cambridge 4 p.
[4] L. I. Mandelstam, *Introduction (From the Prehistory of Radio)*, Moscow, 1948, p. 20 (in Russian).
[5] R. Feinman, *Physics Today*, No. 19, 1966, pp. 31–32.
[6] *Ibid.*, p. 35.
[7] M. K. Mamardashvili, E. Yu. Solovyov and V. S. Shvyrev, *Bourgeois Philosophy: Classical and Modern Stages*, Moscow, 1972 (in Russian).
[8] M. K. Mamardashvili, "Analysis of Consciousness in Marx's Works', *Voprosy filosofii*, No. 6, 1968, p. 19.
[9] Yu. V. Sachkov, *The Style of Thinking in Natural Science Philosophy and Natural Science*, Moscow, 1974, pp. 62–78 (in Russian).
[10] V. I. Arshinov, *The Concept of Integrity and the Hypothesis of Hidden Parameters in Quantum Mechanics. Philosophy and Physics*, Voronezh, 1974 (in Russian).

JOHN STACHEL

# SCIENTIFIC DISCOVERIES AS HISTORICAL ARTIFACTS

In his recent book *Wonderful Life* (Gould 1989, pp. 277–291), Steven Jay Gould notes that Harvard now organizes the sciences "according to procedural style rather than conventional discipline [into] the experimental-predictive and the historical" (*ibid.*, p. 279). While the former, such as physics and chemistry, have often been taken as prototypes for all sciences, Gould emphasizes that:

Historical explanations are distinct from conventional experimental results in many ways. The issue of verification by repetition does not arise because we are trying to account for the uniqueness of detail that cannot, both by the laws of probability and time's arrow of irreversibility, occur together again. We do not attempt to interpret the complex events of narrative by reducing them to simple consequences of natural law; historical events do not, of course, violate any general principles of matter and motion, but their occurrence lies in a realm of contingent detail. (The law of gravity tells us how an apple falls, but not why the apple fell at that moment, and why Newton happened to be sitting there, ripe for inspiration.) And the issue of prediction, a central ingredient in the stereotype, does not enter into a historical narrative. We can explain the event after it occurs, but contingency precludes its repetition, even from an identical starting point (*ibid.*, p. 278).

He then succinctly characterizes the nature of such explanations:

Historical explanations take the form of narrative: *E*, the phenomenon to be explained, arose because *D* came before, preceded by *C*, *B*, and *A*. If any of these stages had not occurred, or had transpired in a different way, then *E* would not exist (or would be present in substantially altered form, *E'*, requiring a different explanation). Thus, *E* makes sense and can be explained rigorously as the outcome of *A* through *D*. But no law of nature enjoined *E*; any variant *E'* arising from an altered set of antecedents, would have been equally explicable, though massively different in form and effect (*ibid.*, p. 282).

Gould adds, perhaps a bit wistfully:

When we have established 'just history' as the only complete and acceptable explanation for phenomena that everyone judges important – the evolution of the human intelligence, or of any self-conscious life on earth, for example – then we shall have won (*ibid.*, p. 283).

When applied to historical sciences that involve human consciousness and agency – to which I shall refer as the telic historical sciences for short – Gould's account must be supplemented by an observation that is quite consonant with his viewpoint.[1] Compared with the non-telic historical



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science,* 139-148.

sciences, a new element enters into the telic ones, the element of intention. Conscious human activity (often called "practice") involves a project, the existence of a conscious goal or goals on the part of the participants in that action. Marx emphasized well over a century ago:

A spider conducts operations which resemble those of the weaver, and a bee would put many a human architect to shame by the construction of its honeycomb cells. But what distinguishes the worst of architects from the best of bees is that the architect builds the cell in his mind before he constructs it in wax (Marx 1976, p. 284).

The reference to inferior architects serves to reminds us that the product of any such goal-oriented human activity, which I shall refer to as an artifact,[2] does not always correspond to the initial intent, even when its production involves only a single agent. Apart from competence, the intent of even the best craftsman will often change in the course of a project. When more than one agent participates in an activity, the goals of some participants may initiate the action, while the goals of others may be defined in the course of their reactions. Negotiations between participants with congruent goals and clashes between participants with incompatible goals assure that the artifact produced hardly ever coincides with the original goal of any single participant. While talk of goals in the non-telic historical sciences is quite misplaced, the element of human intention, of project, of goal should never be ignored in discussing the telic ones.

As noted, goals carry the inherent possibility that they will be modified – perhaps drastically – in the course of an activity, or even abandoned. If not abandoned, they may remain forever unfulfilled because unfulfillable. The objective world confronts humanity with a manifold of potentialities for activity. Clearly, if some project is not in accord with any of these potentialities, no amount of effort will produce the desired artifact – nor can it ever arise serendipitously in the course of striving for other goals. For example, if our present understanding of gravitation is at all adequate – even qualitatively – all projects to construct an anti-gravity machine will fail.[3]

You may have become aware by now that one of *my* goals in this paper is to try to develop a vocabulary for talking about human projects that does not imply that the outcome of such a project is *foreordained* by "objective reality," on the one hand; nor on the other, that "anything goes" – that there are *no* objective constraints on such projects. Whether *my* goal will have to be drastically modified or abandoned depends upon the utility of such artifacts as this paper.

To return to the thread of my argument, which of the densely interwoven web of potentialities will actually be realized as a concrete historical artifact or sequence of artifacts depends upon numerous contingent factors – factors that are not accidental in some absolute sense of that word, but contingent in the sense that the production of these particular artifacts could not be predicted solely on the basis of a knowledge of some initial state and of all the potentialities inherent in it. Indeed, what *some* of the potentialities were often is known only, after the event, and it is doubtful if *all* of them can ever be known; so perhaps it is better to say that the production of artifact, an actualized potentiality, will always appear to depend on factors, the occurrence or non-occurrence of which can be equally well conceived.[4]

What does all this have to do with my topic: "Scientific discoveries as historical artifacts"? If we agree that the history of science is a telic historical science, then all scientific discoveries must be regarded as historical artifacts and studied by the appropriate methods. Perhaps it is not necessary to remind historians of science of this truism, but it is often necessary to remind scientists and philosophers of science that attempts to find a "logic of discovery," a methodology that would fit discoveries into the experimental-predictive mold are fundamentally misguided.

On the other hand – and perhaps here even the historians need a nudge – this observation does not imply that discoveries are not amenable to rational analysis. As applied to historical artifacts, such an analysis must employ the methods of the telic historical sciences. As Gould puts it, patterns of explanation must "take the form of narrative." After the fact, one can make sense of the production of a scientific discovery; its emergence can in principle[5] be rigorously understood as the outcome of some antecedent sequence of events, a sequence that can be causally understood.[6] But if any of these antecedent events had not occurred, or if the events that make up some stage in the process had transpired in a different way, then the event in question would not have occurred, or would have taken a substantially altered form, requiring a different explanation.[7] One must take seriously the lesson of all historical sciences: if we are really constructing a narrative and not a morality play, contingent factors enter in an essential way into the construction of every historical narrative.

How should one go about such a construction? I suggest that we should *not* start from the assumption that there is an ideal world of facts, theories, devices, or what have you, just waiting to be taken off the Platonic shelf ("discovered") by some uniquely creative soul(s). Any attempt to flatten out the rich manifold of intellectual, social, and instutional elements, many of them contingent, that enter into particular discoveries by attributing the entire process to individual creativity uncovering preexistent truth risks ending in tautology and platitude: A creative individual is one who has made a great discovery, while a great discovery is one made by a truly creative individual.

In thinking about how to organize the manifold elements that enter into the production of a scientific discovery, I have been helped by the definition of creativity given by Mihaly Czikszentmihalyi.[8] He raises the question not *what* but *where* is creativity.

> All of the definitions ... of which I am aware assume that the phenomenon exists ... either inside the person or in the work produced ... After studying creativity for almost a quarter of a century, I have come to the reluctant conclusion that this is not the case. We cannot study creativity by isolating individuals and their works from the social and historical milieu in which their actions are carried out. This is because what we call creative is never the result of individual action alone; it is the product of three main shaping forces: a set of social institutions or *field*, that selects from the variations produced by individuals those that are worth preserving; a stable cultural *domain* that will preserve and transmit the selected new ideas or forms to the following generations; and finally the *individual*, who brings about some change in the domain, a change that the field will consider to be creative (Czikszentmihalyi 1988).

By its very nature, then, creativity involves public activity. We might contrast it with talent, the individual capacity to produce artifacts of some type. The act of production of the artifact does not per se constitute part of a creative process. The artifact has to fall within (or in extreme cases initiate) some cultural domain, and receive a positive appraisal by the appropriate audience for the field in question. Of course, more than one person may participate in the production of an artifact; and it is by no means guaranteed that the appropriate individual(s) will receive recognition for their role in its creation. In a somewhat trivialized nutshell, creativity involves talented individuals producing artifacts that are ultimately integrated into some cultural domain by the socially dominant arbiters of that field.

Analysis of a scientific discovery as a historical artifact, then, must involve a discussion of the *goals* of the participants in the discovery. But these goals should not be taken as given. The analysis should consider such questions as:

How the goals of each participant arose in a definite personal, social, and institutional context that reflects the state of the domain and field as it filters through to the participants.

How these individuals went about trying to realize their goals. This involves a process of navigation among the potentialities proferred by nature, some of them already charted but many of them still uncharted – indeed, the object of the search may involve the creation of new charts – in the domain of research; a process of navigation that utilizes the resources – intellectual, instrumental, financial, etc. – of the field that are available to the researchers, i.e., the social and institutional setting of their research.

How each individual's goals were realized, modified (possibly drastically), or abandoned, as the result of successes achieved or obstacles met in the course of this process of navigation; or in the simultaneous or subsequent process of negotiation with others in the domain having congruent goals, and of conflict with still others having opposing or competing goals.

How, out of this complex process of negotiation, there arose a concensus within the domain (or possibly several complementary or even competing sub-concensuses) that defined the nature (or natures, for there are often several more-or-less accepted variants) of the scientific discovery as it is finally accepted into the field.[9] As emphasized above, this outcome (or these outcomes) will be amenable to causal explanation; but it can hardly be considered unique or inevitable: One can often imagine alternate narratives that would have resulted in a different definition of the discovery that is finally incorporated into the field.

Several comments might be added about the nature of such a program for analysing discoveries in the natural sciences as historical artifacts. First of all, it should be clear from what I said that emphasis on goals does not imply a pure methodological individualism; the origins of individual goals would be explained as a result of an interaction of individual temperament and circumstance with the social milieu. Secondly, it does not imply a purely social-constructivist explanation of discoveries. If a particular goal is not in accord with *any* potentialities inherent in the natural world, *no* amount of goal-oriented behavior – no matter how socially reinforced it may be – will ever produce the desired artifact. (To return to my earlier example, the U.S. Air Force was once willing to put a lot of money into anti-gravity devices, but to no avail.)

Finally, such a program does not imply a purely aleatory *or* a purely deterministic explanation of discovery. If their particular social and institutional setting continually motivates a large number of individuals to strive for a cluster of related goals, one or more of which is in accord with some natural potentialities, then the probability is rather high that before too long, one or more of the goals will be realized as a socially useful artifact, in one form or another. The aura of inevitability surrounding many scientific or technological discoveries, often reinforced by their independent occurence several times, can usually be resolved into such a constonance of social motivation conjoined with the inherent feasibility of the resulting project(s).[10]

## II. THE EXAMPLE OF RELATIVITY

I believe that the development of special relativity (SRT) and general relativity (GRT) can be treated with advantage using the approach just indicated. The following brief sketch, drawn primarily from material I have presented in more detail elsewhere, still has to be filled out by considerable further research before making any claim to constitute an account that is adequate when judged by the standards suggested above.

SRT arose from quite a different goal than the one finally attained. Einstein had originally been searching for an electrodynamics of moving bodies that was compatible with all the relevant experimental results, particularly those from optics. This search itself formed part of his wider efforts to give a constructive account of the nature of matter and radiation. The work of Lorentz, Poincaré, Langevin, and others suggests that, if Einstein had not made his contribution, the formalism of special relativity might have been assimilated by the physics community in a context that made no such sharp distinction, as Einstein eventually did, between kinematical and dynamical concepts, a context that probably still would have included the concept of the ether. The shift from the goal of solving problems of electrodynamics to that of a reconstruction of the kinematical foundations of all physics cost Einstein seven years of effort. Even so, in the light postulate, his version of SRT still bears the birthmarks of its origins in electrodynamics, birthmarks that have almost universally been incorporated as an integral element in accounts of the theory. Had Einstein seen clearly in 1905 that he was faced with a purely kinematical problem, he might have realized something that he did not appreciate even after it was pointed out by Ignatowsky in 1909–1910: The light postulate is not needed for a derivation of the Lorentz transformations. The principle of relativity plus suitable assumptions about the homogeneity and isotropy of space and time (assumptions that Einstein also used implicitly), suffice to derive an unique family of kinematic transformations: the relativistic formulae with an arbitrary parameter taking the place of the speed of light.[11] Had Einstein, or the other leaders of the scientific field or community who determined the nature of Einstein's discovery, rejected this final non-kinematical element of the theory, we would have a somewhat different theory than the version today accepted as canonical.

Instead, the transition from Einstein's designation of his work on the relativity *principle* to the domain's characterization of his "discovery" as the *theory* of relativity, and the acceptance of this "theory" by the scientific community, involved a process of negotiation among such augurs of the German-speaking theoretical physics community as Planck, Lorentz, Minkowski, and Ehrenfest, a process that rejected Ignatowski's insight and accepted a certain reading of Einstein's accomplishment. In many cases, no clear distinction was made between Lorentz's "constructive approach" and Einstein's "principle approach" (see the studies of the reception of SR by various national physics communities by Stanley Goldberg and others). What Einstein set out to do and what the community decided he had accomplished were in many respects rather different things.

In contrast to the story of SRT, as early as 1907 Einstein rather clearly formulated the intertwined goals of generalizing the relativity principle beyond the Lorentz transformation group, and of inventing (his favourite term) a relativistic theory of gravitation. While almost all other physicists rejected the first problem, a number of prominent ones worked on the second, but they did so within the confines of SRT. Work on the problem of formulating a special-relativistic theory of gravitation did not stop even after the augers of physics agreed that Einstein had found the "correct" relativistic theory of gravitation by combining the two problems, with the resultant shift to the search for a geometrical, non-flat space-time theory of gravitation. Even so, it was only a series of "lucky accidents" and "fruitful errors" that diverted him from exploring special-relativistic scalar and tensor theories of gravitation that would have been quite compatible with his outlook on the gravitational problem in 1907.

Indeed, after the formulation of special-relativistic quantum field theories, there was an upsurge of interest in such flatspace theories of gravitation because Einstein's non-linear gravitational theory proved so resistant to quantization (it remains so to this day). Suppose Einstein had not been around to give the crucial 'geometrical turn' to the gravitational problem long before the advent of quantum field theory. Could some version of GRT have been developed out of a special-relativistic quantum field-theoretical approach? Feyman actually raised this question in 1957: "Suppose Einstein never existed, and the theory [GRT] was not available" (p. 151). He proceeded to show how a non-linear gravitational theory *formally* identical to GRT could have been developed as the theory of a self-interacting spin-two field in Minkowski space-time – which is just the way most elementary particle theorists insist on viewing Einstein's theory to this day. Of course, as concerns its *conceptual* structure, the resultant theory is *not* equivalent to Einstein's GRT, and this is just my point. Had history taken a different course, we might have had a different version of a non-linear gravitation theory that would not have been a "general relativity theory."

I believe that construction of alternative historical scenarios, such as the ones I have tried to sketch out here for SRT and GRT, serves a valuable purpose in reminding us of the contingent and constructive nature of many features of a scientific "discovery" that, with hindsight, we have become accustomed to regard as inevitable and natural. I emphasize again that the approach sketched does *not* suggest that such contingencies are without their causes, but only that the occurrence of these factors should not be regarded as inevitable. In short, such an approach turns our attention away from attempts to find some sort of *prospective* method of creativity or logic of discovery towards *retrospective* attempts to understand scientific discoveries as historical artifacts.

I have found no discussion by Einstein of the role of social or institutional factors in the development of new scientific ideas. But it is interesting that, rather than "discovery" [*Entdeckung*] (which implies previous existence), or "creation" [*Schaffung*] (which implies complete human control) as the best metaphor for the process that results in a new idea, he seems to have preferred "invention" [*Erfindung*], which implies a process in which human agency acts in accord with something outside of full human control.[12] This preference for the term "invention" rather than "discovery" suggests he was well aware of the role of contingent personal factors in the emergence of such theories. I shall end with another quotation that also suggests he might have favored a retrospective approach:

> A new idea comes suddenly and in a rather intuitive way. That means that it is not reached by conscious logical conclusions. But thinking it through afterwards you can always discover the reasons which have led you unconsciously to your guess and you will find a logical way to justify it.[13]

*Boston University*

## NOTES

[1] I am *not* asserting that *all* sciences involving human agency must be historical. However, the history of science certainly is, so the question of whether there are human experimental-predictive sciences need not be considered here.

[2] I have borrowed this term from Marx Wartosky. (See Wartofsky 1979.) An artifact need not be a material object: a dance, a song, the Greek state, a historical event, are artifacts. On the other hand, a naturally produced object may become an artifact without any physical modification by virtue of its cultural role, e.g., sacred rocks or pools.

[3] An article appeared in the *New York Times* of December 28, 1989, reporting a claim by a two Japanese scientists to have created a weight loss by spinning an object. While this "astounding claim" is "almost certainly wrong," it serves as a salutary reminder of how tentative must be all claims to have exhaustive knowledge of natural potentialities.

[4] This does not imply that they are equally probable, just that the probability for neither appears to be zero.

[5] I say "in principle" because lack of all the relevant historical information (which *could* have been known if adequate records had been kept) usually prevents anything approaching such a complete reconstruction.

[6] Here, the distinction between causality and determinism is important. As I have discussed elsewhere (Stachel 1969), the concept of determinism refers to closed systems, for which a knowledge of the state of the system at any time would permit the prediction of the future (and retrodiction of the past) states of the system at any other time. The concept of causality refers to open systems, for which the effects of external intervention on the system can be lawfully explained. Hence, causal explanation never implies inevitability, since by definition the external intervention is unpredictable on the basis of a complete knowledge of the state of an open system. Universal determinism is the dogma that any open system can be closed by sufficiently enlarging it.

[7] Note that I am here just paraphrasing Gould's account of historical explanation, quoted above, and attempting to apply it to the process of scientific discovery.

[8] The relevance of Cszikszentmihalyi's work was brought to my attention by Howard Gardner (unpublished talk at the Osgood Hill meeting on "Einstein: The Early Years," October 1990).

[9] Here the work of Augustine Brannigan on *The Social Basis of Scientific Discoveries* (Cambridge University Press, 1981) should be incorporated into the program – and will be when this work is elaborated.

[10] For an excellent analysis of the invention of television from this point of view, see Williams 1974. I am much indebted to the ideas of Raymond Williams for my approach to these problems.

[11] Galilean kinematics is just the limiting case of this family when the parameter becomes infinite.

[12] According to Alexander Moszkowski, who reported in Moszkowski 1921 on his extensive conversations with Einstein in 1919–1920: "At first I was almost dumbfounded to hear Einstein say: The expression 'discovery' is itself to be deprecated. For discovery is equivalent to becoming aware of a thing which is already completely formed ... Discovery is not really a creative act!' ... Einstein supplemented this by emphasizing the concept of 'invention,' and ascribed a considerable role to it" (Moszkowski 1921, pp. 100–101).

[13] Einstein to Dr. H. L. Gordon, 3 May 1949 (Item 58–217 in the Control Index to the Einstein Archive).

## REFERENCES


Brannigan, Augustine, 1981, *The Social Basis of Scientific Discoveries* (Cambridge: Cambridge University Press).

Czikszentmihalyi, Mihalyi, 1988, "Society, culture and person: a systems view of creativity," in Robert J. Sternberg, ed., *The Nature of Creativity* (Cambridge: Cambridge University Press).

Richard P. Feynman, 1957, "Critical Comments," in *Conference on the Role of Gravitation in Physics.* Wright-Patterson Air Force Base: Wright Air Development Center Technical Report 57-216, pp. 149-153.

Gould, Stephen Jay, 1989, *Wonderful Life: The Burgess Shale and the Nature of History* (New York/London: W.W. Norton).

Marx, Karl, 1976, *Capital: Volume One*, transl. by Ben Fowkes (Hammondsworth: Penguin).

Moszkowski, Alexander, 1921, *Einstein the Searcher: His Work Explained From Dialogues With Einstein,* transl. by Henry L. Brose (New York: E.P. Dutton).

Stachel, John, 1969, "Comments on 'Causality Requirements and the Theory of Relativity,' " in R.S. Cohen and M.W. Wartofsky, eds. *Boston Studies in the Philosophy of Science* vol. 5 (Boston: Reidel) pp. 179-197.

Wartofsky, Marx W., 1979, *Models/Representation and Scientific Understanding* (Dordrecht/Boston: D. Reidel).

Williams, Raymond, 1974, *Television: Technology and Cultural Form* (London: Fontana/Collins).

PETER MACHAMER

# SELECTION, SYSTEM AND HISTORIOGRAPHY

## 1. INTRODUCTION

This paper is contentious, assertive and programmatic. It attempts to assay certain characteristics of a chronological history of science as practiced. It presents a model based on evolutionary theory that shows how to do that job better. After a brief excursion into the ontological assumptions of history writing, it goes on to sketch a different, systemic model. This systemic model is compatible with the traditional model, but will allow science better to be seen in its complex relation to other human and social aspects. As a *leitmotif* it is argued that intentionality and the mental or their teleological analogues, as an essential part of selectivity, are an ineliminable part of the history of science, however done.

## 2. SELECTION

It is truism that selection is heavily involved in the historian's enterprise. Whether conscious or unconscious, whether given by tradition or made fascinating because bizarre, the historian must attend to some person, episode, period or project that will be studied. What is selected seems to depend differentially on two general kinds of weighted factors: The historian's purpose in writing history, and what is found or picked out as data.

To put this point another way, an historian must start out with some goal. A subject is selected because it somehow interests the researcher. Typical types of interest are that it is now taken, or was sometime thought to be, a crucial event in the history of the world, that it was a person who played an influential role in some set of events, that the institution has not been studied before, that a new document has been found, that new demographics, or new sources for such, shed new light, or that a dissertation deals with a supervisor's cultural hero.

Lying behind topic selection are other intentions. For what more general purpose is this history being undertaken? Three relatively independent motives can be distinguished, though all three tend to operate



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 149-160.

in any given historian at them same time. Pejoratively put, they are: Careerism, antiquarianism and anachronism.

Careerism means the historian selects a topic that will be seen as "hot," will be taken seriously by fellow academics, that probably will be more easily published, that no one has written on before except in an obscure Hungarian language journal, etc. Another name for this is opportunism or positively, professionalism. On this motive, the style in which the history is written will be chosen by what is thought to work best for the interest in question. The underlying motives here are fame, power and/or money.

Antiquarianism means that for some reason the historian finds this part of the past fascinating, just wants to know what happened at that time, or needs this piece to fill in an otherwise relatively complete story. On this motive the style of the history usually will be from the historical actor's point of view or in terms of the institutions as they existed at that time. The basic motive here is that of the collector or connoisseur who finds the history valuable in its own right. But even here what the historian attends to is a function of present interests. This means, for example, there will be contemporary fads of what is of antiquarian interest.

Anachronism means that the selection of past events is determined by what is taken somehow to illuminate the present. This is also called "Whiggish" history or historicism. In these cases, the history is just background to some presentation of the present. There is no typical style. One common form is a linear chronology of events that "lead up to," "explain," or "culminate in" what is of importance now. Another way is to point out parallel events or circumstances between then and now. The basic motive here when noble is that we can learn about the present from the past. When ignoble the motives seem to collapse back into those fads of careerism.

In all these types there is room for many subsidiary motives, for many levels of selectivity. Why does the historian chose to present certain documents and not others? Why did the historian pick out this bit to deal with? Why did the historian find this person important? Even in the case where someone just accidentally finds a document, or stumbles upon something, the question still goes back to why was it deemed to be significant? The basic question is always the same. And the answer always has to be that it was selected *now* because it was taken to be of value.

The ubiquity of selection shows that the doing of history is stuffed through with motives, choices and value judgments. In this sense the doing of history is ineliminably selective and therefore evaluative. There is no "value-neutral" history. There is no "interest-free" history. However, this does not mean that all history is ideologically biased, unless one is perverse about the constitution and scope of ideology.

As an aside, it is worth remarking that selectivity is also present in doing science. Problem choice, selection of experimental design, and the more general fact of the theory laden character of data description, are but some of the selective processes present in doing science. So, minimally we can conclude (*pace* some defenders of the human sciences) that the selective character omnipresent in history *prima facie* does not entail that history cannot be scientific, whatever this claim might turn out to mean on further analysis.

What has been said about history in general also clearly holds for history of science. All these selective intentions and value judgments can be seen in varying degrees in various practitioners. As an exercise to the reader, think about the selections and value decisions made in doing the history of science as it was done by Auguste Comte, William Whewell, Ernst Mach, Pierre Duhem, Paul Tannery, Edwin Burtt, Alexander Koyré, Herbert Butterfield, Giorgio de Santlilliana, Lynn Thorndyke, or Otto Neugebauer.

## 3. AN EVOLUTIONARY MODEL FOR LINEAR HISTORY

Common to most motives mentioned above has been a way of doing history that proceeds chronologically. Actors, ideas, events, institutions, cultural norms and their ilk have been described one following another. In a slight variant, parallel persons, events or institutions have been described showing how they all multiply "caused" the phenomena of interest. These chains have stressed continuities and breaks. Depending upon motive, either the continuity and cumulative effects have been stressed, or the differences, revolutions, or ruptures have been most important.

Yet it seems patently clear that an history that only tells of continuity cannot explain a change (for the difference is not continuous). Nor can a difference be noted without showing the continuity against which it constitutes the different. If there were no change, history would not be interesting or even possible. If there were no continuity, history would be inexplicable or miraculous.

Above, I used descriptions of selective processes and value judgments to talk about historians pursuing their craft. But the selection metaphor also fits well for describing history itself. Chronological histories are stories that relate continuities and differences. The links that make the stories cohere are most often and in some form causal. But causal need not mean law-like. All this fits together intelligibly in an evolutionary model.

Evolution in biology deals with linear continuities and changes and their causal explanations. An adequate evolutionary explanation by natural selection requires the description of the transmission mechanisms which show how the continuity is preserved and the selection mechanisms which show how the differences arise against the background of this continuity (Mitchell 1987, Hull 1988). The continuity mechanisms are spoken of as replicators or transmittors, usually reproduction or genetic identity in biological cases. The selection mechanisms are interactors or adaptors, usually mutation, natural selection, and environmental change in biology. For a complete explanation all these elements must be systematically described to show how they interacted in order to bring about the adaptation and/or newly characterized population.

In dealing with science historians too have recourse to this developmental model, albeit sometimes unknowingly. To talk of transmission, they typically refer to texts cited by the author read or (less strong) available, interactions with other actors, contemporary cultural norms or institutional commitments, systems of patronage or reward, or mutually supporting relations among different contemporaneous patterns of events. Tacit reliance on transmission mechanisms also may mention of the family, education, group membership, shared paradigms or problematics, cultural norms, and, at the extreme, *zeit-geist*. Basically these are persistences in terms of ideas, institutions or cultural necessities.

To explain differences they commonly describe shifts in cultural or personal circumstances, how actors attended to new problems, why certain things at that time became important, and, in general, how an item came to be or came function differently. Sometimes, chance (mutation) has been invoked as in great genius explanations, but this explanatory ploy is now long out of favor. Other different explanations rely on individuals bringing things together in new ways or borrowing from different models or metaphors, individuals or institutions adapting to new political, cultural or social circumstances, or social factors being forced to function in changed material, cultural or ideological settings. Basically these are changes due to internal or external circumstances that lead to an adaptation, which then persists.

What often is missing from these chronological historical narratives is explicit reference to the detailed causal forces that are the mechanisms of transmission or specifics of the forces which interact to bring about differences. How something came to be, what function it has and why it persists are necessary to explain adequately any phenomenon of continuity and change. Anything less is enthymatic, no matter how plausible.

All non-chance mechanisms for explaining difference are selection mechanisms and as such are teleological. This does not require invoking a non-naturalistic teleology, overriding purpose or historical determinism. It does mean that the adaptation has to be seen as attaining some specified goal or end. The goal or end is specified by something outside of the system being studied. But again it need not be specified by something god-like or operating transhistorically. What is necessary is that the new context in which the thing to be explained functions must be specifically described to show how the properties which proved adaptive were selected and how they function in the new context (Machamer 1977).

Selection is always described in terms of choice metaphors, and this licenses the use of intention-like language. It is not that actual choices are made. So it was not that the white moths chose to be eaten in that 19th century smog blackened city near Manchester, England, nor did their avian predators decide that the white moths were tastier than the black. Conditions had changed for both white and black moths as the smoke darkened the environment. So it came to be that the white moths were not camouflaged, and the birds could see them better. The birds had their constant purpose, eating, and conditions enabled them to eat the white moths more readily. This of course meant there were less white moths to reproduce and so the population curve became skewed. Black moths were naturally selected by their adaptive advantage to the changed environment. This is a causal story of changing environmental conditions, moths' attributes (coloring), and birds' hunger. The only intention here belongs to the birds.

That ultimately adequate stories must relate causes only means that the connections or links among the various elements of the story have to be intelligible or understood if they are to be explanatory. Causes related by laws are but one way in which causal linkages are intelligible. Other linkages relate items to human intentions, accepted behavioral patterns or singular observable casual connections.

Further, the necessity of causal intelligibility does not mean that necessary work is not done when one just "fixes" the fact of a change or of a continuity. But this then becomes the data that needs the causal explanation. A historian may chose not to pursue these correlations further. Some believe that there is no more that can be done legitimately than establish the correlations. Causes in history, they say, are so complex that they cannot be determined. But most who say this have in mind some overly restrictive sense of cause. Certainly correlations if adequately established do provide information about what happened. But knowing what happened does not make what happened intelligible. Aristotle, who had a very robust sense of cause, marked this point by distinguishing between a fact and a reasoned fact. What is taken as intelligible depends on the historian's purpose. The historian selects what is intelligible or is better. These principles of selection were outlined above in Section 2.

Suffice it here to draw one interesting entailment. If one is an anachronist historian, or even purports to learn from the past, then the motivations of the actors in the historical episode, or intention-like structures attributed to social institutions or entities will need to be provided. These are the mechanisms of differentiation that can explain changes. Chance (or random mutation) can be a mechanism of differentiation but of course if chance is the cause one cannot learn from it. Neither can one learn from the past merely by noting parallels or correspondences with the present without having ascertained why the past events came out as they did (and not otherwise). Without some differential selection mechanism attributed causally, the citation of contemporary parallels would doom us to repetition of the past or hoping for a chance event. Of course, this position has been held by some.

What needs to be noticed most here, and I repeat it now in structural terms, is that selective causal inter-relatedness is not a two-term relation. The form of the causal relation is not: *A* causes *B*, even where *A* stands for a set of events. It is: An item *i* serves a purpose *c* (fulfills a need, satisfies a desire, etc.) with respect to a system or entity *S* in environment or conditions *C* better than alternatives *j* or not-*i*. The premisses telling us why the purpose is important must be drawn from theories at a different level from or independent of those directly about the system *S*. (Machamer 1977)

## 4. TYPES AND LEVELS OF THINGS SELECTED

Both the types of entities or structures selected for explaining something and those that are to be explained occur at a variety of levels. That is, there are many levels of units of analysis that are used in history. That there are levels seems intuitively clear; exactly how to categorize the different levels is problematic. Below, I use some simple-minded categories, and after will explain why any categories will cause problems. In list form let me remind you of some of the most common types of entities that occur in historical writing:

*Individual human level:* Ideas, cognitive schemas, strategies or goals, intentions, desires for power-fame-money, background beliefs, paradigms, thinking caps, religious beliefs, unconscious needs, leaders, genius, anomie, alienation, sexuality, patriotism, etc.

*Small group level:* Families (mother-son, father-daughter, sibling order), schools, universities, political parties, friends, churches, armies, trade unions, clubs, corporations, etc.

*Larger Institution level:* Educational systems, political structures, legal systems, religious institutions, nations, bureaucracies, transnational entities, alliances, systems of trade, etc.

*Cultural level:* Intellectual fields, habits, shared metaphors, linguistic schemas, languages, kinship structures, economic systems, race, status, rituals, class structures, power, ideology, etc.

*Material condition level:* Climate, diet, agriculture, geographical location, material resources, technology, gender, physicality, etc.

At first sight, what should stand out is how all of these levels are interconnected. In fact, for many items used in explaining, or that are explained, it is not clear to which level they belong, for they operate simultaneously or at different times on many levels. For example, depending upon the construal of the concept of race it could be put in the category of culture or material conditions by virtue of being genetically determined or even at the individual level since it can act as an intentional motive for action. The same is true of gender, power, patriotism, habits, kinship, etc.

Consider the basic unit of some intellectual history, the idea. An idea, in one clear sense, is a property of individuals. That is, individuals have ideas. Yet individuals can share ideas among their peers. They can also pass ideas along to future generations. Indeed, some intellectual history reifies ideas, as sets of abstracted propositions or statements, that are then passed along in baton-like fashion from individual to individual. Still, too, we talk about the ideas or beliefs that characterize a culture. At the extreme we describe an age by the ideas the constitute its *zeitgeist*, or transcendentally the ideas that constitute universal, a-temporal cognitive or rational norms.

Further, most items on any level cannot be adequately explained or defined without appeal to other items on that level and to items appearing at different levels. These are intersystemic concepts. For example, an education system cannot be adequately defined without referring to the schools or universities that compose it, nor without the governmental bureaucracy that includes and empowers it, nor the ideology that directs it, nor the people who work in it, etc.

An actual current example of the inter-relatedness of levels and concepts, which also worries about how these might be applied in doing history is found in Fritz Ringer's discussion of Bourdieu's concept of the intellectual field (Ringer 1990). Ringer characterizes the intellectual field as being made up of agent's taking up various intellectual positions. The intellectual field consists not only of stated theoretical assumptions but also implicit assumptions, preconscious beliefs or mental habits. An historian finds out about intellectual fields by sampling many sources including textbooks, dictionaries, encyclopedias, handbooks, as well as the writings of "great men". The idea is to find structural patterns amongst these and use those to give content to the idea of an intellectual field. The intellectual field is then seen to play the role that the idea did in more traditional intellectual history, but without the reified, transcendental implications.

The necessary existence of such hierarchical levels means that both methodological individualism and social holism are extreme theories that need to be rejected. Neither reduction to individuals, nor simple compositional ploys, nor mindless reification has a chance of making sense out of the various explanatory concepts. In fact, this, too, is a lesson that evolutionary theory teaches us, there are many levels and many types of entities, all of which can be important, relevant and explanatory depending upon the problem or context. How these entities get systematically classified so that they may be used in explanations is a lesson too that we can learn from the history of debates about the taxonomy of evolutionary theories (Hull 1988). But one lesson should be clear, no simple-minded set theoric model of relations of inclusion among these levels and entities will be sufficient to detail the complex relations that exist among them.

As an aside it is also worth noting that given the inter-systemic nature of the concepts and the inter-relation of the levels, it is pointless to embark on projects that attempt to eliminate the subject. Equally foolish are those that naively assume that the individual subject is the main focus of all historical analysis. That is, both a "nameless" history and a "great man" history, if taken literally, are doomed to be inadequate. Here again it is worth noting that human intentions or motives or the intentional equivalents necessary to specify directionality or goal states must be invoked at every level. This is a consequence of selection's being a proper part of every explanation. Selection is what allows us to tell why this happened as opposed to some other thing. The question should then be asked as to what forms selection takes with institutions and social and cultural entities.

## 5. A SYSTEMIC MODEL

In *The Archeology of Knowledge*, Foucault (1969) goes on at length about the "evils" of simple-minded linear history. Foucault writes:

> Archeology, then, takes as its model neither a purely logical schema of simultaneities; nor a linear succession of events; but tries to show the intersection between necessarily successive relations and others that are not so. (Foucault 1969; 1972, pp. 168–9)

Now I am somewhat unclear exactly what Foucault wants, but I think it is reasonably close to what I shall outline below. However, I am going to sketch this alternative model of historiography without restricting it as Foucault does to discursive practices that are constituted by the power and control that institutions have over their members. This would be one interpretation of this model, but I think the model is broader.

The key idea behind the shift to the new model is that too often the portrayal of historical events turns out to be only succession, wherein the "necessary connection" (Kneale 1949) of causal relatedness fails to be shown. Or the portrayal only involves simultaneous events where the relations that are taken as intelligible are part-whole, inclusion hierarchies.

To gain some insight into how to think differently, consider that the normal human way of portraying information (writing or talking) or of gaining information (reading or listening) about things is linear. This not only suggests that any model of information representation will have to be compatible with a linear presentation, but also that humans think in linear terms. So, no alternative model for doing history can do without linear relations, especially linear causal relations. But in history, events and different phenomena, as I have shown in the last section, are highly interconnected. They are temporally simultaneous even as they endure through time, and they are contextually and intersystemically connected at many different levels. Think of them as sets of systems that change through time, where the changes are causally explicable by what happens both within and without of any system selected for study. That at least two systems are needed for any complete explanation follows from our selection schema, as shown in Section 3.

The evolutionary model tries to set standards of adequacy for such changes. The evolutionary model shows us how to establish the necessary connectedness (which is what Foucault wants) by not treating causality as a two term relation. Hierarchies structured in terms of part-whole relations are also too limited, for the simultaneous events we wish most to describe are not moments or atemporal time slices but are themselves occurring in time, and interact among themselves at different times.

Some examples of good history implicitly follow the model I am sketching. I chose for examples some older historical classics in order to show that this model is not new. Take Max Weber's *The Protestant Ethic and the Spirit of Capitalism* (Weber 1904–5; 1958). Here, over a period that runs from the Reformation through the trading practices of the Dutch Republic, Weber tried to weave together skeins that show how the Protestant concept of the calling, the religious status of the "chosen," the practical ethic of work and the emerging structures of capitalist enterprise moved back and forth, in and out across the European continent during that time. I am not interested in how well Weber succeeded in his enterprise for I believe he missed capturing the "necessity." But his goal was to show interconnections among the various religious elements of Protestantism and the practical changes, geographical and functional, that occurred in European life. It was an attempt to show how all these things (and others) moved back and forth among themselves to force a change that would not have occurred otherwise.

An even clearer, and arguably better, example of the same attempt, now mostly confined to England occurs in R. H. Tawney's *Religion and the Rise of Capitalism* (1926). Tawney interrelated the economic revolution in terms of a emerging international money market and forms of produce exchange, with the rise of individualism, the Protestant revolution, land reform, and a dawning sense of economic virtue. The result was that the interests of people, the modes of thought, the patterns of discourse, and humans' conception of themselves is seen to be radically changed by the end of the 17th Century from what it was in the mid-16th.

For an history of science for the period treated by Weber and Tawney, one promising line might select the multi-faceted idea of privacy. Privacy emerged during the late 16th Century in architecture (private bedrooms and hallways), in a newly conceived idea of private (not family) property, and in the epistemology that was part of the new science. In the new science experiments were run by individuals who have to validate results by their own sense experiences. This was part of the burgeoning ideology of individualism. The problem for all domains of the private was how to transcend the subjective and individual in order to make something public and objective. It became an neo-Protagorean age (Machamer 1991).

Further analyzing this systemic model shows that an historian does not know how to reasonably discriminate an element of society, e.g., science, and how it works unless one understands how it is similar to and different from earlier phases of something that might reasonably be called that same element, as well as how it interacts with the co-temporal elements of society, which allow it to be seen as different.

Put simply, if not too clearly, the chronological internal history of a given event is necessary to understand the continuities and differences that make the past different though related to the present state, but also the systemic external history needs to show how this part of life differs from the other parts on which it is yet dependent and with which it develops.

Further, history needs to show what the possible alternatives were to the way in which things turned out, at a time and over time. It is the contrast with alternatives both preceding and simultaneous that ground any claims for necessity and pull the historical narrative away from the merely correlational or from chance. Thus all history must involve the counter-factual as well as the factual. In all explanations the possibilities which did not happen are part of the conditions of intelligibility for what did.

The historian needs to distinguish the state to be explained from past states of this "same" system and show how they are connected, how they evolved. The historian also needs to show co-temporally how this state of this particular system is different from other co-existing systems, or better how these systems (or parts of society) are all inter-related and affect one another as sub-systems within a larger system. The former is strictly in accord with the evolutionary model, while the latter activity is what comes from understanding why all the explanatory terms are only intersystemically intelligible (or definable). One could say, to use the words I heard from my colleague and friend Patrizia Lombardo, that history is both diachronically fluid and synchronically mobile. To which I might add, history is unmitigatedly selective.

*University of Pittsburgh*

## REFERENCES

Foucault, Michel (1969/1972) *The Archeology of Knowledge*, Translated A. M. Sheridan Smith, New York, Pantheon Books.

Hull, David L. (1988) *Science as Process: An Evolutionary Account of the Social and Conceptual Development of Science*, University of Chicago.

Kneale, William (1949) *Probability and Induction*, Oxford.

Machamer, Peter (1977) 'Teleology and Selective Processes', in R. Colodny, ed., *Logic, Laws and Life*, University of Pittsburgh.

Machamer, Peter (1991) 'The Person Centered Rhetoric of 17th Century Science', in Marcello Pera and William Shea, eds., *Persuading Science*, Science History Publications.

Mitchell, Sandra D. (1987) 'Competing Units of Selection?: A Case of Symbiosis', *Philosophy of Science* **54**, 351–67.

Ringer, Fritz (1990) 'The Intellectual Field, Intellectual History, and the Sociology of Knowledge', *Theory and Society* **19**, 269–94.

Tawney, R. H. (1926/1962) *Religion and the Rise of Capitalism*, Glouster, Mass., Peter Smith.

Weber, Max (1904–5/1958) *The Protestant Ethic and the Spirit of Capitalism*, New York, Charles Scribner's Sons.

YORGOS GOUDAROULIS

# CAN THE HISTORY OF INSTRUMENTATION TELL US ANYTHING ABOUT SCIENTIFIC PRACTICE?

1. "While philosophers and historians commonly speak of science in terms of theory *and* experiment, when they speak of the development of scientific knowledge, they speak in terms of theory *alone*." A decade ago, this observation was generally true. The analysis of experimentation and instrumentation is a relatively new trend in the history and philosophy of science. The post-positivistic philosophy of science tended to focus on the theoretical aspects and created a framework unfriendly to the contributions of experiment, instrumentation, and measurement to scientific knowledge.[1] Until very recently, both philosophers and historians of science paid very little attention to the "everyday activities" of scientists in the laboratory.

In the beginning of the 1980s, this "neglect of experiment" and instrumentation was criticized by a group of historians and philosophers of science. Ian Hacking has provided arguments for granting a relative autonomy to both theoretical and experimental practice, and disputed the claim that deliberate experimentation is dominated by theory.[2] The goal of experimenters need not be directed at testing or comparing theories. It may well be directed at investigating the behavior of some unexpected phenomenon and, on occasion, even stretching the available techniques and instruments to the limit, or inventing new ones to do this.

One of the most important consequences of this critique of theory-dominated science is the implication that the successful linkage of theory and experiment is conditioned by the ways in which both are embedded in current technical and social practice. The practical culture of the laboratory and the culture of the laboratory life is, therefore, essential to the production of knowledge. The laboratory as a privileged space for studying science is becoming an important locus for philosophical and historical inquiry. Several recent studies by philosophers, historians, and sociologists of science have focused on scientific experiments and instruments.[3]

2. It now seems extraordinary that historians could consider studying a subject relying heavily on experimentation without feeling the need to



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science,* 161-168.

understand its elements and the complex relationship between the so-called "technical" culture of science and the development of scientific knowledge.

Scientific instruments have been studied mostly as antiquarian objects, cultural artifacts, or even as "heroic devices" responsible for particular scientific breakthroughs.[4] Historians have given detailed descriptions of scientific instruments and their development as an evolution from experimental prototypes to standard laboratory devices and, eventually, to lecture-demonstration apparatus or instruments for everyday use, based on scientific principles. They have studied the structure of the instrument-making trade and its economic history, interactions between the communities of instrument makers, salesmen and experimenters, the role of trade, exploration, navigation, surveying, horology, cartography, astronomy, fortification, etc. in the improvement of instruments, and the establishment of instrument-making centres.[5] They have argued about the value, the authority and repute of various instruments, they even classify them, but there has been little analysis of the various and complex uses of instruments in laboratory sciences. (Their interaction with experiments and the development of scientific concepts, or their impact on scientific method and the changes scientific beliefs are still largely ignored by historians of science.)[6]

Nevertheless, it is true that due to a growing number of recent studies on instrumentation and experimentation by historians and sociologists of experimental science and technology, a big part of Galison's plea for "a historiography with room for a multiplicity of cultures within the much larger rubric of scientific practice"[7] and for

> a history of the material culture of science, but one that is not the dead collection of discarded instruments. [...] a history of the way that scientists deploy objects to meet experimental goals whether or not these were set by high theory; a history of instrument-construction linked to the history of technology; a history that encompasses the relation of instruments to forms of demonstration; a history of the laboratory that tracks the development of the organization of scientific work; and a history of the embodiment of theory in hardware"[8] has been fulfilled.

These studies of the "material" culture of science established a developing experiment-and-instrumentation-orientated history of science. Yet, the problem is not the establishment of a new trend in the historiography of science based on the autonomy of experimental life,[9] or a trend making experiment and instrumentation more important than (or at least as important as) theory. As Hacking says, "of the experimental liberation movement, one of its aims has been not to elaborate the life of experiment, but also to improve the quality of life for theories – along with making the theory/experiment distinction not obsolete but multi-faceted".[10]

Nevertheless, I believe that such studies of the "material" culture of science could deepen our understanding of scientific practice, and that it is beyond any doubt that instruments and instrumental techniques should be regarded as valuable source material in the history of science. Thus, I also believe that it is beyond any doubt that the history of instrumentation "can tell us something not only about scientific practice," but more.

3. Scientific practice encompasses different kinds of activity and has always benefited from the collaboration of theoreticians, experimenters, engineers, and instrument makers. Theories and laboratory equipment evolve in such a way that they match each other and are mutually self-vindicating.[11] In order to make sense of this interrelationship we also certainly need a history of instrumentation, but one that is not a mere description of devices, their invention and construction.[12] We need a history of instruments as embodiments of theories, experiments, and technology.

A device is not in itself a scientific instrument. It can be called an instrument object. If an instrument object is to achieve the status of scientific instrument, it must provide reliable knowledge about nature.[13] It is the scientist who turns the instrument object into a scientific instrument by forcing it to disclose information. Unlike the instrument object, the scientific instrument has to be regarded as information that has been released. The information disclosed by the scientific instrument, and in that sense the scientific instrument itself, becomes an interplay between the instrument object and the scientist.[14].

Thus, it is this interplay between the instrument and the scientist that becomes the subject of the history of instrumentation. Such a history is an inseparable part of the history of science, and it certainly can "tell us many things about (experimental *and* theoretical) scientific practice."

4. A very characteristic and intriguing theme to examine in such a context is the history of gas liquefiers and, generally, the technology of cold. Anything related to the study of the development of low temperature physics is also significant for understanding issues about the relationship between science and technology, since during the last fifty years of the 19th

century, developments in the refrigeration industry brought about lasting changes relating to agriculture, and the distribution and availability of perishable goods in the populous industrial cities.

As far as I know, there is no work on the history of refrigeration, even though there are some manuals that record the main developments, published by the Refrigeration Societies of the U.S.A. and France.

The physical basis for the production of cold is extremely simple: The lowering of temperature depends on the Joule-Kelvin effect, when a gas is suddenly expanded its temperature drops. The trick is to *continuously* decrease the temperature *and* keep it low for a length of time so that it is useful for various purposes.

The history of this method is rather interesting. Apart from Joule and Lord Kelvin, who made a very exhaustive study of this phenomenon both experimentally and theoretically, Dalton, Gay Lussac, Mayer, Rankine, and Clausius contributed largely to the subject. In connection with these works which were conducted for "scientific purposes", a number of experiments were also made with the idea of applying this method to refrigeration purposes and a considerable number of refrigerating machines were constructed. Such machines were adopted successfully for industrial purposes in England as early as 1861.[15]

The essential importance of the "Joule–Kelvin effect" in connection with the liquefaction of gases was first seen by Linde. The cooling due to the "Joule–Kelvin effect" is small; but Linde, by means of the "regenerative" method which was described, developed and extended by C.W. Siemens and E. Solvay, constructed an industrial plant in Germany for the production of liquid air. Low temperature apparatus based on the same principle have been constructed also by Tripler in the U.S.A., by Hampson and Dewar in England, and by Kamerlingh Onnes in The Netherlands.[16]

The problem of efficient heat insulation was solved satisfactorily by Dewar with the invention of the chemically-silvered glass vacuum vessels (the well-known "thermos"), which came to be commonly used in low temperature investigations, as well as in every day life.[17]

The manifestation of this phenomenon in the various cooling machines does not mean that their history is a history of the development in their engineering. And I am questioning whether such a history can best be written using an externalist or internalist approach. I am making the modest claim that the history of the engineering of the cooling machine is about a "different kind" of instrument when compared to the instruments that the history of refrigerators examines, and that both are different from the instruments which are the subject of a history of low temperature physics.

Some of the very same instruments in a particular period may be common to all three pursuits. But that does not make their history the same. Consider the late 19th century. A refrigerator used on a train to carry meat to other parts of the USA, a Linde machine in Germany that supplies liquid nitrogen in large quantities, and a Dewar hydrogen liquefier in London used for measuring the electrical resistance of metals in low temperatures are three *different* instruments since they are involved in three different activities. Thus, they embody three different cultures.

But even two hydrogen liquefiers, based on the same principle and involved in the same activity, should be considered different instruments. Kamerlingh Onnes's hydrogen liquefier, for instance, is not only a technologically advanced instrument in comparison with the above-mentioned Dewar liquefier, it, also, is the embodiment of his 'theorem concerning the law of corresponding states of van der Waals"[18] and of his notion of 'thermodynamically corresponding operations." It is the embodiment of a different experimental *and* theoretical culture. It is the embodiment of a different style of scientific discourse, it is even the embodiment of Kamerlingh Onnes's "potential" that led to the establishment of low temperature physics as a new branch within physics.

In low temperature physics, as in almost every branch (and period) of laboratory science, a complex link was established between instrumentation, theory and the scientist, and the scientific instruments embody this complex relation. Hence, they are inextricably tied to our understanding of key topics in the development of science, and their history should be regarded as an indispensable part of the history of science *and* technology.

*Aristotle University of Thessaloniki, Greece*

NOTES

[1] In fact, both positivists and anti-positivist shared the assumption that there is a universally fixed, hierarchical relation between experiment and theory. Experimentation and instrumentation were, for them, either vulgarized versions of theory, or its primitive building blocks. For positivists observation (a highly schematized and therefore impoverished form of experimentation) forms a cumulative, unproblematic foundation of scientific knowledge and, thus, they devoted practically no effort to the differentiation of types of empirical work.

For many anti-positivists, experiment plays no crucial role in theory choice and, thus, they tended to focus almost exclusively on theory. When experimental work did enter into theory, attention fastened only on experimental results, as they were used to confirm, refute or generate theory, and on the role of theories in the acceptance or rejection of experimental outcomes. In other words, positivists and anti-positivists did not grand a measure of autonomy to the experimental practice. Experimentation had no "life of its own."

[2] (Hacking, 1983). Of course, one may say that by granting relevant autonomy to experimental practice, a philosopher simply recognizes what is obvious to practicing scientists. But, at the same time, he creates the conceptual and historiographical space needed to study the day-to-day culture of experimentation, instrumentation and the procedures of measurement, and to set this culture on an equal footing with highbrow theory.

[3] See for example Galison and Steuwer's papers in Achinstein and Hannaway (1985), Galison (1987), Hacking (1988), Gooding *et al.* (1989), and Baird and Faust (1990). See also the issues of *Isis* (79, no. 3, 1988) and *Science in Context* (2, no. 1, 1988).

[4] See Hackmann (1973; 1989).

[5] See for example, Maddison (1963), McConnell (1980) and Turner (1983a; 1983b).

[6] In fact, only some very recent historical studies (such as Galison (1989), Gooding (1989), Hackmann (1989), and Schaffer (1989)), endorse what Heilbron and others have argued some years for: the history of theory is inseparable from the history of instrumental practices (Heilbron, 1979; Hacking, 1983; and Price, 1984). On the other hand, the role of instruments in bringing about changes in scientific beliefs has been studied only in (generally ahistorical) sociological works (see, for example, Collins (1985) and Pinch (1986)).

[7] "A culture of theory, surely, but also a culture of experimentation, and a culture of instrument building". (Galison, 1988a, p. 211).

[8] *Ibid.*

[9] This autonomy implies that the traditions within experiment and theory do not usually go together. According to Galison, "there are traditions within experiment, theory, and instrumentation; [...] the dislocations within these "subcultures' of physics are not all synchronous; and [...] there are only piece-wise connections between the different strata, not a total convergence or reduction" (Galison, 1988b, p. 526). The invention and perfection of modes of instrumentation define questions for the scientists and ranges of possible answers. Instrumental techniques, among other things, determine what is possible to do in an experimental situation, largely independent of the theoretical structures in the background.

[10] Hacking, 1989, p. 148.

[11] See hacking (1988).

[12] Of course, we can analyse, as Baird and Faust have done in their recent case study of the development of the cyclotron, "the steps in building a scientific instrument – whatever its role in subsequent experiments might be." Such studies, certainly, help us understand some aspects of the instrument-building component. But, considering instruments "in isolation" from the other elements of science is not the best way to deepen our understanding of scientific practice.

[13] Moreover, in some cases it opens up a new world of phenomena. This may be the main difference between a scientific instrument and an "engineering marvel". But, we do not think that there is, or need be, a sharp distinction between science and technology. Watt's steam engine, for example, was certainly an engineering marvel, but it was also a central component in the progress of our "scientific" understanding of heat.

[14] It then follows that while the instrument object may be fixed, the very same instrument can disclose different (and possibly conflicting) information.
[15] See, for example, Wallis-Tayler (1906, pp. 116–125), Marchis (1906), and Voorhees (1909).
[16] See Hardin (1899, pp. 180–213).
[17] It is obvious that the interrelationship between science and technology in every incident related to the development of low temperature physics is so strong, that every effort to write a comprehensive history of this branch of physics without its "material culture" is futile.
[18] Kamerlingh Onnes (1896, p. 5). See also Gavroglu and Goudaroulis (1991, p. xl).

## REFERENCES

Achinstein P. and Hannaway O., eds., 1985, *Observation, Experiment and Hypothesis in Modern Physical Science*, Cambridge: MIT Press.
Baird D. and Faust T., 1990, 'Scientific Instruments, Scientific Progress and the Cyclotron', *British Journal for the Philosophy of Science* **41**(2), pp. 147–175.
Collins, H. M., 1985, *Changing Order*, Beverly Hills: Sage.
Corsi P. and Weindling P., eds., 1983, *Information Sources in the History of Science and Medicine*, London: Butterworths.
Galison P., 1987, *How Experiments End*, Chicago: University of Chicago Press.
Galison P., 1988 a, 'History, Philosophy, and the Central Metaphor', *Science in Context* **2**, pp. 197–212.
Galison P., 1988b, 'Philosophy in the Laboratory', *The Journal of Philosophy* **85** (10), pp. 525–527.
Galison P. and Assmus A., 1989, 'Artificial Clouds, Real Particles', in Gooding *et al.*, 1989, pp. 225-274.
Gavroglu K. and Goudaroulis Y., eds., 1991, *Through Measurement to Knowledge. The Selected Papers of Heike Kamerlingh Onnes 1853–1926*, Dordrecht: Kluwer.
Gooding D., 1989, 'Magnetic Curves and the Magnetic Field: Experimentation and Representation in the History of a Theory', in Gooding *et al.*, 1989, pp. 183–224.
Gooding D., Pinch T., and Schaffer S., eds., 1989, *The Uses of Experiment*, Cambridge: Cambridge University Press.
Hacking I., 1983, *Representing and Intervening*, Cambridge: Cambridge University Press.
Hacking I., 1988, 'On the Stability of the Laboratory Sciences', *The Journal of Philosophy* **85** (10), pp. 507–514.
Hacking I., 1989, 'Philosophers of Experiment', *PSA 1988* **2**, pp. 147–156.
Hackmann W. D., 1973, *John and Jonathan Cuthbertson: The Invention and Development of the Eighteenth-Century Plate Electrical Machine*, Leiden: Museum Boerhaave.
Hackmann W. D., 1989, 'Scientific Instruments: Models of Brass and Aids to Discovery', in Gooding *et al.*, pp. 31–66.
Hardin W. L., 1899, *The Rise and Development of the Liquefaction of Gases*, London: Macmillan.
Heilbron J. L., 1979, *Electricity in the Seventeenth and Eighteenth Centuries*, Berkeley: University of California Press.
Kamerlingh Onnes H., 1896, 'Remarks on the Liquefaction of Hydrogen, on Thermodynamical Similarity, and on the Use of Vacuum Vessels', *Communications from*

*the Physical Laboratory at the University of Leiden* **23**.

Maddison F.R., 1963, 'Early Astronomical and Mathematical Instruments. Brief Survey of Sources and Modern Studies', *History of Science* **2**, pp. 17–50.

Marchis L., 1906, *Production et utilisation du froid*, Paris.

McConnell A., 1980, *Geomagnetic Instruments Before 1900*, London: Harriet Wynter.

Pinch T., 1986, *Confronting Nature*, Dordrecht: Reidel.

Price D. J. de S., 1984, 'Of Sealing Wax and String', *Natural History* **93**, pp. 48-56.

Schaffer S., 1989, 'Glass Works: Newtons Prisms and the Uses of Experiment', in Gooding *et al.*, 1989, pp. 67-104.

Shapin S. and Shaffer S., 1985, *Leviathan and the Air-Pump*, Princeton: Princeton University Press.

Turner G. L. E., 1983a, 'Scientific Instruments', in Corsi and Weindling, 1983, pp. 243–258.

Turner G. L. E., 1983b, 'Mathematical Instrument-Making in London in the Sixteenth Century', in Tyacke, 1983, pp. 93–106.

Tyacke S., ed., 1983, *English Map-making 1500–1650*, London: The British Library.

Wallis-Tayler A.J., 1906, *Refrigerating and Ice-Making Machinery*, London.

Voorhees G.T., 1909, *Refrigerating Machines, Compression, Absorption*, Chicago.

DIONYSIOS A. ANAPOLITANOS AND APOSTOLOS K. DEMIS

# THE ONE IN THE PHILOSOPHY OF PROCLUS: LOGIC VERSUS METAPHYSICS

An important issue concerning the history of logic and philosophy is the appearance of paradoxes in various logical or philosophical systems of the past. This issue is methodologically interesting, from a historiographical point of view, because sometimes one has to make use of logic and metaphysics in juxtaposition to be able to deal with them. That is, one has to make use of criteria which, although not available at the time, are necessary to examine, interpret and understand logical scientific and philosophical systems of the past. This is especially the case whenever one is forced to deal with the issue of examining first principles in a philosophical system, which are postulated so that their indescribability is one of the basic ingredients of their nature.

In the present work we are going to deal with the issue of the describability or indescribability of such a principle, as it appears in the system of the neoplatonic Proclus. More specially:

(a) we will present Proclus's position,
(b) we will refer to a proposal made in the past for overcoming the logical difficulties connected with Proclus's position,
(c) we will propose the adoption of certain requirements before one attempts to deal with such issues and
(d) we will show that such an adoption makes room for Proclus's position and more generally for Proclus's philosophical system.

The basis characteristic of Proclus's philosophical system is that it contains an ontology that consists of a stratified and well-structured collection of classes of individualities, which obey a set of metaphysically ascertained laws. The dominant position in the ontology is occupied by what is called First principle (πρώτη αρχή). This First principle is, according to Proclus, both the ontological and the logical founding of the existence of every individuality, as well as the possibility or impossibility of an epistemological access to every such individuality. Usually Proclus refers to this First principle in his writing by using the name "the One" (τό ῞Εν). According to him the One is beyond any affirmation. It is also



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science,* 169-175.

absolutely indescribable and unknown. Every assertion or negation about it is false.

Proclus's position on the One has caused relevant interpretational and critical reactions.[1] We will examine one of them that we find interesting for the purposes of the present work. According to it the following propositions can properly encode Proclus's positions on the One.

A: *No predicate can be assigned to the One.*
B: *The One is such that no proposition can be true of it.*
C: *The One is such that it is true that the assignment of any predicate to it is a category mistake.*

The above propositions are followed by some criticism.

Concerning proposition A a predicate $\phi$ is introduced in the following way:

> For every $x$,$x$ is $\phi$, if and only if $x$ is such that no predicate can be assigned to it.

It is then obvious that the predicate $\phi$ can be assigned and cannot be assigned to the One. Hence the One is and is not $\phi$ and therefore proposition A cannot be accepted.

Concerning proposition B it is noted that it is self-referential and therefore, as a proposition concerning the One, cannot be true. That is, proposition B cannot be accepted.

Concerning proposition C it is observed that its acceptance is equivalent to the acceptance of the existence of some kind of knowledge of the One.

Furthermore, another problem arises. Let us define a predicate $\phi$ as follows:

> For every $x$, it is true that $x$ is $\phi$ if and only if, for every predicate it is true that its assignment to $x$ is a category mistake.

Using $\phi$ we can give proposition C the form:

> C': The One is $\phi$.

Proposition C', then, (i) will be true since C is true according to the adoption of the propositions A, B, C as faithfully representing Proclus's positions and (ii) will be a category mistake, given that the predicate $\phi$ is assigned to the One.

Since it is absurd that a proposition can be, at the same time, true and a category mistake, proposition C', and therefore C, cannot be accepted.

Following the above, it can be deduced that Proclus's position that the One is absolutely indescribable is logically invalid and therefore, if the One is to have some meaning, it has to be accepted as having some properties of a very special character and that there are some peculiar predicates which can be assigned to it. It is by now quite obvious that the way Proclus treats the possibility of knowing or of describing the One is definitely a problem. Yet, we will show below that this is not a real problem. That is, we will show it is not a problem concerning the substance of Proclus's philosophy, but rather a problem which stems from the way the philosopher uses natural language in order to express himself.

Before attempting to do so, however, we will express some thoughts which concern general matters of method. Following, we will give our own interpretation of what we think is Proclus's philosophy of the One and we will make clear which are the weak points of the previous criticism of propositions A, B and C.

The general methodological requirements which we think have to be adopted for dealing with problems like the above-mentioned are the following:

(i) The first concerns, as much as possible, the careful separations of the ontology of the system under scrutiny (philosophical or other) from the researcher's mental reality. Anachronistic interpretational projections upon the special reality of the examined system should be avoided.

(ii) At the level of the natural language used by the creator of the system, the researcher should carefully attempt to distinguish and examine different linguistic levels which possibly exist therein and which the creator of the system was, probably, not aware of. Such an examination could, quite possibly, lead to an adequate resolution of the apparent self-referential so-called paradoxes that could be discovered at first glance in the ontology of the examined system. For instance, careful examination of Proclus's system would lead, we think, to distinguishing at least two separate linguistic levels in the natural language the philosopher uses to describe his system.

At the first level, for which we can use the term "language", the philosopher talks about different sorts of individualities existing in his system, describing them, assigning properties to them and properly distinguishing them from each other. At this level, no second-order quantification takes place; propositions like A, C and C' do not belong to this level. At the second level, for which we can use the term "metalanguage", the philosopher quantifies over properties of individualities, he uses semantic terms like "truth" and "falsity" and makes semantic evaluations and, we think, he talks about the unknowability and indescribability of the One. We should note that the language is closed and ontologically independent of the philosopher, of course, as soon as this language has been established during the formation of the philosophical system. On the other hand, the metalanguage, as the language of the semantics of the system, belongs to the realm of the evaluation which could be, and most of the time is, open-ended and therefore dependent upon the philosopher. In other words the philosopher is independent of the purely first-order, descriptive part of the philosophical system and the language which is used to describe it; he uses the natural language as a metalanguage dependent upon him to refer to the system and to the use of the language of the system. We call this second-order freedom of choice of the semantics of the philosopher the *Principle of the Independence of the Philosopher*. We could say that, in a sense, the examiner of the philosopher's work shares this independence. As a matter of fact he is even freer than the philosopher in the sense that he can explore the possibilities of semantic development of the system that the philosopher could not have foreseen.

We can now return to and start afresh the examination of Proclus's position concerning the indescribability of the One. According to him the knowing of an individuality is not necessarily equivalent to the possibility of describing it or, even less, of assigning a predicate to it. Describability or the assignment of a predicate to an individuality are, in Proclus's system, dependent upon the particular individuality's logos (λόγος). Knowledge, on the other hand, can exist even without logos (ἄλογος γνῶσις).[2] Such a kind of knowledge is, for instance, the sense perception.[3] Yet, knowledge with logos is the knowledge of individualities, which in Proclus's system are classified as beings (ὄντα).[4] That is, the following proposition is true according to Proclus:

I. *If an individuality is describable or a predicate can be assigned to it then this individuality is necessarily a being.*

Knowledge of a being together with its logos is what helps us explain, describe and present it. The logos is the necessary ontological offspring of every being.[5] The One, on the other hand, is indescribable and unknown. It is logically and ontologically prior to all the other individualities of Proclus's system. So, according to Proclus, a second proposition is true.

II. *The One is not a being.*[6]

An immediate consequence of propositions I and II is the following:

III. *The One is indescribable and no predicate can be assigned to it.*

If we were to express all this through the construction of a formal language, we could start with the language's set of symbols having among them a set of names for all individualities of Proclus's system (the One included) and a set of predicate symbols for all predicates assigned to those individualities. Then, in order to express linguistically the indescribability of the One, we could add a metalinguistic requirement according to which:

> For every predicate symbol $\phi$ the expression of the language "the One is $\phi$" is not a sentence.

In other words, our position is that Proclus's thesis that the One is indescribable does not create a problem because the notion of indescribability belongs to the semantics of the metalinguistic level.

This is not an artificial nor ad hoc construction. It comes out of the ontological requirements of Proclus's system which are requirements not in the system, but about it.

We can now return to our initial discussion of propositions A, B and C. We think that all the criticism of Proclus's system which was based on these propositions is invalid for two reasons:

First, we think that the permissible predicates in Proclus's system are only the ones which can be assigned to beings. Artificially produced predicates like the ones of propositions A, B and C are self-referential, second-order predicates that can be found nowhere in Proclus's system. They can be used, of course, at a metalinguistic level to refer to the system and to the use of the language of the system, always according to the Principle of the Independence of the Philosopher already mentioned, but they cannot be used at the level of the language which talks about first-order properties and relations of the individualities.

Second, we think that propositions I, II and III are second-order propositions which define, in a sense, the permissible ontological and linguistic boundaries of Proclus's system, which cannot be crossed and which cannot be confused with the internal parts of the system. Proclus's system contains only natural predicates which are related to the nature of beings (οντα) in such a way that knowledge with logos can only be of beings.

If we were to return more specifically to proposition B, we could add some more comments. The word "proposition" appearing in B refers to the propositions of a language to which B itself does not belong. Therefore, the semantics of B should not be mixed up with the usual semantics of the propositions the word "proposition" appearing in B refers to. Additionally, we think, that although Proclus could not have had the much later-invented distinction of language and metalanguage, he was aware of the existence of second-order, self-referential propositions like B, as the following quotation from his writings shows:

...in short, it is common to those things to have nothing in common, thus reason itself undermines itself ...[7]

It is clear from the above that Proclus was aware of the dangers of the self-referential, second-order predicates and propositions. So if proposition B had been accepted by Proclus as expressing something basically true, one of the following two things would hold. Either he know the logical antinomy lurking behind it, pretending he did not, or he was unconsciously using the Principle of the Independence of the Philosopher, that is, he had the feeling that propositions like B were not of the usual type and therefore should not be used as such.

Summarizing our position we could say that Proclus's system is hierarchically organized in such a way that self-referential propositions are not internal to the system. When constructed such propositions are highly artificial and belong to the metalanguage of the system. We think the best way to interpret Proclus's contention that the One is indescribable is by admitting that the notion of indescribability is of second-order and, therefore, not expressible by any predicate internal to his system. The interpretation proposed here is based upon a careful distinction of different linguistic levels in Proclus's work and upon the Principle of the Independence of the Philosopher, which can be well accommodated in a consistent reading of Proclus's positions. For instance, it is clear for the philosopher that there are three separate and interconnected levels in his ontology. According to him:

There are three which are connected together, the things, the conceptions and the logoi.[8]

These three levels are the level of reality, the level of the conception of reality and the level of the language which isomorphically maps the conception of the reality, a conception which is of predicative nature. At the level of reality, the philosopher himself is included. According to the

Principle of the Independence of the Philosopher, such an inclusion does not affect the second-order freedom of choice of the Philosopher, a freedom which has to do with this theorizing over and above it all. Second-order propositions are constructed freely and independently by the Philosopher, who is responsible for dissolving the self-referential antinomies by appropriately distinguishing different levels of language, corresponding to different referential levels.

*University of Athens, Greece*

NOTES

[1] See, for instance, Carl R. Kordig: Proclus on the One. *Idealist Studies* **3** (1973), pp. 229–237. Kordig uses the last part of Proclus's commentaries to the platonic Parmenides. [*Corpus Platonicum Medii Aevii,* ed. R. Klibansky; *Plato Latinus III, Parmenides usque ad finem primae hypothesis nec non Procli Commentarium in Parmenides, pars ultima inedita,* eds. R. Klibansky and C. Labowsky, trans. G. E. M. Anscombe and L. Labowsky (London: Warburg Institute,1953).]
[2] In Tim. I 248,30 ff.
[3] *Ibid.* 218, 9–10.
[4] El. Th. pr. 123.
[5] In Tim. I 340,22 ff.
[6] See for instance: In Plat. Theol. lib. III,ϛ'.
[7] In Parm. 725, 37–39.
[8] In Tim. I 339, 5–6.

THEODORE ARABATZIS

# RATIONAL VERSUS SOCIOLOGICAL REDUCTIONISM: IMRE LAKATOS AND THE EDINBURGH SCHOOL

## I. INTRODUCTION

The publication of Thomas Kuhn's *The Structure of Scientific Revolutions* initiated a new era in history, philosophy, and sociology of science. Its influence on history of science, though pervasive, has been indirect. The model of scientific development expounded in *Structure* has never been fully applied (not even by Kuhn himself) for elucidating a past scientific episode.[1] On the other hand, by indicating that the very content of scientific knowledge is amenable to sociological analysis, it had a significant effect on sociology of science and thus indirectly on history of science. Given the well-known and profound transformation that Kuhn's work effected in our philosophical understanding of the nature of scientific knowledge and its considerable effect on sociology of science, its indirect influence on history of science should not be underestimated. Two recent historiographical research programs, Imre Lakatos's *Methodology of Scientific Research Programs* and the 'strong program in the sociology of science' associated with a group of sociologists in the University of Edinburgh, emerged in an attempt to respond to or develop certain aspects of *The Structure*. Their aim was to reconstruct past scientific episodes either, as in Lakatos's case, in the light of a philosophical theory of scientific rationality, or, as in the strong program's, in the light of a sociological theory of scientific practice. Since both programs were considerably influenced by Kuhn and, as I will argue below, were reductionist in that they aimed at reducing historical explanation to a rational or sociological core, it is instructive to make a comparative evaluation of them.

The goal of this paper is to offer a concise critical exposition of these two historiographical approaches, to highlight their differences and common aspects, and to discuss their relevance for contemporary historiography.[2] The core of my argument will be that there are some striking parallels between rational and sociological reconstructions of the history of science. In particular, both are fundamentally ahistorical in that they dismiss the explanatory value of the historical actors' own reasons for upholding their beliefs and pursuing their actions.



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science,* 177-192.

## II. HISTORY OF SCIENCE AND ITS RATIONAL RECONSTRUCTIONS

### II.1. *Rationality and Methodology*

Lakatos's *Methodology of Scientific Research Programs (MSRP)*[3] was developed to meet the challenge that Kuhn's work posed for scientific rationality. According to Lakatos, Kuhn portrayed paradigm replacement as a fundamentally irrational process. Given the then prevailing notion of rationality, Lakatos was, of course, right. This notion had its origins in Reichenbach's codification of the two contexts within which scientific activity takes place, the context of discovery and the context of justification. The former consists in the processes of discovery of scientific hypotheses and theories; the latter in their testing and validation. In Reichenbach's view the context of discovery was the province of historians, psychologists, and sociologists and was not susceptible to logical analysis. On the other hand, the context of justification was an area which could be rigorously explored and formalized and thus fell within the province of logic and philosophy.[4] Standards of rationality applied only to the context of justification. The rules which governed the process of justification defined the canons of scientific rationality. Thus, in the logical positivist view of science 'rational' meant rule-governed. Kuhn, on the other hand, denied that rigid and precise principles of rationality unambiguously determine paradigm-choice. So, if we adhere to the positivist notion of rationality we are led to the conclusion that paradigm replacement is an irrational process.

Kuhn, however, did not develop an alternative theory of scientific rationality, appealing instead to 'socio-psychological' factors to explain paradigm-change. Lakatos, on the other hand, proposed his *MSRP* as a theory which captured the essence of scientific rationality; a theory which could be employed to explain the development of science over the last four centuries in predominantly 'rational' terms. For the purposes of this essay it is not necessary to present in detail Lakatos's methodology, since my assessment of its value for historiography will be relatively independent of its particular features and its philosophical merits.[5] Suffice it to say here that for Lakatos

> The basic unit of appraisal must be not an isolated theory or conjunction of theories but rather a '*research programme*', with a conventionally accepted (and thus by provisional decision 'irrefutable') '*hard core*' and with a '*positive heuristic*' which defines problems, outlines the construction of a belt of auxiliary hypotheses, foresees anomalies and turns them victoriously into examples, all according to a preconceived plan.[6]

It is the machinery of research programs that is to be employed to evaluate the rationality of past scientific episodes. However, one is tempted to ask in what sense Lakatos's methodology would explain those episodes. But before discussing this question we need to elaborate on Lakatosian historiography.

Following Agassi,[7] Lakatos argued that philosophical and normative elements inevitably enter the historian's reconstruction of past scientific practice by influencing her selection and interpretation of historical data. However, Lakatos moved a step further and asserted that historical reconstructions should be explicitly carried out in the light of the best available theory of scientific rationality, namely the theory which portrays most of past scientific episodes as rational. Not surprisingly, Lakatos believed that the best available model of rationality was his *MSRP*. The latter would enable the historian to delineate the domain of internal history, i.e., the domain of past scientific developments which appear as rational in the light of the *MSRP*. When a past scientific episode can be portrayed as rational in the light of the *MSRP* this does not mean that all historical actors who participated in that episode followed the specific methodology proposed by Lakatos. It only means that the methodology would result, in most cases, in the same judgements, decisions, and actions as those of the relevant scientific élite. The actual reasons which led this group to make these judgements and to follow these actions are, according to Lakatos, irrelevant from the point of view of methodology and internal history of science.

One point that should be emphasized for my analysis is that Lakatos's notion of internal history is, as he put it, 'unorthodox' and comprises only a few of the elements which would fall under the customary conception of internal history. Internal history, in Lakatos's sense, not only excludes the institutional, cultural, and socio-economic context of scientific practice but also the "scientists' beliefs, personalities or authority. These subjective factors are of no interest for any internal history." Moreover, it excludes "everything that is irrational in the light of his [the historian's] rationality theory"; its sole concern is the rational 'growth of disembodied knowledge'.[8]

Lakatos recognized that no theory of rationality could ever portray every episode in the history of science as completely rational, either because scientists *qua* humans are not completely rational or because the cultural, political, and social context of scientific activity might unduly influence its rational development. Thus, he held that when the historian focuses on such irrational episodes, i.e., past scientific developments that did not take place as they should have according to the *MSRP*, then she is to explain those aberrant events by resorting to 'external' socio-psychological factors.[9] External history, thus, becomes parasitic on the internal(rational) history of scientific knowledge. The latter determines the subject matter and scope of the former which, incidentally, "is irrelevant for the understanding of science."[10]

## II.2 *Lakatos and Historiography*

If Lakatos's aim were only to rationally reconstruct the majority of past scientific developments in the light of his methodology, the completion of his project, if successful, would have provided a rational explanation of the growth of knowledge. That is, it would have shown in what respects victorious research programs in the history of science were superior to their competitors from the point of view of a timeless methodology. But Lakatos had a more ambitious aim, namely to provide a model for historiographical practice.[11] However, if the *MSRP* is to be of some use for the historiographical enterprise, as usually conceived and practiced by historians, that methodology must have been employed, if only implicitly, by the historical actors in the scientific episodes under scrutiny. Only then it could have constrained their decisions and could, thus, function as an explanatory resource for the historian. Otherwise, despite its potential significance for providing a *posthoc* justification of the historical actors' decisions, it could not be of any use for explaining why the actors themselves were led to those decisions.[12]

Thus, to show the universal historiographical applicability of the *MSRP* one needs to demonstrate that it was shared by the majority of scientists who had an impact on the development of scientific knowledge. Such a demonstration cannot be found in Lakatos's writings. Instead, Lakatos seems to have assumed without argument that scientists have been, for the most part, adherents of his methodology. Otherwise, one can not make sense of his insistence on the need for socio-psychological explanations of all those episodes in the history of science which cannot be portrayed as rational in the light of his methodology. Only if the *MSRP* was shared by the majority of scientists and, thus, constrained their decisions, would any deviation from its prescriptions be in need of 'external' explanation. In the absence of such an assumption, those 'deviant' episodes might be explained adequately – without recourse to socio-psychological factors – by recovering from the historical record the values of appraisal shared by the scientists in question. In that case, the difference between these values and those prescribed by the *MSRP* would account sufficiently for the 'irrational' aspects of those episodes.

The heart of the problem is that Lakatos tended to conflate the justificatory and the explanatory aspects of his methodology. If the triumph of a research program over its competitors could not be justified from the point of view of his methodology he assumed that that victory could only be explained by appealing to socio-psychological factors. If, on the other hand, a past scientific episode took place as it should have according to his methodology he assumed that it was thereby completely explained. The latter assumption pervades most of the historical case studies that have been carried out by Lakatos's followers in accordance with his methodology. But, as Kuhn pointed out, those studies fail to consider "what actually attracted scientists to or repelled them from the various research programmes under study. ... If analysis discloses a philosophically relevant difference between research programmes, then that difference is assumed to have played a role in programme choice."[13] Indeed a successful reconstruction of a particular development would explain why it was a rational episode in the growth of knowledge. What it would not explain, however, is why the actors of that episode adopted certain beliefs and made certain decisions leading to that outcome, which is the historian's question.

These untenable assumptions, which are due to Lakatos's conflation of the explanatory and the justificatory aspects of his methodology, proved fatal for his historiographical approach. Historians are not concerned to justify in an atemporal and context-independent sense past scientific beliefs, actions, and decisions; rather they attempt to describe and explain these beliefs, actions, etc., as 'reasonable' solutions to the specific problem situation faced by the scientist under consideration. To illustrate my argument consider Newton's belief in the existence of absolute space. In order to justify this belief Newton would have given, among other reasons, certain theological arguments. For a modern secular historian these arguments would not carry much weight. Nevertheless, in trying to explain Newton's belief in the existence of absolute space, our secular historian should certainly appeal to these particular theological arguments, since they were employed by Newton himself as warrant for his belief in absolute space. Historical sensitivity demands that the historian should adopt, to the extent possible, the mindset of the historical figure whose beliefs and actions she tries to explain.[14] In the ideal case the recovered rationality would amount to the specific epistemic reasons which the historical actors themselves would have given as warrant for their beliefs, decisions, and actions. Needless to say, what counted as a reason for a specific historical actor need not count as a reason for the historian herself.[15]

The insensitivity of Lakatos's methodology to the categories and criteria of historical actors renders it unsuitable as a historiographical tool. A similar insensitivity, as we will see, characterizes the strong program in the sociology of science.

## III. SOCIOLOGICAL RECONSTRUCTIONS OF THE HISTORY OF SCIENCE

### III.1 *Three Theses of the 'Strong Program'*

*The Structure of Scientific Revolutions* exerted a profound influence on the sociology of science. Pre-Kuhnian sociology of science focused on the study of scientific institutions, the reward system of science, the norms which constitute the ethos of the scientific enterprise, and the social roles of scientific practitioners and abstained from a sociological analysis of the content of scientific knowledge.[16] When Kuhn indicated in *Structure* that certain aspects of the development of scientific knowledge, and especially theory-choice, "are irreducibly sociological",[17] the road was open, so the proponents of the strong program[18] thought, for a total sociological reductionism. It is beyond the scope of this paper to give a satisfactory account of the intellectual origins of the strong program. Suffice it to say that the work of earlier sociologists of knowledge, most notably Durkheim and Mannheim, along with Kuhn's view of scientific knowledge exerted a formative influence on its development.[19]

The strong program is identified with the following theses, "of causality, impartiality, symmetry and reflexivity."[20] I will focus exclusively on the first three, since the issue of reflexivity – the applicability of sociological explanatory models to sociological knowledge itself – though crucial for the viability of the strong program, is without historiographical significance.[21] These tenets assert that the proper sociology of scientific knowledge "would be causal, that is, concerned with the conditions which bring about beliefs or states of knowledge." Furthermore, "It would be impartial with respect to truth and falsity, rationality or irrationality, success or failure. Both sides of these dichotomies will require explanation." Finally, "It would be symmetrical in its style of explanation. The same types of cause would explain, say, true and false beliefs."[22]

The thesis of causality is relatively uncontentious, apart from the fact that the philosophically problematic notion of cause is left unexplicated. The thesis of impartiality emphasizes that all beliefs and actions regardless of their epistemic, rationality, and pragmatic status should be candidates for explanation. As for the impartiality principle, it is, again, uncontentious: no one has disputed, to the best of my knowledge, that all beliefs and actions, irrespective of their status, should be candidates for explanation. The symmetry thesis, as one would expect, has been intensely debated.[23] The core of the dispute concerns the range of factors that could be legitimately invoked to explain a historical actors' beliefs. The symmetry thesis entails a very radical perspective on this issue. Since, by their very nature, epistemic factors cannot be employed to explain 'irrationality' or 'unsuccessful' beliefs it follows from this thesis that they should be dispensed with altogether as an explanatory tool. Thus, the historian and the sociologist of science who adhere to this principle would need to exclude epistemic factors, i.e., the reasons that could be invoked to justify the beliefs in question, from their explanatory repertoire. To put it another way, given that epistemic factors do not bear on the explanation of irrational or unsuccessful beliefs and, moreover, that all beliefs, regardless of their status, should be explained 'by the same types of cause' it follows that 'rational' and 'successful' beliefs should also be explained by non-epistemic factors.[24] Notice how radical and one-sided the symmetry thesis is. Even Lakatos, who was obsessed with rationality, had not suggested that social and psychological factors have absolutely no place in legitimate reconstructions of the history of science.

Lurking behind the symmetry thesis is the underlying question: Whose evaluation of the status of beliefs cannot be employed to explain why those beliefs were adopted? The historian's and sociologist's retrospective evaluations or the historical actors' own appraisals? As far as I can tell, no answer to this question can be found in Bloor's *Knowledge and Social Imagery*. In a later article, however, Barnes and Bloor assert that "regardless of whether the sociologist evaluates a belief as true or rational, or as false and irrational, he must search for the causes of its credibility."[25] Barnes and Bloor's concern is with the sociologist's evaluative stance towards the beliefs that he is trying to explain. On the other hand, they say nothing about the potential significance of the historical actors' own appraisal of the beliefs in question for the explanatory task of the historian or the sociologist. I will have to say more on this point below, since it will become important for my assessment of the symmetry thesis.

### III.2. *The Strong Program and Historiography*

The historiographical implications of the symmetry principle are clear. Historical episodes should not be explained by the reasons that would justify their outcome. Rather, the sole explanatory resources of the historian should be either macro-social parameters, e.g., class membership or micro-social factors, e.g., the professional interests of the participants in the episode under study. The latter are particularly important since, according to Bloor, "Much that goes on in science can be plausibly seen as a result of the desire to maintain or increase the importance, status and scope of the methods and techniques which are the special property of a group."[26] As Thomas McCarthy pointed out, all of these explanatory factors appear to have nothing "in common, except perhaps that they are *not* the sorts of evidencing reasons actors *themselves* would give for their beliefs."[27]

It would be tempting to attempt a sociological reconstruction of Barnes and Bloor's belief in the central principle of the strong program,[28] but for my purposes it is preferable to offer a 'rational reconstruction' of their path to the symmetry thesis. Their central argument goes as follows: There are no context-independent and supracultural norms of rationality. In other words what counts as a reason is highly context-dependent. Thus, in explaining a scientist's belief, decision, or action we cannot appeal to reasons which would constrain any rational agent, because such reasons simply do not exist. The conclusion seems to follow that all beliefs, decisions, etc., should be explained in the same way, i.e., in sociological terms.[29]

There is, however, a crucial flaw in the above argument; a flaw which results from Barnes and Bloor's ahistorical conception of rationality and from their disregard of the historical actors' own appraisals of the beliefs, etc., in question. The fact that there are no transcultural norms of rationality does not imply that the members of a specific community do not share certain values that define and regulate rational behavior. These norms, to the extent that they were shared by the members of the community, would constrain human actions and should thus be taken into account in explaining these actions.

The historiographical implication of the symmetry thesis is that the historian cannot use the historical actors' own appraisal of their beliefs and actions in her explanatory task. In other words, that the reasons that the historical actors would offer in support of their beliefs do not *really* explain why they adopted them and that a genuine explanation should incorporate the social factors which (supposedly) underlie the beliefs in question. I do not deny that cases of 'false consciousness', where an actor's reasons for upholding a belief are merely *posthoc* rationalizations which have nothing to do with what induced him to adopt this belief in the first place, actually exist. However, the ubiquitous presence of false consciousness should not be an *a priori* presupposition of historiographical practice. That false consciousness was operative should be the outcome and not the starting point of a historical reconstruction.

Some proponents of the strong program could grant that contextual reasons constrain and guide scientific behavior and still argue that reasons themselves are not self-explanatory and should, therefore, be explananda of the sociology of knowledge.[30] The premiss of this argument seems correct. The reasons invoked by an actor are intimately connected with the cognitive values of the community to which the actor belongs and, no doubt, one can always ask why a specific community espoused a particular set of values. It does not follow, however, that such an explanation *must* be carried out in sociological terms. It might be the case, for example, that a 'biological' explanation would be more pertinent and that one would be able to explain the dominance of the values in question by showing how they augmented the survival capacity of the community. Furthermore, the historian always has the option to say that the predominance of the values in question is merely a brute fact, and may or may not have an explanation in terms of underlying social, or psychological, or biological factors.[31] Finally, to the extent that the cognitive values associated with the scientific enterprise have been stable throughout the development of science and have been shared by different scientific communities,[32] we have reasons to believe that sociological explanations of their continuing predominance would not be met with success. Explanations of this kind are inevitably tied to local characteristics of a specific community and are, therefore, not applicable to a phenomenon which transcends the boundaries of particular communities, unless one locates sociological factors which are common among different communities and demonstrates that those factors brought about the phenomenon in question – an admittedly daunting task.

If proponents of the strong program were to acknowledge the significance of reasons that historical actors would give for their beliefs, choices, etc. for the historian's explanatory task, they could not be charged with being ahistorical. To the extent, however, that they have chosen to stay close to the original spirit of their program and adhere to the symmetry principle they are led to the same ahistorical predicament that characterized Lakatosian historiography. Remember that for Lakatos the reasons that historical actors themselves would give in support of their decisions, were immaterial for the purposes of internal historiography. In the same way it is an *a priori* thesis of the strong program that these reasons are of little use to historians who aim at reconstructing past scientific practice. These reasons, according to the strong program, are not the 'real' causes of the actors' behavior, which instead must be explained by appealing to underlying social factors. In the same way that Lakatosian scientists were infallible apostles of the *MSRP* the actions of 'strongly programmed' scientists are mere epiphenomena of underlying social realities. *A priori* sociology has replaced *a priori* methodology as a guide of historiographical practice. There is an important difference, however, between the two historiographical approaches. Whereas Lakatos, at least in some of his moods, recognized that his rational reconstructions were philosophical fairy tales, 'strong programmers' insist that their sociological reconstructions are fully realistic.[33]

## IV. CONCLUDING REMARKS

I have argued that both Lakatos's theory of scientific rationality and the strong program's sociological theory of scientific practice lead to a similar ahistorical predicament, since their 'application' to history of science entails a dismissal of the explanatory value of the reasons that historical actors would give in support of their beliefs. This was the negative message of my paper; both reconstructions, to the extent that they are ahistorical, have not much to offer to historiographical practice.

Despite my criticism, there are some valuable elements in both the rational and the sociological approaches. One of the most interesting aspects of Lakatos's *MSRP* was that it tried to capture the internal dynamic of the Popperian world of objective knowledge. Popper introduced a distinction between three different worlds. In Lakatos's words, "The 'first world' is that of matter, the 'second' the world of feelings, beliefs, consciousness, the 'third' the world of objective knowledge, articulated in propositions."[34] It is a hard and unresolved problem whether the third world is entirely reducible to the second. If not, then its internal dynamic might transcend the beliefs, abilities, and wishes of human actors and act as a constraint on the development of scientific knowledge. Scientists might be confronted with problem situations that admit of a limited range of solutions, regardless of the abilities and goals of specific human actors. For instance, one could argue that after the development of Maxwell's electromagnetic theory a tension arose between electrodynamics and mechanics which constrained, albeit not determined, the further development of physics. Although Lakatos's attempt to capture this internal dynamic has not succeeded he has brought to our attention a subject which is ripe for exploration.

The value of the sociological analysis of scientific practice should also not be underestimated. For example, studying the wider cultural and social context of scientific activity can provide an understanding of the conceptual resources upon which the scientists draw for furthering their analyses. Furthermore, a case can be made for a more moderate version of social constructionism. The word 'social' encompasses not only macro-social factors, such as the wider social and cultural milieu, but also the micro-social realm, the social interactions between the members of the scientific community. Even if epistemic considerations of empirical adequacy, consistency, scope, simplicity, and fruitfulness (all these values which, as Kuhn himself has argued, are essentially involved in theory-choice) sufficiently constrain the generation and acceptance of scientific knowledge, their bearing on theory construction and theory choice is eventually decided by the relevant scientific community. Although such considerations play a crucial role in the establishment of consensus within the scientific community, their relative weight is subject to 'social negotiations'. The decision as to their appropriate weight requires such negotiations, since the relative significance of each epistemic criterion is not unambiguously specified.

Micro-social processes are also essential in understanding experimental practice, an aspect of science which has been until relatively recently ignored. Logical positivism presupposed the neutrality and unproblematic status of observational data. Hanson, Kuhn, and Feyerabend among others stressed the 'theory-ladenness' of observation and undermined its privileged status in empiricist epistemology. For those authors, however, the problematic status of data was a consequence of the theory-ladenness of perception and of the fact that observational reports are couched in theoretically contaminated language. The various judgements involved in the experimenter's decision to refine and conclude a particular experiment were not perceived as a potential threat to the validity of experimental results. Not surprisingly, the judgmental aspects of experimentation reinforced the social constructionists' scepticism about the validity of scientific findings. In its most radical version social constructionism maintains that the constraints of nature on the products of scientific activity are minimal and that the historical and sociological study of science should dispense with 'nature' as an explanatory ingredient in the generation and acceptance of scientific knowledge.[35] Data are selected or even constructed in a process which, if we believe the social constructionists, reflects the social interactions within the relevant scientific community. Scientific discoveries are not only or primarily a matter of finding laws or theories which account for the data, but also a matter of selecting or constructing data themselves.[36]

The social interactions and the various 'negotiations' which take place in the scientific community over the validity of experimental findings are undoubtedly an important and relatively neglected area of study. However, I suspect that far from implying scepticism, the social nature of experimental activity should be viewed as one of its main strengths. After all, as an outcome of this activity, "experimental conclusions have a stubbornness not easily cancelled by theory change"[37] and "experimental phenomena persist even while theories about them undergo revolutions."[38] The study of the micro-social aspects of various scientific communities can, thus, enhance our understanding of the judgmental aspects of experimental practice, as well as of the processes of theoretical decision-making.

In conclusion, any attempt to reduce the historiographical enterprise to rational or sociological reconstruction is doomed to fail. Only a metatheoretical account of science which would incorporate the intellectual, social, and material constraints on scientific practice would be a promising historiographical tool.

ACKNOWLEDGEMENTS

I am indebted to Nancy Nersessian for her suggestions which improved substantially both the content and the style of this paper. I have profited as well from comments by Gerald Geison and Norton Wise.

*Princeton University*

## NOTES

[1] In the afterword to the 2nd ed. of his *Black-Body Theory and the Quantum Discontinuity, 1894–1912* Kuhn declared that "I do my best, for urgent reasons, not to think in these terms [paradigms, revolutions, gestalt switches, and incommensurability] when I do history, and I avoid the corresponding vocabulary when presenting my results. It is too easy to constrain historical evidence within a predetermined mold." (T. S. Kuhn, 1987, p. 363) The same feeling, as Gary Gutting correctly pointed out, is shared by the majority of historians of science who "are not currently very interested in general interpretative schemata, which may have a Procrustean effect on their efforts to understand specific episodes." (G. Gutting, 1980, p. 3)

[2] Lack of space prohibits me from extending my analysis beyond the theoretical pronouncements of the two schools, even though their followers have produced a considerable amount of historical case studies. For a collection of rationally reconstructed historical episodes see C. Howson, 1976; and for a comprehensive bibliography of case studies which either support or were directly influenced from a sociological approach to scientific knowledge see S. Shapin, 1982. For a more up to date treatment of the latter see J. Golinski, 1990.

[3] I. Lakatos, 1970.

[4] "The act of discovery escapes logical analysis ... But it is not the logician's task to account for scientific discoveries ... logic is concerned only with the context of justification." H. Reichenbach, 1951, p. 231. Incidentally, Reichenbach's view was shared by Karl Popper: "the act of conceiving or inventing a theory, seems to me neither to call for logical analysis nor to be susceptible of it. ... the logical analysis of scientific knowledge ... is concerned ... only with questions of *justification or validity*". K. Popper, 1965, p. 31.

[5] For an excellent and detailed treatment of Lakatos's philosophy of science see I. Hacking, 1981.

[6] I. Lakatos, 1971, p. 99.

[7] J. Agassi, 1963.

[8] I. Lakatos, 1971, p. 106. For a discussion and critique of Lakatos's idiosyncratic conception of internal history see T. S. Kuhn, 1971, pp. 140–41 and I. Hacking, 1981, pp 138–39.

[9] I. Lakatos, 1971, p. 102.

[10] I. Lakatos, 1971, p. 92.

[11] One of the aims of his "History of Science and its Rational Reconstructions" was "to explain *how* the historiography of science should learn from the philosophy of science" and in particular from the *MSRP*. I. Lakatos, 1971, p. 91.

[12] Cf. E. McMullin, 1984, p. 137.

[13] T. S. Kuhn, 1980, p. 188. For the papers in question see C. Howson, 1976.

[14] Of course this point has been made before. Bertrand Russell, for example, maintained that "In studying a philosopher, the right attitude is neither reverence nor contempt, but first a kind of hypothetical sympathy, until it is possible to know what it feels like to believe in his theories." (B. Russell, 1945, p. 39; cited with approval in T. S. Kuhn, 1977, p. 149)

[15] For an insightful discussion of epistemic factors and their significance for historical explanation see E. McMullin, 1984, esp. pp. 129–136. For this and many other points I am indebted to McMullin's analysis.

[16] See, for instance, R. K. Merton, 1973; and J. Ben-David, 1984.
[17] T. S. Kuhn, 1970a, p. 237.
[18] The tenets of the 'strong program' were initially articulated in David Bloor's *Knowledge and Social Imagery* (D. Bloor, 1976). The focus of my critical remarks will be on this particular formulation of a wider movement in science studies, known as social constructionism. Occasionally, however, I will refer to authors who do not, strictly speaking, belong to the Edinburgh School but who share a similar, and sometimes more radical, perspective.
[19] Both Mannheim and Durkheim, however, did not believe that sociological analysis could provide any insight into the formation and consolidation of scientific beliefs. See J. R. Brown, 1984b, pp 3–6; and S. Woolgar, 1988, pp. 23–24.
[20] D. Bloor, 1976, p. 7.
[21] For a discussion of whether the reflexivity principle poses a threat to the viability of the strong program see D. Bloor, 1976, pp. 17–18; J. R. Brown, 1989, pp. 41–45; and M. Hesse, 1980, pp. 42–45.
[22] D. Bloor, 1976, p. 7.
[23] See, for instance, L. Laudan, 1984; and D. Bloor, 1984.
[24] Note that I have deliberately excluded the dichotomy between true and false beliefs because epistemic factors could very well lead to the adoption of beliefs which eventually might turn out to be false. Thus, both critics and proponents of the strong program agree that this dichotomy is irrelevant for the purposes of historical explanation. Cf. L. Laudan, 1984, pp. 56–57.
[25] B. Barnes and D. Bloor, 1982, p. 23.
[26] D. Bloor, 1984a, p. 80.
[27] T. McCarthy, 1988, p. 90.
[28] Some hints for such an attempt can be found in E. McMullin, 1984, pp. 154–155.
[29] See B. Barnes and D. Bloor, 1982, pp. 27–28.
[30] This is indeed Barnes and Bloor's position. See B. Barnes and D. Bloor, 1982, pp. 28–29; and D. Bloor, 1984b. This position is, however, at odds with the initial formulation of the strong program and, in particular, with the symmetry principle. Explanations in terms of contextual reasons are not of the same kind as explanations in terms of professional interests, class membership and other such social parameters. Moreover, this retreat to 'reasonableness' is not endorsed by all social constructionists. Steve Woolgar, for instance, claims that "SSS [Social Study of Science] favours the conception of rules as a *posthoc* rationalization of scientific practice rather than as a set of procedures which determine scientific action." S. Woolgar, 1988, pp. 17–18.
[31] I have paraphrased here one of van Fraassen's arguments against scientific realism. See van Fraassen, 1980, p. 24.
[32] As Kuhn remarked "such values as accuracy, scope, and fruitfulness are permanent attributes of science." (T. S. Kuhn, 1977, p. 335).
[33] I should point out, however, that the historical reconstructions that have been carried out by historians who were influenced by the strong program have not been characterized by the deliberate falsification of the historical record that was a standard feature of Lakatos's case studies.
[34] I. Lakatos, 1971, p. 127. For a very helpful discussion on the autonomy of the third world see I. Hacking, 1981, pp. 136–38.

[35] Harry Collins, for instance, claims that "the natural world has a *small or nonexistent* role in the construction of scientific knowledge." Cited in M. Hesse, 1988, p. 105.
[36] The classic in this genre is B. Latour and S. Woolgar, 1986.
[37] P. Galison, 1987, p. 259.
[38] I. Hacking, 1984, p. 172.

## REFERENCES


Agassi, J., 1963, *Towards an Historiography of Science*, History and Theory, vol. 2, The Hague: Mouton.

B. Barnes and D. Bloor, 1982, "Relativism, Rationalism and the Sociology of Knowledge", in M. Hollis and S. Lukes, eds., *Rationality and Relativism*, Cambridge, MA: The MIT Press, pp. 21–47.

Ben-David, J., 1971, *The Scientist's Role in Society*, Chicago: The University of Chicago Press. (Second edition, enlarged, 1984).

Bloor, D., 1976, *Knowledge and Social Imagery*, Chicago: The University of Chicago Press. (Second edition, enlarged, 1991).

Bloor, D., 1984a, "The Strengths of the Strong Programme", in J. R. Brown, 1984a, pp. 75–94.

Bloor, D., 1984b, "The Sociology of Reasons: Or Why 'Epistemic Factors' are Really 'Social Factors' ", in J. R. Brown, 1984a, pp. 295–324.

Brown, J. R., ed., 1984a, *Scientific Rationality: The Sociological Turn*, Dordrecht: Reidel.

Brown, J. R., 1984b, "Introduction: The Sociological Turn", in J. R. Brown, 1984a, pp. 3–40.

Brown, J. R., 1989, *The Rational and the Social*, London and New York: Routledge.

Galison, P., 1987, *How Experiments End*, Chicago: The University of Chicago Press.

Golinski, J., 1990, "The Theory of Practice and the Practice of Theory: Sociological Approaches in the History of Science", *Isis* **81**, pp. 492–505.

Gutting, G., ed., 1980, *Paradigms and Revolutions*, Notre Dame: University of Notre Dame Press.

Hacking, I., 1981, "Lakatos's Philosophy of Science", in I. Hacking, ed., *Scientific Revolutions*, Oxford: Oxford University Press, pp. 128–143.

Hacking, I., 1984, "Experimentation and Scientific Realism", in J. Leplin, ed., Scientific Realism, Berkeley and Los Angeles: University of California Press, pp. 154–172.

Hesse, M., 1980, "The Strong Thesis of Sociology of Science", in her *Revolutions and Reconstructions in the Philosophy of Science*, Bloomington and London: Indiana University Press, pp. 29–60.

Hesse, M., 1988, "Socializing Epistemology", in E. McMullin, 1988, pp. 97–122.

Howson, C., ed., 1976, *Method and Appraisal in the Physical Sciences*, Cambridge: Cambridge University Press.

Kuhn, T. S., 1970a, "Reflections on my Critics", in I. Lakatos and A. Musgrave, 1970, pp. 231–278.

Kuhn, T. S., 1970b, *The Structure of Scientific Revolutions*, Chicago: The University of Chicago Press. (Second edition).

Kuhn, T. S., 1971, "Notes on Lakatos", in R. C. Buck and R. S. Cohen, eds., *PSA 1970, Boston Studies in the Philosophy of Science*, **8**, pp. 137–146, Dordrecht: Reidel.

Kuhn, T. S., 1977, *The Essential Tension*, Chicago: The University of Chicago Press.
Kuhn, T. S., 1980, "The Halt and the Blind: Philosophy and History of Science", *British Journal for the Philosophy of Science* **31**, pp. 181–192.
Kuhn, T. S., 1987, *Black-Body Theory and the Quantum Discontinuity* 1894–1912, Chicago: The University of Chicago Press. (Second editon, enlarged).
Lakatos, I., 1970, "Falsification and the Methodology of Scientific Research Programmes", in I. Lakatos and A. Musgrave, 1970, pp. 91–196.
Lakatos, I., "History of Science and its Rational Reconstructions", in R. C. Buck and R. S. Cohen, eds., *PSA 1970, Boston Studies in the Philosophy of Science*, **8**, pp. 91–136, Dordrecht: Reidel.
Lakatos, I., Musgrave, A., 1970, eds., *Criticism and the Growth of Knowledge*, Cambridge: Cambridge University Press.
Latour, B. and Woolgar, S., 1986, *Laboratory Life: The Construction of Scientific Facts*, Princeton: Princeton University Press. (Second edition, enlarged).
Laudan, L., "The Pseudo-Science of Science?", in J. R. Brown, 1984a, pp. 41–73.
McCarthy, T., 1988, "Scientific Rationality and the 'Strong Program' in the Sociology of Knowledge", in E. McMullin, 1988, pp. 75-95.
McMullin, E., 1984, "The Rational and the Social in the History of Science", in J. R. Brown, 1984a, pp. 127–163.
McMullin, E., 1988, ed., *Construction and Constraint: The Shaping of Scientific Rationality*, Notre Dame: Notre Dame University Press.
Merton, R. K., 1973, *The Sociology of Science: Theoretical and Empirical Investigations*, Chicago: The University of Chicago Press.
Popper, K. R., 1965, *The Logic of Scientific Discovery*, New York: Harper & Row.
Reichenbach, H., 1951, *The Rise of Scientific Philosophy*, Berkeley and Los Angeles: University of California Press.
Russell, B., 1945, *A History of Western Philosophy*, New York: Simon & Schuster.
Shapin, S., 1982, "History of Science and its Sociological Reconstructions", *History of Science* **20**, pp. 157–211.
Van Fraassen, B. C., 1980, *The Scientific Image*, Oxford: Oxford University Press.
Woolgar, S., 1988, *Science: The Very Idea*, London and Chichester: Tavistock Publications and Ellis Horwood Limited.

ELENA A. MAMCHUR

# SOCIOCULTURAL FACTORS AND HISTORIOGRAPHY OF SCIENCE

The question of the role of sociocultural factors in the evolution of scientific knowledge is one of the most important in the historiography of science. Should or should not an historian of science take account of the sociocultural context in which scientific knowledge developed? Is he right in limiting himself to the cognitive factors?

A superficial observer may dismiss these questions as idle. Indeed, scientific cognition is social by its very nature. It is based on a system of already-obtained knowledge and uses the language and devices, that is, the results of the previous generations' theoretical and practical activity, or the results obtained by his contemporaries.

Moreover, the sociocultural context influences scientific cognition in another, more profound, sense: Through the ideals of scientific knowledge, styles of thinking and world perception, the social and cultural factors are involved in scientific quest and exert their influence on the content of scientific ideas.

However, the historiography of science deals with a theoretical reconstruction of the history of science rather than its empirical history. The fact that social and cultural factors are included in the cognitive process does not necessarily entail their presence in the process of theoretical reconstruction.

To give a biological analogy, one may say that a living organism is functioning and developing in close interaction with the environment. Moreover, the environment may cause individual changes in this organism. However, not all of these changes will be inherited or result in a change of species.

The same can be said of scientific knowledge. Despite the obvious impact of the sociocultural environment on the process of cognition, the role of sociocultural factors in this process is still undetermined.

Currently there are two trends in the historiography of science – rationalism and sociologism – that provide different answers to this question. Rationalists believe that the history of scientific knowledge requires only cognitive factors for its reconstruction; social factors should



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 193-201.

be employed to explain its external side. Supporters of sociologism argue that no adequate theoretical reconstruction of the history of scientific knowledge is possible outside the sociocultural factors. They believe that such factors are the factors of the *evolution* of scientific knowledge in the Darwinian sense of the word.

To support their arguments they point to a certain parallelism in the contents of scientific theories and intellectual trends of the sociocultural environment. Such parallels occur rather frequently in the history of science, especially in the periods when new paradigms of thinking emerge. In a widely known article P. Forman cites one of the most typical examples of this [1].

While tracing the history of quantum mechanics in Germany in the Twenties, Forman demonstrated that indeterminism of the quantum theory was totally compatible with the indeterminist views then reigning in the German physicists' intellectual milieu. The existentialist "philosophy of life" and Spengler's philosophy were the most powerful vehicles of such ideas, while Spengler's *The Decline of the West* produced an enormous impact on intellectual life in post-war Germany.

L. Feuer cites other examples of parallelism [2]. His analysis of the genesis of Einstein's special relativity theory revealed parallel ideas in art, music, and literature. And his studies of the genesis of Bohr's complementarity conception, the basis of quantum theory, revealed similar phenomena.

What does parallelism point to? Rationalists believe it is a result of random coincidences; their opponents insist it is a manifestation of the sociocultural context influencing scientific knowledge. For example, in the case under study Forman asserts that intellectual ideas and sentiments predetermined indeterminist ideas in physics. In this way, they pushed the system of physical knowledge further. He writes:

> The scientific context and content, the form and level of exposition, the social occasions and the chosen vehicles for publication of manifestoes against causality, all point inescapably to the conclusion that substantive problems in atomic physics played only a secondary role in the genesis of this acausal persuasion, that the most important factor was the social-intellectual pressure exerted upon physicists as members of the German academic community ([1], p. 109).

He insists that "the program of dispensing with causality in physics... achieved a very substantial following among German physicists *before* it was 'justified' by the advent of a fundamentally acausal quantum mechanics" ([1], p. 110). From Forman's point of view, changes in the intellectual environment *predated* changes in scientific knowledge's conceptual context and *brought* them about.

Is it true that social factors were as important in the emergence of the quantum theory as Forman seems to believe? Hendry [7] criticized Forman's stance. He demonstrated that conclusions were inconvincing and that, at best, one could recognize that social factors contribute to quantum mechanics together with the cognitive factors. They were *one more*, not the *main* driving force.

Hendry seems to be right in this special case. But I believe that the main problem is more general than the question of any particular case-study correctness. Postulating that sociocultural factors may cause considerable modifications in the system of scientific knowledge and that they are main driving force somewhat simplifies the real relationships between scientific knowledge and the sociocultural milieu. These relationships are much more complex. There is no one-sided impact of the social context on scientific knowledge, but rather a mutual influence, a multistage strengthening of the intellectual tendencies of the sociocultural environment, on the one hand, and of the system of scientific ideas, on the other. Very often it is hard to establish what is the cause and what is the result, what was first and what came after. This is especially difficult in the case of a new scientific paradigm emergence, a phenomenon of paramount importance.

To illustrate I shall turn to the history of Galilean–Newtonian physics, isolating only one aspect, namely, the changes in the status and nature of ideas about space. Typologically close changes also occurred in other cultural spheres in painting, sculpture and architecture. This enables me to have a closer look at the parallelism of ideas.

What changed in this respect in physical knowledge? The main changes occurred when the concept of space developed from the Aristotelian interpretation of it as a "place" to the conception of the infinite homogeneous geometrical space of Galilean–Newtonian physics. It was an important intellectual advance without which the principle of intertia – the privat of Galilean-Newtonian physics – could not have been formulated. The idea of force as a cause of movement (Aristotle) ceded "place" to a picture of the world in which a body was indefinitely in a state of straight-line uniform motion, as long as no forces were applied to it (Newton).

New ideas about motion required of the conception of a void homogeneous and infinite space. This concept and its assimilation by the academic community was a true revolution in science.

It was irrelevant that Aristotle used the term "place" to describe space and never used the term "space". In turn, Newton used the same word to designate the same idea. It should be said that Aristotle and Newton used similar terms to describe ideas of space. This is a trap for careless historians of science who trend to judge conceptions by their names rather than their essence.

Indeed both Newton and Aristotle described space as a "place"; both insisted that "place" existed together with bodies, both described motions as change of place. Still, there is a deeply dividing difference between what Aristotle and Newton understood by space. This is related to the idea that place can be separated from bodies. Differing from Newtonian physics, in Aristotelian physics "places" and bodies cannot be separated.

It is totally irrelevant that Aristotle used such expressions as "place can be left behind by the thing and is separable" ([3], p. 290) or "the place of a thing is neither part nor a state of it but is separable from it" ([3], p. 290). Aristotle used the words "leaved", "separable" and "side by side" in a sense that differed from Newtonian physics. "Place" for Aristotle was a body's limit. What is more, this was not a body that was being discussed but the body that embraced and contained it. Where there is no containing body there is no "place". In this way, according to Aristotelian physics, if all bodies embracing the given body are removed there is nothing left. In Newtonian physics there is space left. It is not fortuitous that Aristotelian cosmos was immovable – there was nothing to embrace it, hence, it had no place by the changing of which it could move. According to Aristotle, motion was a consecutive change of bodies that embraced the given body, while in Newtonian physics motion occurred as related to space.

Aristotle was aware of aberrations caused by an attempt to grasp his conception of space ("place"):

> Place is thought to be something important and hard to grasp, both because the matter and the shape presents themselves along with it, and because the displacement of the body that is moved takes place in a stationary container, for it seems possible that there should be an interval which is other than the bodies which are moved. The air, too, which is thought to be incorporeal, contributes something to the belief: it is not only the boundaries of the vessel which seem to be place, but also what is between them, regarded as empty ([3], p. 291).

To prevent these aberrations he specifies: "...the innermost motionless boundary of what contains is place" ([3], p. 291).

The transition to new physics occurred when the idea of "place" as a boundary of the containing body gave way to the conception of "place" as something that is left behind after all the bodies that filled it are totally removed from it.[1]

Today it is rather hard to assume the Aristotelian point of view and imagine space in the way it was perceived in antiquity: We are in the grip of Galilean–Newtonean physics. It would have been equally impossible for an ancient scientist to visualise space in the modern way. Not only scientific but also artistic consciousness of antiquity saw space in the same way – it was not presented in ancient paintings. Spengler, in his time, pointed out this: "The Classical relief is strictly stereometrically superimposed on plane, and there is an interplace between the figures but no depth" ([4], p. 184).

The spatial ideas of antiquity can best be seen in ancient sculpture. In certain respects Greek sculpture was unique; as opposed to statues in Gothic cathedrals that were placed in niches, Greek sculptures were placed to be seen from all sides. This is quite natural when seen in the context of Classical spatial ideas. Spengler had the following to say on this score: "Choosing of a certain position of statue, made in order to create a necessary impression, would mean introduction of definite space relations between spectators and the work of art in the language of the latter's form" ([4], p. 240).

Mediaeval painting also ignored space – beginning with early Christianity, painters consciously limited themselves to the foreground. This device survived the Romanesque style and extended to the Gothic. Mediaeval painting did not use linear perspective. Boris Raushenbach, a Soviet researcher, has demonstrated that axonometry served the cornerstone of Byzantine (and Old Russian) art ([5], pp. 384–416). As distinct from linear perspective, where parallel lines meet at the horizon, in axonometry such lines retain their parallelism in perspective geometry. Not infrequently, axonometry was disrupted by the so-called reverse perspective, with parallel lines diverging at the horizon and meeting at the side of the spectator. When confronted with the task of presenting depth in their pictures, mediaeval painters reverted to different methods more than their colleagues of the New Times. In the Renaissance, distance was designated by objects' smaller sizes. In the Middle Ages, the same task was resolved by placing distant objects above the horizon.

Similar changes could be observed in other cultural fields. Spengler wrote that starting with the 13th and 14th centuries, unified and infinite geometrical space was the key principle shared by all arts. Classical sculpture that symbolised the body gave way to oil painting, the most important elements of which were linear perspective and spatial relations between objects. He argued that "in the actual picture there is transvaluation of all the elements. The background, hitherto casually put, is regarded as a fill-up and, as space, almost shuffled out of sight, gains a preponderant importance..." ([4], p. 239). The horizon as a symbol of the eternal and infinite universe appeared in painting. "...There emerges in the picture the great symbol of an unlimited space-universe which comprises the individual things within itself as incidentals – the horizon" ([4], p. 239). Clouds, used to produce an effect of distance, became similarly symbolic.

A style in oil painting where outlines reigned was replaced with "chiaroscuro" that created representations. Starting with the late 14th century and reaching its peak with the Impressionists' painting, chiaroscuro shed the outlines that limited objects. Its aim was the feeling of distance and perspective.

Elongated ponds, walks, alleys, perspectives and galleries became popular in garden architecture serving the same purpose as perspective in oil painting.

Similar intellectual tendencies were present in architecture. The sophisticated skeletal constructions of Gothic cathedrals that replaced Romanesque architecture allowed architects to overcome the bulky Romanesque shapes, to lighten walls and vaults and create a dynamic unity of spatial cells. Infinite space is born by the contrast between the high and light central nave and darkened side naves typical of Gothic architecture. This effect was emphasised with stained-glass windows.

In music, the victorious polyphony and purely instrumental forms, free of bodily associations, pushed back the traditional arrangements dominated by texts and human voice.

How can the nature and mechanisms of similar parallel developments be explained? There is no lack of explanations, ranging from the opinion that in Classical Antiquity and the Middle Ages painters had not yet learned how to draw[2] to the states, that mediaeval paintings' religious content and aim explains the specific treatment of perspective in them.[3]

While not rejecting either of these factors outright, many researchers believe that neither can be regarded as all-important. Each provides but a partial explanation and leaves many parallels in Classical and Mediaeval culture unexplained, such as flat frescos, round Greek statues and specific spatial ideas and conceptions in the science of Classical Antiquity and the Middle Ages.

For my part, I believe in the following explanation as the most plausible: Not a single one of the above-mentioned cultural parallels is a cause or foundation for the rest. They were more or less synchronous and mutually determined; their common sources should be sought in human activity specific to that period. It is natural to suggest that painters of Classical Antiquity and the Middle Ages sought to paint the world as they saw it – and they saw no linear perspective.

There are numerous data (including the way vision develops in children) that testify that man today would continue to perceive the world around him in the similar way, had he not learned to use his previous experience to correct what he saw. The picture of the world is born in man's mind – vision is created by the joint effort of the mind and eyes. The brain works hard to transform the representation on the retina, it mobilises man's previous experience. Unlike children in Classical Antiquity and the Middle Ages, contemporary children see representations (photography, paintings, cinema and so on) that use linear perspective. This decides their perception as different from that of mediaeval people.

Galilean–Newtonian physics made its contribution to this process: today, it continues to shape the vision of schoolchildren. It is hard to believe, however, that the development of physics resulted in perspective being employed in painting and other graphic arts. It is equally improbable that spatial ideas in science were changed under the impact of Renaissance painting. It is impossible to isolate one of the three phenomena (science-art-technology) as genetically predating the others. They were synchronous and genetically unconnected.

Since causal determination presupposes that one phenomenon predates and gives birth to others, one has to recognise that there is no causal relation in this case.

In order to reconstruct an interrelationship of the mentioned phenomena one has to take into account that the system of scientific knowledge is a subsystem of the system of culture. The most appropriate type of explanation of a system behavior is the so-called functional model. Its main point is that in a complicated dynamic system, composed of different subsystems acting one within another, the behavior of each can be explained by the necessity of preservation of the whole system ([8], pp. 165–176). In the framework of the functional model phenomenon of ideas, parallelism should be explained by the functioning of science and other subsystems of culture in a wider context: They behave so that the whole system of culture can be preserved and provided with an optimal regime of functioning.[4]

It does not mean that causal explanation cannot be used in a reconstruction of the above-mentioned parallelism. The system of culture is very complicated and many-sided, and the theoretical reconstruction of the interrelations of all its parts cannot be reduced only to the functional model. At times the functional model should be supplemented by causal and other explanational schemes.

Moreover, giving a certain interpretation to the notion "cause", one can agree with those who believe that an explanation of the parallelism of ideas by all means requires the search for a cause.

In our time, cause is interpreted as something making an effect, providing a result. But such interpretation did not always exist. In antiquity, four types of causes (efficacious, material, formal and final) were admitted. Only one of them (efficacious) was close to the modern sense of the word, and only to some extent. As was noted by Heidegger [10] the Greeks did not comprehend cause as an action. For them cause was something that was responsible for the appearance of something else. The four kinds of causes were interpreted as four types of responsibility.

From the point of view of antiquity, in the appearance of an object made by a human being, equally responsible are the material from which it is made, the form of the object, the master who made it and its destination. All of these are necessary, with the loss of one the object would not appear. If antiquity's view on cause were admitted, it would be possible to state that sociocultural factors are one of the causes of a new paradigm's appearance. Without giving birth to the new knowledge they bear a part of the responsibility for its genesis.

*Institute of Philosophy, Russian Academy of Sciences*

## NOTES

[1] The treatment of the idea of "place" turned out to be closely connected with the problem of void (Aristotelian physics rejected the idea of "vacuum" both inside and outside the world). The discussions on the possible existence of void that were raging in the 14th century among theologians made it easier to switch to the idea of an empty homogeneous space. In 1277, a Franciscan tribunal headed by Etienne Tempier ruled that, contrary to what the great Stagirite had asserted, the system of celestial bodies could have been set in motion through a certain straight-line motion. (The very idea of this motion was treated by Aristotelians as

an absurdity due to an absence of void.) Clothed in scholastic rather than scientific robes, the edict was still a step towards new spatial ideas by opening the possibility to discuss the formerly taboo question due to the canonisation of the Aristotelian teaching.

[2] Art critics today agree that this explanation is ridiculous, to say the least. It is impossible to imagine that the nations that produced sculptural masterpieces failed to master the most primitive methods of drawing perspective.

[3] There is an opinion that varied scales and the flatness of Byzantine and Old Russian icons can be explained by the fact that the spectator was *aware* of the presence of a saint in them. The figures were oriented on the spectator and were maximally closed to him. This forced the icon painter to make the central figures bigger than the rest (see [6], p. 400).

[4] Functional explanation is not identical to a teleological one. Doubt can appear because in both cases the notion "goal" is used. However, in true teleological explanation, goal is a cause of phenomena. In the functional model the notion of goal is used in the sense of a final result of the system's development and not in the sense of a cause. In the functional model, one should speak rather of a teleonomical explanation when it is presupposed that the system behaves *as if* it had a goal ([9], p. 75).

## REFERENCES


1. Forman, P., "Weimar Culture, Causality and Quantum Theory, 1918 1927. Adaptation by German Physicists and Mathematicians to a Hostile Intellectual Environment". In: R. McCormmach (ed.), *Historical Studies in the Physical Sciences*, No. 3, Philadelphia, 1971.
2. Feuer, L. S., *Einstein and the Generations of Science.* New York, 1974.
3. *The Works of Aristotle.* William Benton, Publisher. Encyclopaedia Brittannica, INC Chicago, The University of Chicago. The Great Books, 1952, Vol. 1.
4. Spengler, O., *The Decline of the West.* Vol. 1, Authorised translation by Charles Francis Atkinson. L., Sydney, 1980.
5. Raushenbach, B. V., *Prostranstvennye postroenia v zhivoposi*, Moskva, 1980 (in Russian).
6. Saltykov, A., "O prostranstranstvennykh otnosheniyakh v viscantiiskoi i drevnerusskoi zivopisi". In: *Drevnerusskoe iskusstvo: Zarubezhnve svayzi*, Moskva, 1975 (in Russian).
7. Hendry, J., "Weimar Culture and Quantum Causality". In: *History of Science.* Vol. 18, 1980.
8. Nikitin, Ye., *Objasnenije-funkzija nauki.* Moskva, Nauka, 1970 (in Russian).
9. Cheklend, P., *Systems Thinking, Systems Practice.* John Wiley and Sons, Chichester, N.Y., Brisban, Toronto. 1981.
10. Heidegger, M., Die Frage nach der Technik. In: *Vorträge und Aufsätze.* Pfüllingen, 1959.

SABETAI UNGURU

# IS MATHEMATICS AHISTORICAL? AN ATTEMPT TO AN ANSWER MOTIVATED BY GREEK MATHEMATICS

"There is no religious denomination in which the misuse of metaphysical expressions has been responsible for so much sin as it has in mathematics".
Ludwig Wittgenstein, *Culture and Value.*

## I. In the *Nichomachean Ethics Aristotle said:*

... Wisdom is both knowledge (ἐπιστήμη) and intuitive intelligence (νοῦς) ... Prudence (φρόνησις) on the other hand is concerned with the affairs of men and with things that can be the object of deliberation. For we say that to deliberate well is the most characteristic function of the prudent man; but no one deliberates about things that cannot vary nor yet about things that are not a means to some end and that end a good attainable by action; and a good deliberator in general is a man who can arrive by calculation at the best of the goods attainable by man.

Nor is Prudence a knowledge of general principles only: it must also take account of particular facts, since it is concerned with action and action deals with particular things. This is why men who are ignorant of general principles are sometimes more successful in action than others who know them ... And Prudence is concerned with action, so one requires both forms of it, or indeed knowledge of particular facts even more than knowledge of general principles. ...

Prudence is indeed the same quality of mind as Politics (πολιτική), though their essence is different. ... Prudence also is commonly understood to mean especially that kind of wisdom which is concerned with oneself, the individual ... For people seek their own good and suppose that it is right to do so ... Moreover, even the proper conduct of one's own affairs is a difficult problem, and requires consideration.

Further evidence for this is furnished by the fact that the young may be geometers and mathematicians and be wise in such subjects, but they are not thought to be prudent. The reason is that Prudence (practical wisdom) is concerned with particular facts as well, which become known as the result of experience, while a young man cannot be experienced, for experience is the fruit of years. One might indeed further inquire, why a boy may be a good mathematician, but cannot be wise (σοφός) [i.e., a philosopher or a metaphysician] or a natural philosopher. May we not say, it is because the subjects of mathematics are reached by means of abstraction, while the principles of philosophy (metaphysics) and physics come from experience; and the young have no conviction of the latter, though they may speak of them, while in mathematics the principles are plain and free of ambiguity (οὐκ ἄδηλον)? Again, in deliberation there is a double possibility of error: you may go wrong either in your general principle or in your particular fact ...

Prudence then stands opposite to intelligence ... (ἀντίκειται μὲν δὴ τῷ νῷ (1141b2-1142a25)



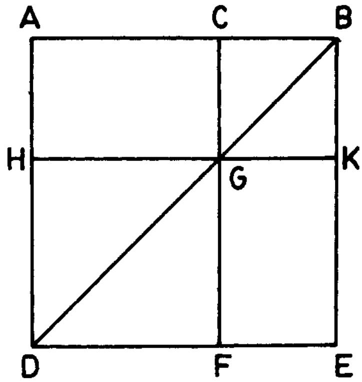
*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science,* 203-219.

This is a very interesting distinction, I think, and it serves me well as an appropriate beginning of my paper. Be that as it may, there is nothing in Aristotle's analysis to lead one to conclude that old men (experienced people) may not be imprudent, i.e., lack in practical wisdom. But the distinction between *Prudence* and *Intelligence* meshes smoothly with that other famous Aristotelian distinction in the *Poetics*, between *Poetry* and *History*, about which I will say a few more words shortly.

First, however, let me dispose of some necessary geometrical examples, upon which hinge both my criticisms of the traditional historiography of Greek mathematics and my alternative interpretation of ancient Greek mathematical texts. "Le bon Dieu est dans les détails."

II.4. *Elements:*

If a straight line be cut at random, the square on the whole is equal to the squares on the segments and twice the rectangle contained by the segments.[1]



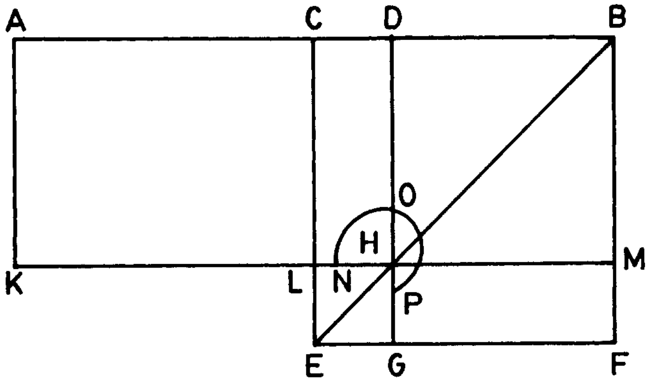

$$SqAB = sqAC + sqBC + 2\ \mathrm{Rect}\ (AC, CB)$$

The truth of the claim is obvious. And this is how Van der Waerden assesses this proposition:

II.4 entspricht der Formel $(a + b)^2 = a^2 + b^2 + 2ab$.[2]

II.5 *Elements:*

If a straight line be cut into equal and unequal segments, the rectangle contained by the unequal segments of the whole together with the square on the straight line between the points of section is equal to the square on the half.[3]




The claim is then:

$$\text{Rect }(AD, DB) + sqCD = sqBC$$

And here are the main steps of Euclid's proof:

$$\begin{array}{ll} & \text{Rect } CH = \text{Rect } HF \\ \therefore & \text{Rect } CM = \text{Rect } DF \\ & \text{Rect } CM = \text{Rect } AL \\ \therefore & \text{Rect } AL = \text{Rect } DF \\ \therefore & \text{Rect } AH = \text{Gnomon } NOP \\ \therefore & \text{Rect } AH + sqLG = sqCF, \text{ q.e.d.} \end{array}$$

Now, algebraically, if

$AB = a$ and
$BD = x$, then it is clear that
$ax - x^2 = \text{Rect } AH = \text{Gnomon } NOP$

If Gnomon $NOP = b^2$, then $ax - x^2 = b^2$. And this is how Heath assesses this proposition:

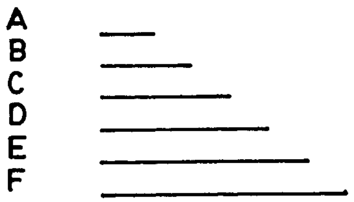
...the problem of solving the equation $ax - x^2 = b^2$ is, in the language of geometry, *To a given straight line (a) to apply a rectangle which shall be equal to a given square ($b^2$) and fall short by a square figure,* i.e., to construct the rectangle *AH* or the gnomon *NOP*.[4]

For my text example I need to say a few words about *mathematical induction*. The essence of *complete induction*, or of mathematical induction, is:

> If $P(1)$ and $P(n) \rightarrow P(n + 1)$ for all $n$ are both true (valid), then $P(n)$ is true (valid) for all $n$.

Now, it has been claimed (by Stamatis, Hans Freudenthal and others) that proofs by induction occur in the domain of Greek mathematics. Since I shall take issue with this claim, I should discuss some of the alleged instances of induction by means of one, or more, textual examples. Let us begin with IX. 8 *Elements*:

If any multitude of numbers, starting from an unit, be in continued proportion, then the third from the unit will be square, as will all those that leave out one successively; the fourth will be cube, as will all those that leave out two; and the seventh will be concomitantly cube and square, as will all those that leave out five.[5]




The numbers $A$, $B$, $C$, $D$, $E$ and $F$, in continued proportion starting from an unit, are given. The claim is that $B$, the third from the unit, is square, as are all those that leave out one successively ($D$ and $F$); also that $C$, the fourth from the unit, is cube, as are all those that leave out two ($F$); moreover, $F$, the seventh from the unit, is concomitantly cube and square, as are all those that leave out five.

*Proof*:

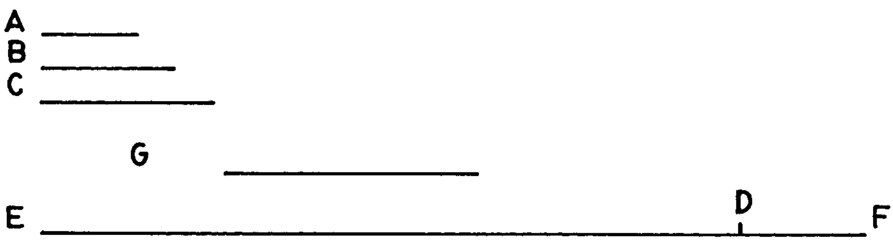
$1 : A :: A : B$, by definition of continued proportion. 1 measures $A$ the same number of times that $A$ measures $B$, by def. VII.20 ["Numbers are *proportional* when the first is the same multiple, or the same part, or the same parts, of the second that the third is of the fourth"].[6] But the unit measures $A$ according to the units in it; consequently, $A$ measures $B$ also according to the units in $A$, i.e., $A$ times $A$ equals $B$ ($A * A = B$) and so $B$

is square. Now, since $B$, $C$, $D$ are in continued proportion, i.e., $B : C :: C : D$, and $B$ is square, there follows from VIII.22 that $D$ is also square ["If three numbers be in continued proportion and the first be square, the third will also be square"].[7] By the same token, $F$ is also square. *In the same fashion it can be shown that all those numbers that leave out one are squares.*

What about $C$, the fourth from the unit? Since as the unit is to $A$, so is $B$ to $C$, i.e., $1 : A :: B : C$, the unit measures $A$ the same number of times that $B$ measures $C$. But the unit measures $A$ according to the units in $A$. Therefore $B$ also measures $C$ according to the units $A$, that is $A$ times $B$ equals $C(A * B = C)$. Now since $A * A = B$ and $A * B = C$, it is obvious that $C$ is cube. And because $C$, $D$, $E$ and $F$ are in continued proportion, i.e., $C : D :: D : E :: E : F$, and $C$ is cube, there follows, by VIII.23, that $F$ is also cube. [VIII.23: "If four numbers be in continued proportion and the first be cube, the fourth will also be cube"].[8] But it was already shown that $F$ is square also; consequently, the seventh from the unit is concomitantly square and cube. "Similarly we can prove that all the numbers which leave out five are concomitantly cube and square".[9] Q.E.D.

Finally, my last mathematical example, although there are others, perhaps weightier and more trenchant, but less appropriate in the present circumstances, is the famous proposition IX.20 claiming that
"Prime numbers are more than any assigned multitude of prime numbers."[10]

The beautiful proof is well known. I shall broach it here in order to make my points.



Let $A$, $B$, $C$ be the given prime numbers. The claim is that there are more primes than the given ones. Let the least number measured by $A$, $B$ and $C$ be taken and let it be $DE$. (It is clear that this number equals $A * B * C$.) Let the unit $DF$ be added to it. $EF$ is either prime or composite. If prime, the proof is completed. If composite, then by VII.31, it is measured by some prime number [VII.31: "Any composite number is measured by some prime number".][11] Let it be measured by $G$. $G$ is a *different* prime than any of the given primes $A$, $B$, $C$. Otherwise, since $A$, $B$, $C$ are such that each measures $DE$, $G$ will also measure $DE$. But $G$ measures $EF$, too. Consequently, since $G$ is a number, it will also measure the remainder, the unit $DF$, which is absurd. Hence, $G$ is different from any of the numbers $A$, $B$, $C$ and it is prime. It follows, therefore that there are more prime numbers than the given multitude of primes, $A$, $B$, $C$. Q.E.D.

So much by way of historical, mathematical examples. Let us return now to Aristotle. In the *Poetics*, Aristotle distinguishes between *historian* and *poet*, pointing out that the difference between them lies in the former telling us *what happened* while the latter is concentrating on *what might happen*. "For this reason," he goes on, "poetry is something more scientific and serious than history, because poetry tends to give general truths while history gives particular facts" (1451b 3).

This difference, I submit, is essentially also the difference between *mathematics* and *history*.

*History* is the study of the present traces of past events from the standpoint of change and the particular, the idiosyncratic. Although long-lasting structures, stable frameworks and durable, quasi-constant features are legitimate topics of historical investigation, they are not what makes history what it is. History is primarily, essentially interested in the event *qua* particular event, in the specific happening, in change from an identifiable, individual characteristic to another identifiable, individual characteristic. History is not (or is primarily not) striving to bunch events together, to crowd them under the same heading by draining them of their individualities. On the contrary, history is the attempt at understanding each past event in its own right. The domain of history, then, is the idiosyncratic.

The historian of ideas does not discharge his obligation by showing merely the extent to which past ideas are like modern ideas. His main effort should be in the direction of showing the extent to which past ideas were unlike modern ones, irrespective of the fact that they might (or might not) have led to modern ideas. This is a wise methodological tack, since it enables the historian to avoid reductive anachronism while channelling his historical empathy toward an understanding of the past in its own right. It is also wise to take the written documents of the past to mean precisely what they say, short of clear-cut proof to the contrary. There is no historical advantage whatever growing out of the gratuitous assumption that the men of old played tricks on us by systematically hiding their line of thought.

I shall not presume to define here what mathematics is, as that is best left to practicing mathematicians. Besides, there are plenty of definitions available, running the gamut from Bertrand Russel's "... mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we are saying is true"[12] to Nicolas Bourbaki's "A *mathematical theory* ... contains rules which allow us to assert that certain assemblies of signs are *terms* or relations of the theory and other rules which allow us to assert that certain assemblies are *theorems* of the theory".[13] Every reader can easily take his pick. But I can safely say what mathematics is not. It is certainly not history. The domain of mathematics is not the idiosyncratic, but in a very real sense, the nomothetic, since what mathematicians do is to show that from certain assumptions about as yet unidentified objects some conclusions about the same objects will follow necessarily, by rule.

The history of mathematics is history not mathematics. It is the study of the idiosyncratic aspects of the activity of mathematicians who themselves are engaged in the study of the nomothetic, that is, of what is the case by law. If one is to write the history of mathematics and not the mathematics of history, the writer must be careful not to substitute the nomothetic for the idiosyncratic, that is, not to deal with past mathematics as if mathematics had no past beyond trivial differences in the outward appearance of what is basically an unchangeable hard-core content.

It is high time that we return to the examples appearing above. What is wrong with the standard interpretation of these examples?

Briefly, in the case of II.4 and II.5, the *symbolic, algebraic* reading of those propositions. The kind of manipulative, abstract, generic, ideal symbolism involved in modern algebra was not available to the Greek mathematicians. Specifically, if you have at your disposal

$$(a + b)^2 = a^2 + b^2 + 2ab,$$

you *also* have at your disposal, as a bonus, a practically unlimited number of other identities obtainable mechanically from the given one by the simple procedures of elementary, algebraic transformations. Some of these additional identities simply do not make sense within the confines of Greek mathematics. Furthermore, saying that what really lies behind Euclid's geometrically couched statements are algebraic reasonings, appearing in geometrical garb because of the lack of an appropriate algebraic symbolism strikes me as both inaccurate and unilluminating. To see this, let us ask the following question: how illuminating would it be to propose that Euclid *really* thought in Sanskrit, but because of his ignorance of the Sanskrit alphabet, had to use the Greek one and consequently expressed himself in Greek? Greek mathematics must be understood in its own right. This can be done by refusing to apply to its analysis foreign, anachronistic criteria. The only acceptable meta-language for a historically sympathetic investigation and comprehension of Greek mathematics seems to be ordinary language, not algebra.

We are coming now to IX.8 and IX.20. Concerning the former, Hans Freudenthal said:

> In einer geometrischen Reihe $a$, $a^2$, $a^3$, ... sind $a^2$, $a^4$, $a^6$, ... Quadrate. Es findet ein echter (quasi-allgemeiner) Induktionsschluss statt, nämlich von $a^2$ auf $a^4$.[14]

This is ludicrous. Leaving aside the objectionable procedure of transcribing symbolically the rhetorical Euclidian proposition, to which I alluded earlier, I can only add here that Freudenthal's procedure, to my mind, actually disposes of the propositional, theorematic character of the proposition under discussion. As Heath pointed out, "The whole result is of course obvious, if the geometrical progression is written, with our notation, as 1, $a$, $a^2$, $a^3$, $a^4$, ... $a^n$."[15] Indeed!

What concerns us here, however, is to grasp that the mathematical thought-process displayed by the proof of proposition IX.8 is *not* an inductive mathematical process. *First* it is clear that Euclid did not, and could not, rely on the axiom of mathematical induction which was not yet invented and formulated. Hans Freudenthal himself acknowledges this. Its first formulator seems to be Blaise Pascal, who lived somewhat later than Euclid, in his *Traité du triangle arithmétique* written in 1654 and published posthumously in 1665. *Second*, not only is there no reliance in IX. 8 on the axiom of complete induction, but, *and this is the crucial consideration*, Euclid's grasp of what he is engaged in in the demonstration of the claim contained in the enunciation of IX.8 seems to me to furnish evidence that his understanding of his procedure is mathematically noninductive. Otherwise it is impossible to comprehend his explicit statement, *repeated twice*, in connexion with squares and cubes that "Similarly we can prove that all those which leave out one are square" and "Similarly we can prove that all the numbers which leave out five are also both cube and square".[16]

In other words, Euclid does not appear to have felt that he had given a demonstration covering *all numbers* leaving out one and *all* those leaving out five, but rather that he had shown how the proposition could be proved for any assemblage of numbers in continued proportion, the proof in the *Elements* being a token, an example for any future proofs. And this, to my way of thinking, is *not* an instance of a proof by mathematical induction, a method that does not stand in need of reinforcement and supplementation by an indefinite number of proofs with identical conceptual structure.

Mathematical induction, the substance of which consists in the inference from $n$ to $n + 1$, is not a faulty method of proof. On the contrary, it is flawless. Its conclusions are not illustrations or tokens and, as is well known, they are valid for all natural numbers. Needless to emphasize that the concept of "all natural numbers" is foreign to Greek mathematics.

Hence there is no escape from the conclusion that proposition IX.8 was not, *and could not have been*, proved by mathematical induction. (I shall say more about the second part of my conclusion later.) Let me merely point out at this juncture in my paper that all I said about IX.8 applies, *mutatis mutandis*, to IX.9 also. Its enunciation is:

> If as many numbers as we please beginning from an unit be in continued proportion, and the number after the unit be square, all the rest will also be square. And, if the number after the unit be cube, all the rest will also be cube.[17]

It is NOT proved by mathematical induction and it could not have been proved by this method.

What is the method of proof of IX.20? In spite of the fact that it actually uses only three given primes, it is clear that this is a quasi-general proof, wholly innocent of inductive considerations. Even Freudenthal himself acknowledges this: "Der Beweis ist quasi-allgemein ... Liegt hier ein Induktionsschluß vor? Natürlich nicht!" The conclusion is unavoidable, then, that Evangelos Stamatis' conclusion in his article *ὁ ἀναδρομικὸς συλλογισμὸς παρὰ τῷ Εὐκλείδῃ*, which appeared in 1953 in ΠΡΑΚΤΙΚΑ

ΤΗΣ ΑΚΑΔΗΜΙΑΣ ΑΘΗΝΩΝ, to the effect that IX.20 is an instance of demonstration by mathematical induction, is mistaken and that the source of this mistake resides, *inter alia*, in the treacherous transcription of the Euclidean proof by means of an historically unacceptable algebraic symbolism, coupled with the readiness to modify, more or less, Euclid's enunciation, steps enabling to shunt the proof onto inductive tracks. That such a sidetracking is indeed possible (though superfluous and pointless, I might say) is something that Freudenthal too asserts immediately after the previous citation from his article:

Wollte man aber den Satz beweisen 'zu jedem *n* gibt es eine *n*- te Primzahl', so käme man unter Verwendung des gerade Bewiesenen von selber zu einem Induktonsschluß.[18]

As I said this seems to me entirely superfluous and pointless. But this is not all, since Freudenthal goes on to say: "Soll man Euklid diesen Gedankengang zuschreiben? Ich neige zu einer bejahenden Antwort."[19]

And I, contrariwise, out of considerations that, *mutatis mutandis*, were already mentioned in connection with IX.8, "neige zu einer verneinenden Antwort"! Without recanting anything I said so far, I am willing to agree with Freudenthal that it is possible to enunciate anew the proposition *and* to modify accordingly the proof, in such a manner that the latter would be performed by induction. *But this is not history*! Marx already knew that though men make their own history, they do not, and cannot, make it in circumstances of their own choosing. And the circumstances of Greek mathematics simply do not permit the manipulations required to contrive mathematical induction within its confines.

Number for the Greeks, *arithmos*, is always determinate. It is a *limited multitude* (πλῆθος τὸ πεπερασμένον), a collection of units, or as Euclid puts it, "a multitude composed of units,"[20] multiplicity or manyness or plurality, being, as it were, the genus of number: 'Plurality is as it were the class to which number belongs; for number is plurality measurable by one, and one and number are in a sense opposed, for inasmuch as one is measure and the other measurable, they are opposed."[21]

Number is always a definite number of definite things, even when the things are not *aisthetá*, objects of sense, like dogs, sheep and horses but pure units, graspable only by thought. The defining feature of *arithmos* as a "definite number of" remains intact in this case too. As Alexander of Aphrodisias puts it in his commentary on the *Metaphysics*, "for every number is of something".[22] It is this fundamental characteristic that enables number, *arithmos*, to be covered by the category πρός τι – "in relation to something." And this implies that *arithmos* is abidingly inseparable from that of which it is the number.

Even when designated by a letter, as is done not only by Euclid and Aristotle but also by many others, the determinateness of numbers, their definiteness does not cease. Such a designation does not transform Greek numbers into abstract symbols, into concepts of general magnitude, into parameters, – it merely baptizes them. All this means that there is a wide and historically unbridgeable gap between the Greek and the modern conceptualization of number and that any interpretation of Greek arithmetic that overlooks this gap, be its reasons noble, upright, and progressive as they might be, is bound to be distorting. Euclid's numbers are *definite* aggregates of units of measurement. The presentation is geometric, by means of continuous lines, precisely because the constitutive units are units of *measurement* and lines are directly measurable, and because, after the discovery of incommensurable magnitudes, the Pythagorean monads, the *psephoi*, became unacceptable as means of numerical representation in mathematics, which underwent a thorough geometrization. As Klein puts it, "The result of this development is the reversal of the 'Pythagorean' thesis that the *mensurability* of things is grounded in their numerability, ... now *numerability* is, conversely, understood as a – not even always complete – expression of mensurability."[23]

As against this, modern mathematics is characterized primarily by its symbolic, operational formalism. Modern number is a concept, the concept of general magnitude. This is why, in the Greek context, speaking simply of "numbers" without qualification is, strictly speaking, misleading. What characterizes the Greek concept of number, is the relation it has to the "thing" it intends. The modern concept of number, on the other hand, is the progeny of an abstraction that is symbol-generating. Modern number *is* a symbol. It is the symbol of an *indeterminate* multitude, which, starting with Viète (1540 – 1603), is designated by a letter on which one can and does operate. These letters are called by Descartes (and what a term this is!) "termini generales" and referred to as "puri et nudi".[24] They figure in a discipline that is dubbed by Descartes "Mathesis Universalis" and by Viète "Ars Analytica", the supreme goal of which is "Nullum non problema solvere". The subject matter of "Mathesis Universalis" is "magnitudines in genere" conceived by the intellect as "entia abstracta."[25]

All this is a far cry from Greek mathematics. Elsewhere I referred to the sea change undergone by mathematics in the 16th and 17th centuries as the transformation of the *mos geometricus* into the *mos per symbola*. Indeed Klein speaks in this context of the "intimate connexion between the mode of *'generalization'* of the 'new' science and its character as an *'art'* ". He goes on: "The most characteristic expression of the above connexion is to be found in the symbolic formalism and calculational techniques of modern mathematics." He then points out the important distinction between what he calls *"the generality of the method"* and the *"generality of the object"* of investigation in Greek mathematics, calling attention to the typical tension between them. According to Klein, Greek mathematical methods are dictated by the ontology of the *mathematiká*, the mathematical objects, while modern mathematics starts with a general method and is led by it to the features of the mathematical objects. Although this may sound awkward, I think it is a very significant distinction indeed. It is fundamentally the same distinction as that between "saying" and "showing" in Wittgenstein's *Tractatus*, or the equivalent one between "telling" and "showing" in Wayne Booth's *Rhetoric of Fiction*. It has to do, in the Euclidean context, with what is, on the one hand, the generality of the proof procedure, say in IX.20, *versus* the determinateness of the data and of the specific results reached at the end of the proof. Thus, even though Euclid's proof is undertaken only for three primes and these primes are themselves determinate, specific, $A$, $B$ and $C$, it is a completely general proof, i.e., the procedure is independent of the specific number of primes taken and of their peculiar identity. It is this that explains why, unlike the case of IX. 8, Euclid does not say at the end of the proof that the claim can be shown to be true in the same manner for *any* multitude of prime numbers.

The Euclidean procedure, then, is not symbolic. Euclid, just as the other ancient Greek Mathematicians, did not have at his disposal the concepts of parameter, variable, and unknown. As Jacob Klein puts it,

> In *illustrating* each determinate number of units of measurement by measures of distance it [*i.e.*, the Euclidean presentation] does *not* do two things which constitute the heart of the symbolic procedure: It does *not* identify the object represented with the means of its representation, and it does *not* replace the real determinateness of an object with a *possibility* of making it determinate, such as would be expressed by a sign which, instead of *illustrating* a determinate object, would *signify* possible determinacy.[26]

The fact that number is always *determinate* in Greek mathematics prevents it from playing the role of independent variable in any numerical expression. Greek number is always a function of the units of measurement, it is always a *"number of"*. For a genuine, not a potential, proof by induction, a "pure", "unadulterated", indeterminate, general, abstract number is needed and not a "number of". In other words, without a number that can serve as independent variable, it is impossible to formulate a true proof by mathematical induction, in which the claim requiring proof is a function of the natural numbers.

Generally speaking, the history of mathematics typically has been written as if to illustrate the adage "anachronism is no vice". Most contemporary historians of mathematics, being mathematicians by training, assume tacitly or explicitly that mathematical entities reside in the world of Platonic ideas where they wait patiently to be discovered by the genius of the working mathematician. Mathematical concepts are seen as eternal, unchanging, unaffected by the idiosyncratic features of the culture in which they appear, each one clearly identifiable in its various historical occurrences, since these occurrences represent different clothings of the dame Platonic hypostasis.

Various forms of the same mathematical concept or operation are not considered merely mathematically equivalent but also historically equivalent. Indeed mathematical equivalence is taken to represent historical equivalence. Since the mathematical Forms are eternal and since in their works mathematicians of all ages share in the expression of the same Forms, the specific mathematical idiom used by a mathematician has no bearing on the content of his thought. Mathematical language is at best a secondary appendage of the mathematical culture of any epoch. The mathematical kernel is untouched by the peculiar language used, since all mathematical languages lead back to the same ideal forms. This makes the various casts in which the same mathematical truth has been expressed throughout the centuries completely equivalent. Under such an ontology, the object of the history of mathematics becomes the task of identifying the ideal Forms present in the work of each historical author and apportioning out proper credit to that mathematician who first gave expression to one of these eternal FORMS, i.e., who first brought it out of the eternal Platonic realm into the world of human consciousness. this is precisely the task performed traditionally by the historian of mathematics. But if scholars continue to neglect the peculiar specificities of a given mathematical culture, whether as a result of explicitly stated or implicitly taken-for-granted assumptions, then, by definition, their work is ahistorical and should be recognized as such. It is impossible for modern man to think like an ancient Greek. Historical understanding, however, involves the attempt at faithful reconstruction of the past. In intellectual history, this necessarily means the avoidance of conceptual pitfalls and interpretive anachronisms. Thought it is impossible to think like Euclid, it is rather facile to think obtrusively unlike him. We cannot know what went through Euclid's mind, when he wrote the *Elements*. But we can determine what Euclid could not have thought when he compiled his great work. He, most likely, did not employ concepts or operations for which there is no genuine evidence either in his time or in the works of his predecessors. This much is safe to conclude. Furthermore, he clearly could not have foreseen what mathematicians and historians of mathematics were going to do in the long run to his *Elements*; he could not have used mathematical devices and procedures which were invented many hundreds of years after his death. This much is obvious too. Given that we cannot think like Euclid, we should, nevertheless, strive to avoid thinking unlike him, when elucidating and commenting on his writings. This is, and must remain, the historian's goal. One way of thinking *unlike* Euclid is to use the algebraic approach in interpreting his works.

In mathematics, like in anything else, form and content are not independent variables. On the contrary, they mutually condition one another and neither is immune to change. A certain form permits only a certain content, and a new content requires a new form. This is why the methodological approach, which casts indiscriminately the algebraic shadow over the garden of Greek mathematics, obscures precisely those features which make it *Greek* mathematics. Instead of showing the degree to which it was unlike modern, post-Renaissance mathematics, that approach, by greatly overemphasizing the similarities, prevents an understanding of Greek mathematics in its own right. It also leads in the long rung to the untenable view that the Greek mathematicians did not mean what they said, but that they hid "admirably" their line of thought. Coupled with this is the great danger of easily "discerning" problematic or nonexistent influences between mathematical cultures a world apart, simply because when submitted to the algebraic cure, all mathematical cultures look alike.

Entrenched as it is, the traditional interpretation of the history of ancient mathematics must give way to a new, more sympathetic, and historically responsive, interpretation, simply because the old interpretation has outlived its usefulness and is now an obstacle on the road to a sensitive historical understanding of ancient mathematical texts. After all, like scientific theories, historical theories are tentative attempts to make sense of the past; they are provisional by their very nature, and consequently their authors should not be dreaming hopelessly of endowing them, in God-like fashion, with eternal life and immaculate beatitude.

In every interpretive textual task one can distinguish between *(1)* an "objective", semantic interpretation that conceives the text as physical object, a material, neutral, self-contained and independent entity incorporating all its lessons and messages and (*2*) an intentional, voluntaristic interpretation that denies the possibility of understanding the text properly, i.e., historically, severed from the intentions of its author. The anti-intentional approach is the approach of the mathematician, the declared or undeclared platonist, who, justifiedly, calls attention to the unavailability of the author's intentions and, injustifiedly, alleges the semantic autonomy of the text.

The intentional approach, which claims that the uncovering of the author's intentions is necessary for the proper understanding of the text, is open to the criticism, that it identifies the meaning of the text with a mental entity existing only in the mind of the author before it materializes in the linguistic object, the written text, and that the access to such a mental entity is impossible and anchored in incoherence.

Still the historical approach is an interpretive approach which cannot, by its very nature, divorce itself from *the attempt* to unravel the original intentions of the texts's author, a text, moreover, which is *not* just a well-delimited physical object, autarchic and well-defined with physico-mathematical exactness, enabling one to determine with ease what does and what does not pertain to it. The text is the product of human action, which is inherently intentional, expressing volition, intent, desires, and goals. And so, the only road open to the historian, the interpreter of the past, who refuses to lose her identity as historian, is the road that takes into account the historico-cultural context to which the text belongs, a context, moreover, which acts as a sieve, a philter discarding various interpretations inconsistent with it as unacceptable, not kosher.

The mooring of the text to the context, supplying the historian with the cultural background of the extant socio-professional knowledge, endows historical interpretation with the key that permits the retrieval of intentions which, strictly speaking, are not purely mental, being anchored in an objective textual world available to comparative historical analysis. This analysis is apt to establish with a great degree of likelihood what the text's author could *not* have intended when he wrote his text. This is a *quasi*-intentional approach, since it takes into account circuitously the author's intentions, i.e., the proper intentions of the text, in the historical analysis of his work. Renouncing *ab initio* to look at the text from this perspective, the goal of which is to go after the original meaning of the text, is tantamount to an obliteration of the historical approach.

The reader should not be surprised by the fact that trade typically leaves an imprint on character. An opinion consonant with this fact is already in the *Nicomachean Ethics*. It is also there, by the way, that Aristotle said that: "The saying that 'no one is voluntarily wicked nor involuntarily happy' seems to be partly false and partly true; for no one is involuntarily happy, but wickedness *is* voluntary" (1113b 14–15). It is this passage, I take it, that Gonzalez-Crussi has in mind, when he says in his *Notes of an Anatomist*: "virtue (as well as vice) is literally a matter of habit and 'on-the-job-training'."[27] Be that as it may, I am sure the reader will agree that character, too, typically leaves an imprint on trade.

*University of Tel Aviv*

NOTES

[1] Thomas, L. Heath, *The Thirteen Books of Euclid's Elements*, 3 vols, vol. 1 (New York, 1956), p. 379.
[2] Cf. the English version of the statement in B. L. van der Waerden, *Science Awakening* (New York), p. 118.
[3] T. L. Heath, *op. cit.*, p. 382.
[4] *Ibid.*, p. 383.
[5] Cf. Heath, *op. cit.*, vol. 2, p. 390.
[6] *Ibid.*, p. 278.
[7] *Ibid.*, p. 379.
[8] *Ibid.*
[9] *Ibid.*, p. 391.
[10] *Ibid.*, p. 412.
[11] *Ibid.*, p. 332.
[12] Bertrand Russell, *Mysticism and Logic and Other Essays* (New York, 1971), pp. 59–60.
[13] Nicolas Bourbaki, *Elements of Mathematics: Theory of Sets* (Paris/London, 1968), p. 16.
[14] Hans Freudenthal, "Zur Geschichte der vollständigen Induktion." *Archives Internationales d'Histoire des Sciences*, vol. XXII (1953), pp. 17–37, at p. 28.
[15] *Op. cit.*, vol. 2, p. 392.
[16] Cf. n. 9 *supra*.
[17] Cf. n. 15 *supra*.
[18] *Op. cit.*, p. 30.
[19] *Ibid.*

[20] *Op. cit.*, vol. 2, p. 277.
[21] Aristotle, *Metaphysics*, 1057 a 2–5.
[22] Πᾶς γὰρ ἀριθμός τινός ἐστι (*Alexandri Aphrodisiensis in Aristotelis Metaphysica Commentaria*, ed. M. Hayduck (Berlin, 1891), p. 86, ll. 5–6).
[23] Jacob Klein, *Greek Mathematical Thought and the Origin of Algebra* (Cambridge, Mass., 1968), p. 240 n. 121.
[24] René Descartes, *Regulae ad directionem ingenii* in *Oeuvres*, eds. Ch. Adam and P. Tannery, vol. X (Paris, 1974), p. 457 l. 20; p. 455, l. 21.
[25] J. Klein, *op. cit.*, p. 209.
[26] *Ibid.*, p. 123.
[27] F. Gonzalez-Grussi, *Notes of an Anatomist* (London, 1986), p. 62.

## REFERENCES

Alexander of Aphrodisias. *Alexandri Aphrodisiensis in Aristotelis Metaphysica Commentaria*. Ed. M. Hayduck. Berlin, 1891.

Aristotle. *Metaphysica*.

Nicolas Bourbaki. *Elements of Mathematics: Theory of Sets*. Paris/London, 1968.

René Descartes. *Regulae ad directionem ingenii. Oeuvres*, vol. X. Eds. Charles Adam and Paul Tannery. Paris, 1974.

Hans Freudenthal. "Zur Geschichte der vollständigen Induktion." *Archives Internationales d'Histoire des Sciences*, vol. XXII, 1953.

F. Gonzalez-Grussi. *Notes of an Anatomist*. London, 1986.

Thomas, L. Heath. *The Thirteen Books of Euclid's Elements*, 3 vols. New York, 1956.

Jacob Klein. *Greek Mathematical Thought and the Origin of Algebra*, Cambridge, Mass., 1968.

Bertrand Russell. *Mysticism and Logic and Other Essays*. New York, 1971.

B. L. van der Waerden. *Science Awakening*. New York, 1963.

DAVID H. FOWLER

# THE STORY OF THE DISCOVERY OF INCOMMENSURABILITY, REVISITED

I take as my opening text the kind of thing my colleagues – certainly the mathematicians and often the historians of mathematics – might say about the beginnings of Greek mathematics. Something like this:

The early Pythagoreans based their theory of proportion on commensurable magnitudes (or on the rational numbers, or on common fractions $^m/_n$), but their discovery of the phenomenon of incommensurability (or the irrationality of $\sqrt{2}$) showed that this was inadequate. This provoked problems in the foundation of mathematics that were not resolved before the discovery of the proportion theory that we find in *Elements* V.

You must, at some time or another, have heard, or perhaps even have said, something like that. I've certainly said it; I even got into *Punch* for saying it![1] But I shall try explain why I now disagree with everything in this line of interpretation. My space is limited so I will have to refer you for many of the crucial details in what follows to the thorough discussions I have cited or given in my book, *The Mathematics of Plato's Academy*; in fact, I hope this article will form the basis of the opening chapter of a sequel to this book.[2] I shall arrange my comments under various headings. First we have the matter of:

*The nature of our evidence*. Our evidence about Greek mathematics in general comes in very disparate forms, and almost all of it has been subject to an unknown amount of editing and interference. In particular our late sources – editions, compilations and commentaries dating from the 2nd century AD onwards – are manifestly of very variable quality. So, in the first phase of my reconstruction, as represented by my book, I put it all to one side as far as possible, and ignore it.[3] (This is very drastic, and my hope is to consider some of this later evidence in the sequel.) Moreover, some of the relevant evidence in early sources, in particular Euclid and Plato, comes in homogenous slabs which often fit rather awkwardly in the various versions of the received interpretation. In some measure to redress the balance after my radical approach to the late texts, I endeavour to follow the principle that, if any piece of such an early slab enters the



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 221-235.

reconstruction in a significant way, then the proposal should also engage with the whole of its context. For example, any important use of any aspect of the curriculum in Plato's *Republic* VII should also connect with the whole of this curriculum, especially since Plato insists on its unity; any significant application of a proposition in *Elements* II should eventually involve all 14 of them; anything about *Elements* XIII, the book which contains the construction of regular solids and a lot more, should say something about Book X, the classification of incommensurables, and perhaps also about Books IV and II.

Please note: I am not saying early evidence is all good, late evidence is all bad. I am saying that our evidence is a hodge-potch that we cannot sort out until later in the project, but the best evidence is likely to be found in the coherent and obscure chunks of early provenance (this is analogous to the principle *difficilior lectio* of textual criticism), so let us start from this material, taken all in one piece. It is a methodological principle, not a simple value judgement of the evidence.

Back now to my opening text. In fact, my principal objection to it is that it is founded entirely on late evidence and speculation, uncorroborated to a remarkable extent by any of our earlier sources, so it forms a very insubstantial base on which to start our reconstruction of Greek mathematics. I shall spell that out in more detail, and then propose an alternative starting point.

*On Pythagoras and the early Pythagoreans*, see, for example, W. Burkert, *Lore and Science in Ancient Pythagoreanism*. I subscribe fully to his general conclusion, that the only kind of scientific activities and discoveries we can attribute with confidence to the early Pythagoreans are some remarkable findings in music theory and acoustics, the most remarkable being that consonance is associated with small integers. Most, if not all, of the mathematical stories have the ring of later legendising. In particular, consider:

*The Pythagorean theory of proportion based on commensurable magnitudes*. I know of no explicit evidence for this, early or late. From what I can work out, the thinking goes something like this: we, since medieval, times have learned at school about common fractions, that is $^{m}/_{n}$ s, so that is what the Pythagoreans must have used, especially since we find these fractions later in Greek mathematics and Greek accounting. Occasionally this opinion gets expressed in print; see, for example, B. L. van der Waerden, *Science Awakening*, pp. 49–50 & 115–6. He writes:

It is probable that it was *calculation with fractions* which led to the setting up of *Elements* Book VII [which van der Waerden attributes to the early Pythagoreans]. Fractions do not occur within the official Greek mathematics before Archimedes but, in practice, commercial calculations had of course to use them... .

Once again, I disagree with everything in this opinion. Fractions do occur in 'official mathematics', even, in a limited way, in the *Elements*: see the use of the word *meros*, plural *merê*, translated as 'part' and 'parts', especially in Book VII. For example, VII 37 & 38 talk of 'homonymous parts', of three & the third, four & the quarter, etc., and these *merê* are an ingredient of the way fractions are described in Greek. But I do not think there is any unambiguous evidence for *common* fractions, these $^m/_n$ s, in any of our early texts including commercial calculations, and their apparent appearance in the late Byzantine copies which account for 99% or more of our evidence may be instead as scribal abbreviations. Our plentiful surviving explicit evidence is that Greek fractional practice was exactly the same as Egyptian practice: fractions were expressed as sums of the *merê*, as so-called unit fractions. The details of this argument are long, painful, and contentious because, while we have lots of different kinds of evidence, most of it is of the wrong sort or it comes from the wrong place or the wrong time. The details are given in Chapter 7 of my book and summarised in my article 'Logistic and fractions in early Greek mathematics'. And if you find the full conclusions of my overall thesis too much to swallow, we need only a much weaker version of it here, that the Greek mathematicians before Plato and Eudoxus used not common fractions, but Egyptian fractions. I'll return to fractions and arithmetic later.

Note that, here and elsewhere, 'Greek' as in 'Greek mathematics' simply means 'written in Greek'. Almost all of our evidence has been transmitted via Egypt, and then via the whole eastern Mediterranean, and that, of course, complicates the argument. Also, concerning fractions and division, I think that Diophantus needs a separate discussion.

*The topic of incommensurability.* At the end of my book, I review all of the evidence on this topic for the thousand year period up to Proclus and, there and here, I refer you also to another such review in W. R. Knorr, *The Evolution of the Euclidean Elements*, Chapter 2. Here are some opinions from these reviews.

• In the first surviving explicit mention of incommensurability,[4] in Plato's *Theaetetus* (147a ff), the topic is handled confidently as a source of interesting mathematical research. Incidentally, as to the date of Theaetetus' death, which is generally regarded as one fixed point, perhaps the only secure fixed point, in the shifting sands of the incommensurability issue, I also note that there is now one expert who dissents from the common view. H. Thesleff, specialist on pseudo-Pythagorean texts, writes: "I find it essential to note that historians of mathematics who take it for granted that Theaitetos was still alive in the 370s must be wrong. He made some important discoveries as a young man, and Plato and his friends were deeply impressed by this. But he is likely to have died in 390 BC."[5]

• The celebrated passage in Plato's *Laws* (817e ff) where Plato talks of "ignorance ... not worthy of human beings but pigs" is probably not referring to incommensurability in our sense here, but something else, very possibly the kind of techniques used in land measurement, where again things are not what mathematicians and historians of mathematics seem to assume they ought to be when, for instance, they parade stories from commentators about Egyptian land measurement as the origin of mathematics.

• Aristotle's favourite mathematical illustration is "the incommensurability of the diagonal" as something all mathematicians know; but he never suggests that it is or ever was a disaster for any mathematical theory, even though, in closely related passages, especially in the *Metaphysics*, he is highly critical of the Pythagoreans. Curiously, Aristotle never specifies that he is talking of the diagonal of a square. Twice, both times in *Prior Analytics*, at 41a23 ff & 50a35 ff, he says something like: "from the assumption that the diagonal is commensurate, it follows that odd numbers are equal to evens", but Aristotle gives no more details of what he means by this. I find completely convincing Knorr's proposal (*Evolution*, pp. 228–232) that the so-called Pythagorean proof, revolving around this statement, was tacked on to the end of *Elements* X in two clumsy versions sometime after the time of Alexander of Aphrodisias in the 3rd century AD, in response to the needs of Aristotelian commentators, and that Alexander himself cobbled together the variant of this proof to be found in his commentary. However other natural mathematical explanations of the remark by Aristotle about odd and even numbers are possible, and our evidence for any interpretation of it so tenuous as to be unreliable as a basis for further reconstruction. I will come back again to Aristotle later.

• No surviving fragment or testimony of Eudoxus mentions incommensurability: the word-index to Lasserre, *Die Fragmente des*

*Eudoxos,* does not contain the words *(a)summetros*, *(ar)rhetos*, or *alogos*! To us, it may seem blindingly obvious that the principal achievement of the Eudoxan theory of *Elements* V must have been to accommodate ratios of incommensurable magnitudes, but no ancient source says that, and I intend to include a chapter on Eudoxus in the sequel that tells this story differently. For the moment, I'll have to leave it at that, and add that you'll already find most of the ingredients in my book, worked out in some detail.

• Proclus never quotes anything from Eudemus on incommensurability, though he cites Eudemus by name several times, and also writes about incommensurability several times. The one apparent exception to this, the passage in the catalogue of geometers where Eudemus appears to refer to Pythagoras, is almost certainly an interpolation; and the remark there that Pythagoras discovered the "doctrine of proportionals" is a modern editorial emendation of the text, where all of our manuscripts unanimously have "the doctrine of the *alogos*".

• A proper discussion of this word *alogos* would take us on a very long excursion into *Elements* Book X; instead, let us here just look very briefly at Euclid: As far as I am aware, the only time the topic of incommensurability appears in Euclid's works is in *Elements*, Books X & XIII, which form our only coherent slab of early evidence on the topic, and a very massive, very coherent slab it is, in bulk and content well more than a quarter of the *Elements*. Any discussion of incommensurability should deal with Book X; but if you have never looked at Book X, I can promise you that you will find it difficult, so I recommend you to try to get hold of a pamphlet by Christian Taisbak called *Coloured Quadrangles*. You may find that difficult too – I mean getting hold of Taisbak's pamphlet! – so I have given a version of it in Chapter 5 my book, and subsequently written an improved version of this: 'An invitation to read Book X of Euclid's *Elements*'. As to the role of Books X and XIII in my reconstruction of the incommensurability story, I must refer you to the rest of Chapter 5 of my book.

• The source of most of the stories about Pythagoras, Pythagoreanism, and incommensurability is the book *On the Pythagorean Life* by Iamblichus, so perhaps it is worth quoting the relevant passages in full:[6]

§18 (88) ... As for Hippasos, he was indeed a Pythagorean, but because he was the first to make public the sphere constructed from twelve pentagons he was lost at sea for his impiety: he got the reputation of having discovered it, but it all came from 'that man' – that is what they call Pythagoras: they do not use his name.

§34 (246) ... The first man to reveal the nature of commensurability and incommensurability[7] to those unworthy to share his teaching was so much detested, they say, that not only was he excluded from their common life and meals, but they built him a tomb as if their former companion had left human life behind. (247) Some say the supernatural power took revenge on those who published Pythagoras' teachings. The man who revealed the construction of the 'twenty-angled shape' was drowned at sea like a blasphemer. (He told how to make a dodecahedron, one of the 'five solid figures', into a sphere.) Some say this fate befell the man who told about irrationality and incommensurability.

This farrago of mutually inconsistent stories, which appear for the first time in a source of doubtful reliability and relevance dating from some nine centuries after the time of Pythagoras, is the main evidential base for much of what has been written about the discovery of incommensurability! I add the slightly perplexed comment that the most recent and authoritative study of the role of mathematics in neo-Pythagorism, D. J. O'Meara's book *Pythagoras Revived, Mathematics and Philosophy in Late Antiquity*, does not even seem to mention incommensurability.

*The foundation crisis.* Again I find Freudenthal, Knorr,[8] and other writers very convincing when they argue that, far from being a period of crisis and confusion, the early fourth century was an extraordinary period of creativity, especially in Plato's circle; we have no historical evidence for any of the postulated difficulties of a 'foundation crisis'. But I want to go further and explore the possibility that the discovery was an incidental event in the early development of mathematics. So let us now look at the evidence concerning the effects of the discovery. This will involve tracking over some of the same material again.

*The implications of the discovery of incommensurability.* Just what precisely are the supposed problems raised by the discovery of incommensurability? As far as I know, no Greek text, early or late, tells us clearly of the mathematical difficulties raised by the phenomenon. Aristotle wrote of the innocent's surprise, and the way it then gives way to a more informed appreciation:

All men begin ... by wondering that the matter is so (as in the case of automatic marionettes or the solstices or the incommensurability of the diagonal; for it seems wonderful to all men who have not yet perceived the explanation that there is a thing which cannot be measured even by the smallest unit). But we must end in the contrary and, according to the proverb, the better state, as is the case in these instances when men learn the cause; for there is nothing which would surprise a geometer so much as if the diagonal turned out to be measurable (*Metaphysics* 983a 12–20).

Pappus' *Commentary on Book X of Euclid's Elements*, I.2, writes, much later, of how thereafter:

the soul ... wanders hither and thither on the sea of non-identity ... immersed in the storm of the coming-to-be and the passing-away, where there is no standard of measurement

from which passage I shall grasp, below, on the only substantial straw, the last four words: "there is no standard of measurement". There is the similar Scholium 1 to Book X, quoted in part in the Introduction to Book X in Heath's translation of the *Elements*, where you will find a long discussion that is remarkable for its lack of any hard evidence.[9] And Proclus, in his *Commentary on the First Book of Euclid's Elements*, 60, writes:

The statement that every ratio is expressible (*rhetôs*) belongs to arithmetic only, and not to geometry, for geometry contains inexpressible ratios (*arrhetos logos*)

but this is not, very much not, the terminology of *Elements* X, which is built on a completely different meaning for expressible incommensurable lines and ratios.

So, prompted by Aristotle, let us try to "perceive the explanation" and "learn the cause" of incommensurability. But first, may I persist a bit longer with a variation on my question at the beginning of this section:

What precisely, to modern commentators, are the supposed difficulties revealed by incommensurability? A whole range of possibilities now seems to be on offer. With some overlap, there are the following:

- The Proclus objection: Incommensurable ratios cannot be expressed in (or by) numbers. Surely that is nonsense! We express them in numbers, and much of mathematics is, ultimately, numbers. Early Greek mathematicians could have expressed them in numbers. We have no explicit evidence that they did this, but I am arguing for a speculative interpretation in which they were expressed in numbers. My book is full of examples.
- Incommensurable ratios could not be fitted within the pre-Eudoxan style and scope of mathematics. That is false! The definition for lines that $a{:}b{::}c{:}d$ means rectangle $(a,d)$ = rectangle $(b,c)$ uses only believed-to-be early ingredients in a believed-to-be early way, and will handle most of what we need for the *Elements*, perhaps everything except for compounding, on which Euclid's own treatment is strange and unsatisfactory. But we have no evidence that this was an early definition.
- Incommensurability showed the inadequacy of the Pythagorean

doctrine that "all is number". Therefore, some add, it had to be concealed. Curiously, Aristotle, our principal early witness, from whom we learn of this "all is number", never advances this criticism, even though all of this material comes together in the *Metaphysics*. We find there lots of mentions of incommensurability, we find summaries and harsh criticisms of Pythagorean philosophy, but we don't find this objection. Zhmud has recently even proposed that this formulae "all is number" was Aristotle's invention.[10] And, I would add, incommensurability need not show the inadequacy of the doctrine that "all is number", *pace* Proclus; indeed, in my reconstruction, the behaviour of the anthyphairetic ratios of $\sqrt{n} : \sqrt{m}$ might even reinforce such a doctrine, and give a mathematical explanation of the 'expressible' lines that underlie *Elements* X and XIII.

- The discovery of incommensurability showed that the [Babylonian?] arithmetical basis of geometry was inadequate, so geometry had to be reformulated purely geometrically, for example, as in *Elements* II. We have no evidence of this supposed early, possibly Babylonian, arithmetical basis of early Greek mathematics, so this is pure speculation. Also, there are other explanations of the role of *Elements* II. I shall return to this issue later, when I discuss arithmetisation.
- Incommensurable ratios can only be approximated, while Greek mathematics aimed for precision. This assertion runs counter to our evidence: Archimedes is interested in approximation. Also Aristarchus and then, later, Hypsicles, Hero, Ptolemy, ... . Moreover the fraction 1/7 can only be approximated in sexagesimal arithmetic, and the fraction 1/3 can only be approximated in the Greek system of land measurement which, by convention, only uses the parts 2', 4', 8', 16',... . So Greek mathematics was involved in approximation and the necessity of approximating simple numerical ratios was a well-known phenomenon.
- The discovery put into question the basic idea of ratio. I prefer to reformulate this, since no early text seems to express this concern: hints of it seem to surface first in later commentators like Iamblichus and Pappus, and then grow thereafter. So I shall change the question once again and ask instead:

Why do none of the early testimonies seem concerned with the manifest difficulties that incommensurability poses to our way of defining ratio? This emphasises that the difficulties may be *our* difficulties, not necessarily theirs; they may have had different ways of thinking about ratio. I finish by developing this theme a bit.

Incommensurability does present a problem to arithmetised geometry, though ultimately I think it turns out that the problem was already lurking there even for commensurable manipulations. Let me try to explain. Arithmetised geometry is how we tend to think of geometry today: a line has a length, a number; a rectangle has an area, again a number which is equal to the product of the lengths of its sides; ratios are defined arithmetically, as quotients of numbers; and so on. So the geometry becomes translated into the arithmetical manipulation of numbers – addition, subtraction, multiplication, division, taking roots, etc. – and then this arithmetic is later abstracted into algebra. For example, the so-called Pythagoras' theorem becomes $a^2 + b^2 = c^2$, where the usual interpretation of this statement involves the lengths (i.e. numbers) $a$, $b$, and $c$ of the sides of the triangle. It then seems that the numbers associated with things like $\sqrt{2}$ are rather complicated, so complicated that filling in all of these irrational numbers properly could not be done before the middle of the nineteenth century. Dedekind, the first, tells us that he succeeded on 24 November 1858; it was a Wednesday.

But read Dedekind carefully, and you will see that there are already serious problems with arithmetic. I go yet further than this, and argue that there need be no real problems in defining the set of numbers, for example, as decimal sequences, or sexagesimal sequences, or anthyphairetic sequences, or my astronomical sequences, or other such descriptions. However, a precise description of the arithmetic, even the arithmetic of the rational numbers when they are not being conceived as common functions, is intractable.[11] Hence the crucial role of my belief that early Greeks did not use common fractions, so they would not think of ratios in any way like our rational numbers, and so would not think that arithmetic was a natural and obvious basis on which to build their mathematics.

To pursue this theme further, early Greek geometry seems to me to be not arithmetised to a remarkable extent,[12] but later Greek mathematics is arithmetised. There are two different traditions in this later arithmetisation. The first is the astronomical one, using Babylonian sexagesimal numbers, which is not attested in Greece before Hypsicles and Hipparchus in the 2nd century BC. It may have been transmitted much earlier, but that is pure speculation, as was remarked earlier. The second is the Graeco-Egyptian unit fraction tradition which we find in our earliest testimonies like the Hibeh Papyrus, then later in the Heronian Corpus, and in Ptolemy alongside the sexagesimal astronomical calculations. The arithmetic of both poses theoretical problems, and both are spectacularly absent from the geometry of the *Elements*: they are incompatible with the spirit of the proportion theory of Book V and remote from the spirit of the treatment of incommensurability of Book X. No wonder commentators from antiquity onwards, who seem to work in a vaguely arithmetised geometry, have a hard time, and no wonder there seems to be some confusion.

So let us try to purge our mind of this arithmetised geometry. (I did not find this easy and it took me some years spent with some good alternatives.) One problem we then face is defining ratio or proportion. (Note that this problem is not now directly concerned with incommensurability.) I have already pointed out that we have no real difficulty in fitting proportionality into the pre-Eudoxan style of mathematics. But our only evidence on how early Greek mathematicians actually handled ratio or proportion is the celebrated passage in Aristotle's *Topics* 158b29ff on *antanairesis/anthyphairesis* which suggests the use of the so-called Euclidean algorithm:

Given two numbers or two lines (or, with a bit of technique at our disposal, two more complicated geometrical objects), then see:
how many times the second can be subtracted from the first;
how many times the remainder can be subtracted from the second;
how may times the next remainder can be subtracted from this remainder;
etc.
and this gives a string of numbers that characterise the relationship of size between the two things.

For example, the ratio of 60 to 26 will be twice, three-times, four-times exactly. Do the process for two numbers, and you will quickly see that it must stop after a finite number of steps, because if not "an infinite series of numbers will [arise], each of which is less than the other, which is impossible in numbers" (as formulated in *Elements* VII 31). Now do it with two lines. Early Greek geometers seem to take little notice of the philosopher's problems, and their lines can be chopped up indefinitely, so the possibility of the anthyphairesis going or indefinitely presents itself. If you are a geometer, possessed of a bit of technique, and this question poses itself, you will soon find examples of it happening, as I shall illustrate below. Thus, as Aristotle might be saying, we could "learn the cause" and "perceive the explanation" of incommensurability in this way of thinking about the ratios of lines. But if, as Pappus might be saying, we involve geometry in "standards" of measurement", which I take here as a hint of arithmetisation, then we encounter problems, as I have tried to explain.

Here then are some illustrations. First consider the problem of 'the diagonal and the side': draw any regular polygon, and evaluate the anthyphairetic ratio of one of its diagonals and its side. The easiest and best known example is the pentagon, so I leave that for the reader, and will consider here the square. We easily see that the side S of a square goes once, but not twice, into its diagonal D (this follows from Socrates' comments at Plato's *Meno*, 82a-85d), and so the ratio (diagonal to side) is once, followed by the ratio (side to diagonal minus side). We are now faced with the evaluation of this ratio of the side to diagonal minus side, and some may feel that, like Meno at 84a, that we have made little progress: "It's no use Socrates, I just don't know". But, just as Socrates unblocks that impasse by conjuring a clever figure out of thin air, so I here draw Figure 1: Starting from the small diagonally placed square in the left-hand corner, with side s and diagonal d, we construct a larger square whose side S is s+d, and check that its diagonal D is 2s+d.[13]

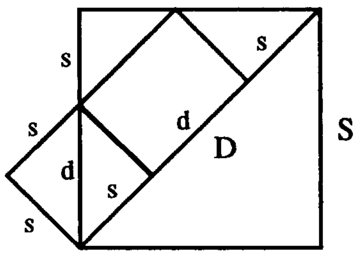



Fig. 1.

(Note how the symbols, here and below, are pure shorthand for lines and contain no hint of arithmetisation.) With one eye on Figure 1, and with the insight that the ratios are unaffected by the scale or orientation of the figures involved, we take up our problem again, and see that the ratio (big side to big diagonal minus side) is the same as the ratio (little side plus diagonal to little side); and we can now evaluate this as twice, followed by the ratio (little side to little diagonal minus side) which is – scale and orientation aside – what we just started from. Hence the ratio (side to diagonal minus side) is twice, twice, twice, twice, continuing thus indefinitely, and so the ratio (diagonal to side of a square) is once, twice, twice, twice, ... .

Let me evaluate this same ratio another way; or, to be more exact, let me give another proof that ratio (diagonal plus side to side) is twice, twice, twice, twice, ... . Start with a square P and, by adding on a gnomon Q+R+S = P, as in Figure 2, construct a larger square of size 2P, whose side will therefore be the diagonal of the original square; and then append a rectangle T equal to Q, as shown. Then ratio (diagonal plus side to side)

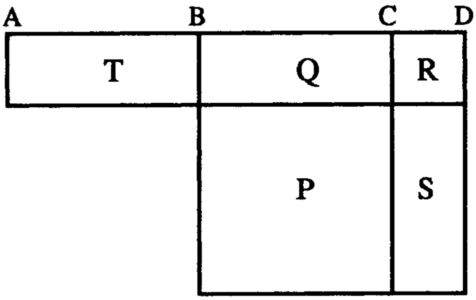



Fig. 2.

will be ratio (AD to BC), which we immediately see is twice, followed by ratio (BC to CD). The required proof will now follow if we can show that ratio (BC to CD) is equal to ratio (AD to BC); or equivalently, by a manipulation of geometrical proportions (see *Elements* VI 16; this manipulation was alluded to earlier), that rectangle (AD, CD) = square (BC), that is T+Q+R = P; but this underlies our construction, since T = Q = S. Q.E.D.

The case of the ratio of the longer diagonal to side of a hexagon gives rise to the ratio $\sqrt{3}$ to 1, where $\sqrt{3}$ denotes the side of the 3-fold square (which can be constructed, for example, using *Elements* II 14); and this example can be generalised to the investigation of the ratios like $\sqrt{n}$ to $\sqrt{m}$ – what I propose to call the problem of 'the dimensions of squares'. We first use numerical techniques very close to those found in *Elements* VII to explore and conjecture what the answer might be. The second kind of proof above then deals with the simpler cases; a comparison of the mechanisms of Figure 2 and *Elements* II 11, gives us insight into the particular example of the ratios $n$-times, $n$-times, $n$-times, ...; and yet more heuristic exploration, followed by a generalisation of Figure 1 which involves *Elements* II 13 & 14, gives a complete solution of this remarkable problem and a new interpretation of the whole of *Elements* II.

An analogous problem of 'the dimension of cubes' beckons, but proves to be of such redoubtable difficulty that in fact it remains unsolved today; Plato's remarks at *Republic* 528b-c are still perfectly applicable! We can also explore related problems such as 'the circumdiameter and side' of polygons or polyhedra and see where they lead: the can provide explanations of the roles of *Elements* IV, X, and XIII. And the similar problem of 'the perimeter and the diameter' might be what lay behind Archimedes' original calculation in his *Measurement of a Circle.*

Developing these ideas takes us into a different world. The ingredients are all early Greek, but they are fitted together to create a completely different picture. It is mathematically appealing, amazingly coherent (too coherent, in fact, for my comfort), and consonant with a quite extraordinary breadth of our evidence. And the topic with which I started, the simple discovery of incommensurability, plays no significant part in it, which is why my book contained no discussion of it beyond the bald and sceptical catalogue of our evidence at the end of its penultimate chapter.

I am constantly tinkering with the details of my book, so I shall finish with one final little addition to it, in the first of my dialogues, where my slaveboy discovers this anthyphairetic definition of ratio under Socrates' prompting. They start doing it on numbers (heaps of stones, in fact) and the slaveboy realises that the process must terminate. Then Socrates introduces the possibility of doing it with lines. At this point, at the end of $B_{36}$, on p. 28, add: 'I wonder if it now can go on for ever'. That's my slaveboy realising one of the causes and explanations of incommensurability, and it is just a passing remark on the way to discovering the much more remarkable things alluded to in $S_{37}$-$S_{43}$ and described above, alongside which this simple fact of incommensurability fades into insignificance.

*University of Warwick*

NOTES

[1] *Punch*, April 24, 1974. This magazine provided a humourous commentary on British life for 150 years until its closure in March 1992. In recent years, it ran a regular column 'Country Life' which explained itself as follows: "Not everything that happens in Britain gets into the national press. This feature presents some of the news that never made it." One reader sent

in the following clipping: "The programme, which is about the development of number systems, will include an interview with David Fowler, of Warwick University, on the historical crises associated with the square root of two." This must ultimately have come from some Open University publicity about a TV programme I had helped make for their first History of Mathematics course, which had been then been passed on by my university, picked up by the local newspaper, *The Leamington Spa Courier*, and submitted to *Punch* by a local reader. This programme was my first and reluctant venture into Greek mathematics, and I later disowned it, only permitting the Open University to continue broadcasting it if it also circulated a disclaimer by me to the students doing the course!

[2] Hereafter I shall refer to these as 'my book' and 'the sequel'.

[3] At the outset, I must admit that one piece of evidence of late provenance plays a crucial role in the reconstruction, namely the material on 'side and diagonal' numbers and lines, found in Theon of Smyrna, Iamblichus, and Proclus. See the discussion in my book, pp. 100–4, which however does not deal with the material in Iamblichus (ed. Pistelli, 91.3–93.6).

[4] H. Thesleff, Platonic chronology, on p. 18, n. 47.

[5] I here leave to one side the notorious 'nuptial number' at *Republic* 546b, with its talk of the 'rational and irrational diameters of five'.

[6] These quotations are taken from a new English translation, by G. Clark: *Iamblichus: On the Pythagorean Life*, Liverpool, 1989.

[7] The translation has "symmetry and asymmetry", but surely (in)commensurability is the appropriate translation here of *(a)suummetros*, as in the phrase "irrationality *(alogos)* and incommensurability" at end of the passage. The later "twenty-angled shape" is *eikosagonon*, a non-standard name (perhaps found only here in Iamblichus) for *dodekaedron.*

[8] See Freudenthal, 'Y avait-il une crise de fondements...' and Knorr, *Evolution*, 306–12.

[9] Contrast this with the sparse sections in Thomas, *Selections* i pp. 110–11 & 214–17, on 'The Irrational', where this and a passage from Aristotle are the only texts excerpted.

[10] L. Y. Zhmud, '"All is number"? "Basic doctrine" of Pythagoreanism reconsidered'.

[11] For a discussion and examples, see my '400 years of decimal fractions', '400.25 years of decimal fractions', and 'Dedekind's theorem: $\sqrt{2} \times \sqrt{3} = \sqrt{6}$'.

[12] The only arithmetised passage I know, anywhere up to Archimedes and beyond, is in Plato's *Meno*, 82c ff, where Socrates says to the slaveboy: "Now if this side is two feet long, and this side the same, how many feet will the whole be"; and the passage continues in this arithmetised vein for the slaveboy's first two attempts. (I thank Wilburg Knorr for pointing this out to me in January 1991 in a train somewhere between Verona and Venice; I had used this passage for the introduction to my book without appreciating this feature!) The switch to geometry is then indicated by Socrates when he tells the slaveboy: "If you don't want to count it up, just show me on the diagram" (84a). And I think one can explain this singular exception by observing that the slaveboy is not a mathematician – that is the whole point of the episode. Note that an arithmetisation of aspects of everyday life occurs when a barter economy gives way to the use of money, which seems to have happened in Greece by the 7the century.

[13] This my proposed interpretation of the figure being described in the texts on side and diameter lines; see note 3, above.

## REFERENCES

W. Burkert, 1972, *Lore and Science in Ancient Pythagoreanism*, Cambridge, Massachusets; translation by E. L. Minar of *Weisheit und Wissenschaft*, Nurenberg, 1962.

R. Dedekind, 1872, *Stetigkeit und die irrationale Zahlen*, & 1888, *Was Sind und was sollen die Zahlen*, translated by W. W. Beman in *Essays on the Theory of Numbers*, Chicago, 1901.

D. H. Fowler, 1985, '400 years of decimal fractions', *Mathematics Teaching* **110**, pp. 20–21, & '400.25 years of decimal fractions', *ibid.*, **111**, pp. 30–31.

D. H. Fowler, 1987, *The Mathematics of Plato's Academy: A New Reconstruction*, Oxford.

D. H. Fowler, 1992a, 'Logistic and fractions in Greek mathematics: a new interpretation', pp. 133–47 in P. Benoit, K. Chemla, & J. Ritter, *Histoire de Fractions, Fraction d'Histoire*, Basel.

D. H. Fowler, 1992b, 'Dedekind's theorem: $\sqrt{2} \times \sqrt{3} = \sqrt{6}$', *The American Mathematical Monthly* **99**, pp. 725–33.

D. H. Fowler, 1992c, 'An invitation to read Book X of Euclid's *Elements*', *Historia Mathematica* **19**, pp. 233-64.

H. Freudenthal, 1966, 'Y avait-il une crise de fondements des mathématiques dans l'antiquité', *Bulletin de la Société Mathématique de Belgique* **18**, pp. 43–55.

T. L. Heath, 1926, *The Thirteen Books of Euclid'S Elements*, 2nd. ed., Cambridge.

W. R. Knorr, 1975, *The Evolution of the Euclidean Elements: A Study of the Theory Incommensurable Magnitudes and Its Significance for Early Greek Geometry*, Dordrecht.

F. Lasserre, *Die Fragmente des Eudoxos von Knidos*, Berlin.

D. J. O-Meara, 1989, *Pythagoras Revived, Mathematics and Philosophy in Late Antiquity*, Oxford.

C. M. Taisbak, 1982, *Coloured Quadrangles: A Guide to the Tenth Book of Euclid's Elements*, Copenhagen.

H. Thesleff, 1989, 'Platonic chronology', *Phonesis* **34**, pp. 1–26.

I. Thomas [=Bulmer-Thomas], 1939, *Selections Illustrating the History of Greek Mathematics*, vol. i, London/New York.

B. L. van der Waerden, 1954, *Science Awakening*, Groningen; translation by A. Dresden of *Ontwakende Wetenschap*, 1950, Groningen.

L. Y. Zhmud, 1989, '"All is number"? "Basic doctrine" of Pythagoreanism reconsidered', *Phronesis* **34**, pp. 270–92.

JEAN CHRISTIANIDIS

# ON THE HISTORY OF INDETERMINATE PROBLEMS OF THE FIRST DEGREE IN GREEK MATHEMATICS

One of the most interesting issues in the early history of algebra is that of the formation of the methods for solving the first-degree indeterminate equations. In the traditional historiography such methods are primarily associated with the Chinese and the Indian mathematical traditions. The fact that the Remainder Theorem is commonly called "Chinese Remainder Theorem" in almost all the textbooks on Number Theory, strikingly expresses the traditional viewpoint. The same holds about the so called problem of the "hundred fowls", the origins of which are also reduced to the Chinese mathematical tradition.[1]

In the "picture" of Greek mathematics which is emerged from the manuals on the history of mathematics, the issue of linear indeterminate analysis is totally absent.[2] This gives rise to a rather misleading impression that the Greek mathematicians of the classical era were not preoccupied with linear indeterminate problems. It is considered that such problems appear in Greek mathematics in the late Hellenistic or even in the Byzantine era, and echo mathematical traditions rather alien to the Greek one.

Although this "picture" does not seem natural to a number of historians, the objections that have at times expressed bear the character of conjectures rather than of real arguments. Thus, during the last century J. F. Wurm suggested an alternative interpretation of Archimedes' *Cattle Problem*, according to which it concerns not an equation Fermat–Pell, but a system of linear indeterminate equations of the form $ax\text{-}by=1$.[3] This interpretation has recently been subordinated by D. H. Fowler to the Greek anthyphairetic ratio theory.[4] In a work devoted to the lost books of Diophantus' *Arithmetica*, Tannery has also suggested that, one of them had, possibly, contained problems of the form of the "remainder problem".[5] However, the historian who has expressed the strongest opposition to the view that the solution of linear indeterminate equations was unknown to the Greeks is B. L. van der Waerden. In his view "It is true that the solution of linear Diophantine equations is not found in any Greek text. On the other hand, the method is based on the Euclidean algorithm. As we shall see presently, the Greeks were able to solve much more difficult number-theoretical problems by means of the Euclidean



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 237-247.

algorithm, namely problems connected with the rational approximation of irrational ratios and with 'Pell's Equation'. Excellent Greek mathematicians like Archimedes or Apollonios would have no difficulty in solving linear Diophantine equations by means of the Euclidean algorithm."[6] As regards the question of the origins of the linear Diophantine problems, he writes: "I believe that in an ancient Greek source or in a pre-Greek tradition problems giving rise to linear Diophantine equations were treated systematically by means of the Euclidean algorithm, and that this ancient tradition was the common source of all later treatments."[7]

The first historian who explicitly stated the view that the Greeks knew how to solve first-degree indeterminate equations and, moreover, that their study was the subject of logistic (λογιστική), is A. M. Eganyan. He has noticed that two problems in the Akhmîm papyrus (6–7th century), in which it is required to express the result of a division of the form $a : bc$ into the sum of aliquot parts, are reduced to the solution of equations of the form

$$b = mc \equiv 0 \pmod{a}.$$[8]

We will see below that the same holds for one more problem of the same papyrus.

These remarks cast doubt upon the traditional view that the elaboration of methods of solving linear indeterminate equations should be attributed exclusively to the Chinese and Indian mathematicians. In our paper we are going to survey the instances of linear indeterminate problems that occur in Greek sources from Antiquity to the Byzantine era. We intend also to consider the question of the place of the Diophantine tradition in Greek mathematical science. Our main thesis is that Diophantine analysis was subject-matter of the logistic. This view widens the concept of logistic, which is now understood as embracing also the treatment of Diophantine equations and enables us to get a more adequate understanding of the role that logistic has played in the development of the so called numerical algebra.

Diophantine equations are today studied by many branches of mathematics. But the structure of mathematical science, like of any other science as well, is historically determined. In the course of history we are faced, in any given period, with the formation of new branches of science, the gradual transformation of the structure of already existing ones, or even the obsolescence of some others. In Greek mathematics the solution of Diophantine equations was, primarily a subject of logistic. This thesis, which, as we have already said, is not new in the historiography of Greek mathematics, fits with the fact that to this field is generally assigned the most significant work on Diophantine analysis in Antiquity, i.e. Diophantus' *Arithmetica.*

The inclusion of *Arithmetica* to logistic is not an arbitrary choice. It is a view accepted by many historians and scholars of Diophantus' work (Tannery, Heath, Vogel, Klein, etc.) and is based on evidences and on the comparative examination of this work with other survived works on logistic.

An instance of evidence is the anonymous commentary entitled *On Arithmetic* [Περὶ 'Αριθμητικῆς] which in some manuscripts precedes Nicomachus' *Introductio Arithmetica* as a sort of *Prolegomena.* The author of this work makes an important distinction concerning the character of Nicomachus' work in comparison to that of Diophantus. According to his view, Nicomachus' *Arithmetica* studies the "measuring number" [*τὸν μετροῦντα ἀριθμόν*], i.e. the abstract number and its properties, whereas Diophantus' work treats the "measured number" [*τὸν μετρούμενον ἀριθμόν*], i.e. the number counted off on concrete objects.[9]

The affinity of Diophantus' *Arithmetica* to logistic is, finally, confirmed by the comparative analysis of their content to other works of logistic, like the arithmetic epigrams of the *Greek Anthology.* The analysis shows that, in some cases, behind the different formulation lies one and the same content. This fact has repeatedly stressed by ancient commentators of the epigrams. For example, the problem I, 2 of *Arithmetica* states: "To divide a given number into two having a given ratio.[10] This problem, which is equivalent to the system.

$$x + y = a,$$
$$x : y = b,$$

coincides in fact with the following epigram of the *Anthology*: "What violence my brother has done me, dividing our father's fortune of five talents unjustly! Poor tearful I have this fifth part of the seven-elevenths of my brother's share. Zeus, thou sleepest sound."[11] The affinity between these two problems has been noticed by the anonymous commentator of the epigrams, who writes: "The problem is solved in the same manner like the second problem of the first Book of Diophantus, in which it is required to divide a given number, in our case number 5, into two numbers, such that the one of them to be the fifth part of the seven-elevenths of the other."[12]

The characterization of Diophantus' *Arithmetica* as a work of logistic, should not lead however to underestimation of its new features in comparison to the other known works of logistic. The most important such features are the abstract character of formulation of the problems, the application of algorithmic methods with high degree of generality and effectiveness and, finally, the rich algebraic apparatus which also embraces the methods for solving the Diophantine equations.

On the aforementioned grounds we can argue that the study of indeterminate problems in Antiquity was part of the subject-matter of logistic. This position, which also concerns the first degree indeterminate problems, is confirmed by instances of such problems that are contained in survived works of the ancient and the Byzantine periods.

The first degree indeterminate problems that occur in ancient Greek and Byzantine texts can be classified into the following groups:

1. *Problems of some special form*, the solution of which has nothing to do with the solution of the linear indeterminate equation in two unknowns. Such problems are, for example, the two generalizations of the "bloom of Thymaridas", mentioned by Iamblichus in his *Commentary on Nicomachus' "Introductio Arithmetica"*, which are equivalent to systems of the form:

$$x = y = a(z + u)$$
$$x + z = b(u + y)$$
$$x + u = c(y + z).$$

The solution of these systems which is described by Iamblichus is based on the method of the "bloom of Thymaridas".[13]

Two more examples of indeterminate problems of special form are contained in the arithmetic epigrams of the *Greek Anthology*, but their solution is rather trivial.[14]

2. *The problems of the "hundred fowls"*. These problems, which have a long history in the Greek logistic tradition, trace back to the 2nd century A.D. Indeed, in a Greek papyrus of that time is stated a problem which is equivalent to the system:[15]

$$x + y + z = 100$$
$$10x + 20y + 30z = 2500.$$

It should be stressed that the solution of this problem amounts, to the solution of an equation of the form

$$ax + by = c$$

*The solution of this equation is the oldest evidence of a solution of a linear indeterminate equation in the entire history of mathematics.* The method applied for this solution is a kind of solution by inspection.

But this is not the only instance of a problem of the "hundred fowls" that occurs in Greek mathematics. Four similar problems with their solutions are contained in a collection of problems on logistic dated from the 15th century.[16]

The same collection contains also a problem which is equivalent to a linear indeterminate equation in three unknowns,[17] the solution of which has unfortunately been lost.

3. *The remainder problems.* The most known example is the problem of *Codicis Cizensis*, ascribed to Isaak Argyros, which is contained in an appendix of Hoche's edition of Nicomachus' *Introductio Arithmetica.*[18] This problem amounts to the system:

$$\begin{aligned} N &\equiv 1 \pmod 3 \\ N &\equiv 3 \pmod 5 \\ N &\equiv 0 \pmod 7. \end{aligned}$$

Another instance, which is contained in the aforementioned Byzantine collection, is equivalent to the system:

$$\begin{aligned} N &\equiv 1 \pmod m \\ N &\equiv 0 \pmod 7 \end{aligned}$$

for $m = 2, 3, \ldots, 6$.[19] This problem is widely known in the history of mathematics, since it is mentioned by many medieval mathematicians, like Bhaskara I, Ibn al-Haytham, Fibonacci etc.

Besides the above problems, in the history of Greek mathematics some other problems also occur, the solution of which presupposes the capability to solve the first-degree indeterminate equation in two unknowns. The Akhmîm papyrus, for example contains three such problems (designated by the numbers 18, 39 and 40 in the Baillet edition)[20] in which it is required to express the results of divisions of the form $a : bc$ where $bc > a$, $(a, b) = (a, c) = 1$, into a sum of aliquot parts. This is attained by means of the formula:

$$a : bc = \frac{1}{c\,\dfrac{b + mc}{a}} + \frac{1}{\left(b\,\dfrac{b + mc}{a}\right) : m}$$

where $m$ is a positive integral number, greater than 1, such that

$$b + mc \equiv 0 \pmod{a}.$$

In the papyrus the method of finding $m$ in this equation is not explained. May be it was found by inspection.[21]

Another problem for the solution of which the first-degree indeterminate analysis was probably been made use of is, as we have already said, Archimedes' *Cattle Problem.*

Finally, strong evidences, already noticed in the historiography of Greek mathematics, in favour of the view that the Greeks were able to solve linear indeterminate equations even during the classical era, is the knowledge of the algorithm of "anthyphairesis" and the interest in the problems related to the "Great Year", that is the time interval between two successive conjunctions of the planets. The interest in the "Great Year" explicitly manifests itself in the fragments of the pre-Socratic philosophers, in passages of Plato's Dialogues, in the fragments of the Stoic philosophers etc.

The existence of such cases of first degree indeterminate problems as well as the identification of logistic as the branch of Greek mathematics which studies these problems, has drawn us to search for other traces of linear indeterminate analysis in texts of ancient authors, not necessarily mathematicians, relevant to Greek logistic. Such traces can, finally, be found in Plato's *Laws* 819b 4–6, where it is described, in our view, the remainder problem, in the form of a simple exercise on distribution of apples or crowns among children. In the passage 819b-c Plato describes three mathematical exercises in which, according to the Egyptian viewpoint on learning, the students get acquainted with counting. The text is generally considered as obscure. The first exercise is stated as following:

First, as regards counting, lessons have been invented for the merest children to learn, by way of play and fun, as for example modes for distributing [a given number of] apples and crowns, to groups [of children the multitudes of which are] at the same time greater and lesser than the same proper numbers.

According to the most of the philologues and translators of the *Laws* the matter of this lesson concerns the operation of division between positive integers. In particular, the problem is reduced by them to the finding of the quotients obtained as the results of divisions of a given number $N$ (the number of apples or crowns) by certain divisors (the number of students in each group). In our view, however, the above described problem is just the converse one. What is required here is to find out the number $N$ on the basis of the quotients obtained by its divisions by given numbers. In support of our point of view we quote the following commentary of Athenaeus, from the 15th book of his *Deipnosophists* (670f–671a):

But more: the most saintly Plato, in the seventh book of *Laws*, propounds a problem relating to wreaths which is worth solving; the philosopher says: 'They distribute apples and wreaths, the same number being used to fit a larger or smaller number of persons.' These are Plato's words. What the means is something like this: One number, of apples or wreaths, is to be found such that all shall have an equal quantity down to the last person who enters.[22]

Athenaeus takes the number 60 for the required number of crowns which should be distributed, successively, into the first, the second, the third, the fourth, the fifth and the sixth person who enters in a banquet so that, at any moment, the number of crowns shared by each person to be the same (671a-b). In this case, the problem is reduced to the determination of a number $N$ such that,

$$N \equiv 0 \pmod{m},$$

when $1 \leqq m \leqq 6$. This problem is, of course, a special case of the remainder problem and the number $N$ coincides with the Least Common Multiple of the modules of the congruences.

Concerning Athenaeus' interpretation it should be noted:

1. According to Athenaeus, the problem is to find out the number $N$, which should satisfy some conditions related to the operation of division. We will consider below what these conditions are. It should be stressed here, however, that, according to Athenaeus, the problem concerns neither the finding of the quotient of a division (when the dividend and the divisor are given), nor the finding of the divisors of a given positive integer $N$, but the determination of $N$ on the basis of the quotients obtained by the divisions of it by some given numbers. This problem is connected with the following two problems of number theory:

**P1.** With the finding of the Least Common Multiple of some positive integers (when the divisors of $N$ are given).
**P2.** With the problem of remainders (when the remainders of the divisions are also given).

2. According to Athenaeus' interpretation, the proper numbers [ ἁρμόττοντες ἀριθμοί] mentioned in Plato's passage, are the divisors of $N$. This interpretation fits to the meaning of the verb ἁρμόττω and the participle ἁρμόττων which means appropriate, proper, convenient.
3. For the finding of $N$ one only division is not enough. It is necessary to make "distributions" [διανομαί], that is at least two divisions.

As we have already said, the mathematical interpretation for the Platonic text which Athenaeus suggests is the above problem P1. In our view, however, the Platonic text admits a more general interpretation, which is just the problem P2. Such an interpretation does not contradict the philological examination of the text from which follows:

a. The dative cases "greater and at the same time lesser" (πλείοσιν ἅμα καὶ ἐλάττοσιν), are adjectives which refer to the implicit noun "children" ([παισί]) and what is meant is that the distributions are carried out into greater and, at the same time, lesser groups of students.

b. Taking into account, that both the verbs διανέμω and ἁρμόττω can be followed by dative case, we can conclude that the words "greater and lesser", which occur in the dative case, may be grammatically and semantically connected with both the genitive absolute ἁρμοττόντων ἀριθμῶν τῶν αὐτῶν (i.e. greater and lesser than the same proper numbers) and the "distributions" [διανομάς] i.e. to children which at the same time are greater and lesser). In the first case, the distributions are taken to get carried out according to the proper numbers. This yields Athenaeus' interpretation. In the second case, the adjectives "greater and lesser" must have another term of comparison: greater and lesser *than*... For this term can be taken the "one and the same proper numbers". In this case, the divisions should be understood rather as incomplete, i.e. yielding a remainder. Then the problem P2 should be taken as adequate interpretation of the Platonic text, that is the problem of remainders.

According to the second interpretation, the exercise described by Plato can be reconstructed as follows: The teacher suggests to the students to find the number $N$ of apples or crowns, if they knew how many apples or crowns remain each time, when they are distributed into some groups of students, each of which is composed by a number of students that does not coincide with the divisors of $N$, but is greater and lesser than them.

In modern terms the problem can be stated as follows: Let $N$ be the sought number (that is the number of apples or crowns) and let $d$ be a divisor of $N$ (that is a proper number). Then the problem is reduced to the finding of the solution of the system:

$$N \equiv r \pmod{m}$$
$$N \equiv s \pmod{n}$$

where $m$ and $n$ are positive integers) that is the number of students of each group), such that $m < d < n$ (that is, *at the same time* greater and lesser than the proper number $d$) and $r$, $s$ are the remainders which would be given by the teacher.

This interpretation confirms our view that the Greeks were able to solve linear indeterminate equations, to which we have come based on the appearance of the aforementioned instances of such equations.

In the traditional historiography of Greek mathematics, the relation between logistic and arithmetic is commonly viewed through the prism of the relation between practical arithmetic and number theory. In our paper we have made an attempt to show that this scheme is not adequate. It does not fit to the content of the Greek logistic as a whole, since the latter includes also the treatment of the indeterminate problems of the first as well as of higher degrees. The inclusion of the indeterminate analysis under logistic, enriches the content of the latter and throws new light on its role as a source of the numerical algebra.

*Greek Naval Academy, Athens*

NOTES

[1] Concerning the question of priority and the supposed influence of the one mathematical tradition on the other, numerous discussions took place in the past, which were not free from nationalistic prejudices of the participants in them. Today it is known that the methods used by the Chinese and the Indian mathematicians in order to solve the first-degree indeterminate equation in two unknowns are different. For a detailed examination of this question see, Libbrecht, U.: *Chinese Mathematics in the Thirteenth Century*, pp. 213–381.

[2] Libbrecht in the aforementioned work makes a general survey of the development of the first-degree indeterminate analysis outside China (Libbrecht: *Op.* cit., pp. 214–266), but concerning Greece he mentions only one instance of first-degree indeterminate problem

which is of the form of the remainder problem, belongs to the Byzantine era and is ascribed to the monk Isaak Argyros (1318/1372).

[3] Heath, T. L.: *The Works of Archimedes*, pp. 320–323.

[4] Fowler, D. H.: *The Mathematics of Plato's Academy. A New Reconstruction*, pp. 60–62, and also *Archimedes' Cattle Problem and the pocket calculating machine.*

[5] Tannery, P.: La perte des sept livres de Diophante. *Mémoires Scientifiques*, vol. 2, p. 85.

[6] Van der Waerden, B. L.: *Geometry and Algebra in Ancient Civilizations*, p. 133.

[7] *Ibid.*, p. 134.

[8] Eganyan, A. M.: *Grescheskajia Logistika*, pp. 194–196.

[9] Tannery, P. (Ed.): *Diophantus Alexandrinus Opera Omnia,* vol. 2, p. 73, l. 20–28.

[10] *Op. cit.* n. 9, vol. 1, p. 16, l. 24–25.

[11] *Op. cit.* n. 9, vol. 2, p. 61, l. 22–25. English translation in Paton W. R. (Ed.): *The Greek Anthology*, vol. 5, epigram No 128.

[12] *Op. cit.* n. 9, p. 62, l. 2–6.

[13] Iamblichus: *In Nicomachi Arithmeticam Introductionem liber*, Ed. H. Pistelli, pp. 63–68.

[14] Heath, T. L.: *A History of Greek Mathematics*, vol. 2, p. 443.

[15] Winter, J. G. (Ed.): Papyri in the University of Michigan collection. Miscellaneous papyri. *Michigan Papyri*, vol. 3, pp. 34–52.

[16] Hunger, H. and Vogel, K. (Eds.): *Ein Byzantinisches Rechenbuch des 15. Jahrhunderts*, pp. 44–47.

[17] *Ibid.*, pp. 52–53.

[18] Nicomachus: *Introductio Arithmetica*, Ed. R. Hoche, pp. 152–153.

[19] *Op. cit.* n. 16, pp. 72–73.

[20] Baillet, J. (Ed.): Le papyrus mathématique d' Akhmîm, pp. 72–73 & 83–84.

[21] Another interpretation of the same problems has been suggested in, Knorr W. R.: Techniques of fractions in ancient Egypt and Greece. *Historia Mathematica*, vol. 9, 1982, pp. 133–171.

[22] Athenaeus: *The Deipnosophists*, Ed. Ch. Barton Gulick, vol. 7, pp. 92–94.

## REFERENCES

Athenaeus, 1961, *The Deipnosophists* (Ed. Ch. Barton Gulick) 7 vols. (The Loeb Classical Library). First printed 1927–1941.

Baillet, J. (Ed.), 1892, Le papyrus mathématique d' Akhmîm. *Mémoires publiés par les membres de la Mission Archéologique Française au Caire*, vol. IX, fasc. 1. Paris: Ernest Leroux.

Diophantus, 1893–1895, *Diophantus Alexandrinus Opera Omnia* (Ed. P. Tannery) 2 vols. Leipzig: Teubner.

Eganyan, A. M., 1972, *Grescheskajia Logistika.* Erevan: Aiastan.

Fowler, D. H., 1980/1981, *Archimedes' Cattle Problem and the Pocket Calculating Machine.* Coventry: University of Warwick, Math. Institute Preprint.

Fowler, D. H., 1990, *The Mathematics of Plato's Academy. A New Reconstruction.* Oxford: The Clarendon Press. First printed in hardback 1987.

Heath, T. L., 1897, *The Works of Archimedes.* Cambridge: University Press.

Heath, T. L., 1981, *A History of Greek Mathematics*, 2 vols. New York: Dover. First printed 1921.

Hunger, H. and Vogel, K. (Eds.), 1963, *Ein Byzantinisches Rechenbuch des 15. Jahrhunderts*. Österreichische Akademie der Wissenschaften. Philosophisch-Historische Klasse. Denkschriften, Bd 78, Abh. 2. Wien.

Iamblichus, 1894, *In Nicomachi Arithmeticam Introductionem liber* (Ed. H. Pistelli). Leipzig: Teubner.

Knorr, W. R., 1982, Techniques of fractions in ancient Egypt and Greece. *Historia Mathematica* **9**, 133–171.

Libbrecht, U., 1973, *Chinese Mathematics in the Thirteenth Century*. Cambridge: The M.I.T. Press.

Nicomachus, 1866, *Nicomachi Geraseni Pythagorei Introductionis Arithmeticae libri II* (Ed. R. Hoche). Leipzig: Teubner.

Paton, W. R. (Ed.), 1960, *The Greek Anthology*, 5 vols. (The Loeb Classical Library). First printed 1916–18.

Tannery, P., 1912–1915, La perte des sept livres de Diophante. *Mémoires Scientifiques* (Eds. J.-L. Heiberg & H.-G. Zeuthen) **2**, 73–90. Toulouse: E. Privat & Paris: Gauthier-Villars. First printed in *Bulletin des Sciences Mathématiques* **8** (1884), 192–206.

Van der Waerden, B. L., 1983, *Geometry and Algebra in Ancient Civilizations*. Berlin/Heidelberg/New York/Tokio: Springer-Verlag.

Winter, J. G. (Ed.), 1936, Papyri in the University of Michigan collection. Miscellaneous papyri. *Michigan Papyri*, vol. 3. Ann Arbor: University of Michigan Press.

IZABELLA G. BASHMAKOVA AND IOANNIS M. VANDOULAKIS

# ON THE JUSTIFICATION OF THE METHOD OF HISTORICAL INTERPRETATION

"The manuscripts do not burn"
M.A. Bulgakov

Today, when all over the world the interest in history of science and the research in this field have grown, the methodological questions of the historical studies have become especially sharp and actual. Around the problem of the advantages and admissibility of interpretations, of the "translatability" of older texts into modern language, whole parties, who hold diametrically opposite views have appeared; let us agree to call them *antiquarists* and *modernists*. The modernists boldly translate the classical texts into the language of modern mathematics, appealing to far finer and more complicated parts of it than the simple language of notations usually used by the authors of the past. The antiquarists, on the contrary, declare that such interpretations are illegitimate, distort the meaning of the text, bring in it concepts and methods alien to it.[1] Who is right?

The translation that is used in the history of science is an instance of translation from one natural language into another, and so it shares of the same advantages and disadvantages like any translation in general. It provides an ample opportunity for those otherwise unable to read older texts in the original to get acquainted with the classical works and, in case of a successful translation, it helps them not only to understand the ideas and the images of the author, but even to enjoy the beauty of exposition, in spite of the well-known irretrievable loss of nuance and inevitable distortion of meanings. The "translations" or, generally, the interpretations in history of mathematics have also their own, specific only to them, advantages, although there is a definite danger behind them.

The admissibility of interpretations in history of mathematics is deeply rooted in the specifics of mathematics itself as a science. We can distinguish two components in it: *formal* and *contentual.*

From a formal point of view, mathematics is an aggregate of deductions and results, evolving in time, for the theorems proved establish facts which remain constant in the subsequent course of history, beginning from the



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 249-264.

most ancient times. There is no mathematical theory or theorem which once has been proved has been rejected or refuted afterwards.

However, one and the same mathematical theory, one and the same mathematical fact can be started in different languages. Thus we can talk in different languages about ones and the same, in essence, objects. This is a feature specific to mathematics.

The Lobachevskian plane, for example, can be defined axiomatically, as the inventor of the hyperbolic geometry does, and its plane geometry can then be developed like in Euclid's *Elements*. For piece of the Lobachevskian plane we can also take the surface of a pseudophere, the geodesic lines of which will play the role of straight lines. It is also possible, for Lobachevskian plane to take the interior of a circle, where its chords (without end-points) would play the role of straight lines and the transformations which translate a circle to itself the role of motions. Other interpretations are also possible. Thus, one and the same theory, i.e. the Lobachevskian geometry, in the considered case, can be stated in different languages, while the essence of the theory remains the same.

It is worth noticing that in history of mathematics we are always faced with "translations" of older theories into new terms and with their subsequent incorporation into the context of the new period of development of science. The mathematicians of every new age reconsider the previous material and restate it in new terms, thus making it readily available and applicable for the contemporary scientist. That is the case with the works of Euclid and Archimedes, which were originally translated into the language of infinitesimals and then into the language of differential and integral calculus, making use of the concepts of upper and lower integral sums, of the differential (characteristic) triangle, of limit etc. All the mathematicians of the 17th century, including the teacher of Newton Barrow, had engaged in translating Archimedes' works, by their attempt to explicate the general methods of calculating areas and volumes, as well as the methods of finding tangents that are present by it. Euler, Lagrange and Gauss have constructed the general theory of quadratic forms, while its main results were "translated" by Kummer, Dedekind, Zolotarev and Kronecker into the language of the theory of fields of algebraic numbers. Consequently, "translations", or rather, interpretations in new terms are encountered in mathematics in every step of its development.

The ontological and metaphysical views on mathematical objects and their relations, the methodological, metamathematical and, in general, philosophical views, as well as the expressive power of the forms of mathematical thought (symbolism etc.) constitute the content of a mathematical theory. This component is liable to change in the course of history.

Various periods of development of mathematics differ from each other in the depth of the ideas and methods elaborated, in their generality (degree of abstraction), in the rigor of the proofs and the foundations of their fundamental concepts and principles, in the effectiveness of the application of their theories in other fields of science, etc. The consideration of the contentual component of mathematical science is absolutely necessary in order to preserve the historical distance between the modern approaches and the older concepts and theories and their specific function in the context of a certain stage of development.

At this level of methodological analysis of the historical process it is possible to study the problem of the formation of new mathematical notions, of the change in meaning of older concepts or the extension of the scope of previously used operations, the problem of the growth of abstraction of mathematical concepts, the appearance of new methods and ideas, as well as changes in the conceptual structure of mathematical knowledge, in the views on the foundations of mathematics, on the standards of mathematical rigor, etc.

In accordance with these peculiarities of the subject-matter, the history of which we study, any investigation in the history of mathematics should, in our view, be carried out in two steps:

First, the text under consideration should be *"translated"* into the prevailing contemporary mathematical language, i.e. an adequate model for it should be constructed. This is absolutely necessary in order to *understand* the text, to reveal its mathematical content. To understand a mathematical text means to know *what is the case* provided that it is true.

In principle, various models can be constructed, since mathematical truths do not depend on the language in which they are expressed. However, this does not mean that the mathematical truths do not interact at all with language: one language can be more suitable for their expression, while another less suitable. For example, more suitable for algebra has turned out to be the language of the literal calculus which was created only in the end of the 16th century, while the history of this science begins from the times of ancient Babylon (19–18th centuries B.C.) and during all this enormous period algebra has repeatedly changed its language. Consequently, the mathematical truths are not reducible to language and history of mathematics to the history of the various languages and their specifics.

The understanding of the material under consideration is carried out by the hypothetico-deductive method. From a pack of permissible interpretations we choose one or we invent a new one in the following way: we suggest a hypothesis what meaning should be assigned to some term of a theory; our hypothesis is further checked taking into account the meaning of the remaining terms of the theory and corrected or refined according to the behavior of the named object within the mathematical theory as a whole.

One might wonder whether we could make a mistaken choice of interpretation, whether we can talk about a "correct" or "incorrect" understanding. Let us assume there is a text. The author of this text has already assigned a certain meaning to it. Must the interpretation of this text by the interpretator coincide with the understanding of the author himself in order to accept his understanding as "correct"? On reflexion it becomes rather clear that such a requirement could not be accepted. The only thing we can expect is that our interpretation would be consistent with all known facts, i.e. the meaning which we assign to separate terms would be consistent with the content of the theory as a whole and the interpretation of the text would fit the other texts of the same author as well as the historical facts of the cultural life of his epoch.

The construction of a model does not serve only to grasp the meaning of a text, but also to communicate it. To explain the meaning of an older theory or concept to somebody else can be done *only in terms familiar to him*, i.e. in terms of the modern metatheory.

As a matter of fact, the first step of any study in the history of mathematics should be "translation" of the considered text into the language of modern science. However, this is but the first step which by itself is not enough. After the current understanding of the text follows a second, more difficult task. It is necessary to *embed* the considered work into the context of science of its own day. The suggested interpretation should be *justifiable* when the text is considered in the context of contentual notions of its own time. We ought to elucidate what studies preceded the considered work, what tasks the science of that time had confronted, how the research on the considered topic was continued and by whom, how the one or the other notion could then be understood, to clear up the question of identification or difference in meaning of the fundamental concepts and conceptual structures under consideration with the modern ones, etc. For example, if the matter concerns the general theory of proportions of Eudoxus–Euclid, then many questions arise concerning how the Greeks could understand the relation between ratios and numbers, when and how the concept of number was first explicitly extended, whether there were other definitions of ratio of magnitudes, what reasons have possibly led to the acceptance of the Eudoxian theory by the Greeks, etc.

Consequently, the primary purely mathematical interpretation of an earlier theory should necessarily be followed by its *"historical-mathematical interpretation"* by the determination of the place of the considered work in that model of development of mathematics which we construct for the given period (or a shorter segment of time). It is necessary for the suggested interpretation of the text to be *effective*, that is to explain much more than only those facts, for the sake of which it had been constructed, to throw light not only on specific aspects of the historical perspective, but also on the general "picture" of development of mathematical knowledge.

An earlier piece of mathematics can be "embedded" in the context of its own time in various ways, depending on its understanding by the interpretator, on its methodological and philosophical views. Correspondingly, various conceptions of history of science are possible and various models of its development. The history of mathematics, considered from the point of view of the history of mathematical truths, presupposes a critical assessment of what has been attained in science in the past from the standpoint of the contemporary science. The "picture" of the course of development cannot be created once and forever, but it should be being reconstructed all the time. Already in the beginning of the current century V. I. Vernadsky wrote that "the understanding of the past of science, even of the 18th century by a scholar of the beginning of the twentieth century sharply differs, in many respects, from the notions, worked out ten years ago. The course of time and the work of scientific thought perpetually give birth to reappraisal of the values in the scientific outlook. The past of scientific thought is depicted each time in a completely different and always new perspective. Each scientific generation discovers in this past new features and loses the settled notions about the course of scientific development. The history of scientific thought, like the history of philosophy, religion and art, can never give a completed, unchangeable picture, really reproducing the factual course of events, since these long ago bygone events appear each time in different light and reflect, in the one or the other way, the contemporary to the investigator state of scientific knowledge." ([17] 191–92). Every new great discovery in mathematics widens the possibilities for interpretations and gives us the opportunity to read afresh the classics of science, in order to see what has been overlooked, misconceived or passed unnoticed, what has not been able to be expressed in the previous interpretations. In this sense the classical works of Archimedes, Apollonius, Fermat, Euler and the other mathematicians are also everlasting and immortal, like the poetic masterpieces of Homer, Shakespeare and Pushkin.

Let us consider two examples illustrating our thesis about the necessity of *two levels of historical analysis.*

## I. DIOPHANTINE EQUATIONS

In Diophantus' *Arithmetica* the foundation of literal algebra was laid. Diophantus introduced a notation for the first six positive, the first six negative powers of the unknown and the zero power, as well as signs for subtraction and equality. He does not present his methods in general form, in the known Greek style, but illustrates them by means of the solution of concrete problems.

In the pursuit of an understanding of these methods, already F. Viète begun to translate Diophantus' problems and methods into the language of literal calculus, which took its shape only in the work of Viète himself. It was Viète who introduced notation for arbitrary constant magnitudes (parameters) and wrote the first mathematical formulas. This "translation" enabled Viète, and Fermat afterwards, to discover the first general methods contained in *Arithmetica.* These methods concern the solution of indeterminate equations and systems in rational numbers. Thus Viète formulated in general form the methods of solution of the second-degree equations of the form

$$y^2 = ax^2 + bx + c^2 \text{ and } y^2 = a^2x^2 + bx + c$$

and found the method of solution of the equation

$$x^3 + y^3 = a^3 - b^3,$$

which, as Diophantus asserts, is always solvable.[2]

Thus the "translation" into the language of literal calculus helped us to understand the first layer in Diophantus' work. Unfortunately, untill recently the historians of mathematics did not apply in the study of

*Arithmetica* any other methods except those of elementary algebra and elements of number theory, the tradition of application of which goes back to Fermat. This has given rise to incomprehension of Diophantus' methods and to a wrong appreciation of his work. Becker and Hoffman, for example, assert that "Diophantus does not give any general method, but it seems that in each case he applies some unexpected artificial method, resembling the Eastern ones" ([6] 90). Analogous views have been stated also by B. L. van der Waerden, H.-G. Zeuthen, M. Cantor, Th. Heath, P. Tannery, G. Wertheim and many others.

Only when concepts and methods of algebraic geometry were used for the interpretation of Diophantus' *Arithmetica* it was made possible to classify Diophantus' problems and to uncover his methods. This was attained by means of a "translation" into another language, more suitable to the methods of Diophantus himself. The analysis has demonstrated that *Arithmetica* is devoted to the systematic exposition on concrete material of highly general methods of solution of indeterminate equations, defining curves of genus 0, i.e. such curves which admit parametrization in rational functions of one parameter. Diophantus' results can be formulated, in modern language, in the following way: if on a curve of genus 0 lies a rational point, then there exist infinitely many rational points on it, the coordinates of which can be expressed as rational functions of one parameter. From the Book IV onwards, curves of genus 1 of orders 3 and 4 are, in essence, also examined and it is shown that if on such a curve lies a rational point, then by means of the "tangent method" it is generally possible to find one more rational point, while if two rational points are given then a third one can be found by the "secant method" (this method is applied by Diophantus when one of the two given rational points is finite and the other infinite). It was also made possible to separate two essentially distinct classes of "double equations of the second degree", i.e. systems of the form

$$a_1x^2 + b_1x + c_1 = u^2,$$
$$a_2x^2 + b_2x + c_2 = v^2.$$

The first of them defines space curve of genus 0, which is shown to be always parametrizable, while the second defines a space curve of genus 1, for which Diophantus applies completely different methods ([18] 2).

Let us consider an example. In Proposition 8 of the Book II of Diophantus' *Arithmetica* the following problem is stated: "To divide a given square number into two squares". After the statement of the problem in words Diophantus writes the condition in his own notation. In absence of elaborated symbolism for the denotation of arbitrary parameters he takes a concrete square, namely 16, for the "given square".[3] Further, he denotes by the symbol of the unknown the side of one of the sought for squares–let us use the symbol $t$ here, although Diophantus uses a different sign. The side of the other square is formed "from any number of $t$ minus as many units as there are in the side of 16" and let it be $2t - 4$". Here, the number 4 is defined by the choice of the given square, while the number 2 plays a double role, as an ordinary number and as a sign standing for arbitrary rational parameter. We find this second function in algebra until the end of the sixteenth century.

In the 16th century Viète introduced symbols to denote arbitrary parameters ($B, C, D, \ldots$) and unknowns ($A, E, J, \ldots$) and stated the same problem in the Book IV of his *Zeteticorum* in another way: he denotes the arbitrary square by $B$quad., the side of the one of the squares by $A$ and the other by

$$B - \frac{\sin A}{R}.$$

In the 17–18th centuries the same problem was formulated by means of the equation

$$x^2 + y^2 = a^2, \tag{1}$$

and the following substitution

$$y = kx - a \tag{2}$$

was made. Thus the interpretation of this problem into the language of the literal algebra is completed. And it is just in this form that has become widely known.

Nevertheless, another interpretation, which is closer to the geometric style of thinking of our century, is also possible. The equation (1) can be considered as defining a circle and the substitution (2) as the equation of the straight line passing through the point $(0, -a)$ of its circumference. It is then clear that the straight line (2) will intersect the circumference (1) in one more point $(x, y)$, the coordinates of which must be rational. Then

$$a^2 = t^2 + (kt - a)^2,$$

or

$$x = t = \frac{2ak}{1 + k^2}$$

$$y = kt - a = \frac{k^2 - 1}{k^2 + 1}\, a$$

(in Diophantus $x = \frac{16}{5}$ and $y = \frac{12}{5}$).

The geometric interpretation is suitable for the study of the problem how general Diophantus' method is and whether he finds *all* the solutions of it, or only part of them.

It is noteworthy that the interpretation by means of algebraic geometry has further been used by A. Weil [20] in his analysis of certain results of Fermat and Euler as well as by R. Rashed [14].

An analogous interpretation by means of the language of algebraic geometry was also used for the study of the works on Diophantine analysis of the mathematicians of the Arabic East. Here new results were obtained again. Let us discuss one of these. Abū Kāmil (c. 850–930) in his *Book about algebra and al-muqābala* lists 38 problems which are reduced to indeterminate equations. Moreover, the problems, which are equivalent to the finding of rational points on plane curves, were alternated with analogous problems for space curves, so that the principle, according to which the sequence of the problems was made up remained obscure. E. I. Slavutin, by means of "translation" into the language of algebraic geometry, has established the following astonishing fact: the first 25 problems, both plane and space, concern methods of finding rational points on algebraic curves of genus 0, while the next 13, on curves of genus 1 ([2] 132–35). Of course, Abū Kāmil *had not* the modern notion of genus of an algebraic curve at his disposal, but he was, seemingly, aware of the difference between the two classes of problems, maybe, in the sense that the curves involved in the problems 1–25 are parametrizable in rational functions, while the curves involved in the next 13 problems are of completely different nature: it is impossible for them to specify the whole set of infinitely many rational points, but only to provide a method of specification of one more point out of already known rational points. Once again we express a thought which we ascribe to Abū Kāmil into a language unusual to him – into the language of geometry. This very same thought (if he had ever had such a thought in mind) might have been started by him in a purely algebraic language, i.e. he might have talked not about curves and points lying on them, but about indeterminate equations and their rational solutions.

The work of Diophantus had such a far-reaching significance for the development of algebra and number theory like the works of Archimedes for the development of infinitesimal calculus. Unfortunately, we know nothing about the predecessors and the successors of Diophantus up to the 9–10th centuries. Undoubtedly, the algebraic tendency, the most eminent exponent of which would be Diophantus, had other exponents too in the science of the subsequent centuries. Traces of this influence are quite evident in the Arabic mathematical tradition of the 9th century, after Qusṭā ibn Lūqā's translation of Diophantus' *Arithmetica* into Arabic.

As regards the development of algebraic geometry itself, with which the work of Diophantus is now associated, one can trace the influence of the ideas of the great algebraist of antiquity on it up to the beginning of the current century. After Fermat the methods of Diophantus are found in Euler, Jacobi and other mathematicians. However, the deepest application they assume in the works of H. Poincaré, in particular, in his work *Sur les propriétés arithmétiques des courbes algébrique* (1901) where he studies the algebraic structure of the set of rational points of algebraic curves [1]. It should be noted that the methods of Diophantus were also understood and further elaborated by P. Fermat.

These investigations on the history of Diophantine equations enable the reconsideration of the history of algebra. It turns out that the mainsprings of the development of algebra are not only problems, which have been expressed by definite equations, but also the study of and solutions to Diophantine equations [3–5].

## II. PLATO'S THEORY OF FORMS

In Plato's *Parmenides* the widely known Third Man Paradox is stated which has special interest for the history of logic, because of the phenomenon of self-reference that is involved in it. The problem of interpretation of this paradox has raised many discussions, especially in the 50's, and the literature devoted to it has become enormous. However, hitherto many points have remained obscure in Plato's reasoning. Already in 1954 G. Vlastos, who has made a considerable effort to clear it up, noted that "if any progress is to be made at this juncture it must come from some advance in understanding the logical structure of the argument" [19].

Hitherto it has been generally believed that the paradox was simply stated by Parmenides, but not really solved and, even more, that Plato scarcely could have solved it, since he confused the categories of substance and attribute (see, for example [11] 19). Russell also argued that Plato violated in his arguments the restrictions imposed on language by the theory of logical types. This view is encouraged by the linguistic difficulties which Plato has faced in his attempt to formulate an ontology of abstract entities, i.e. that in Greek language abstract and concrete terms are formally indistinguishable: τὸ λευκόν (literally 'the white') may signify both 'the white thing' and 'whiteness"' [*op. cit.*]

The root of this misconception, in our view, largely stems from the fact that in English literature one and the same word, 'Form' or 'Idea', is commonly used interchangeably to render a whole cluster of Greek terms, i.e. *eidos, idea, ousia, genos* which, although close in meaning are not identical in any given context. In particular, it is common faith that no significant difference exists between the Platonic terms *eidos* and *idea*, which are usually rendered by the self-same word. However, a closer examination shows that this is not so. Plato himself makes a quite clear distinction between these two terms.

In *Theaetetus,* for example, Plato declares that "an *eidos* is a whole that has generated out of parts, but is different from the sum-total of the parts": τὸ ὅλον ἐκ τῶν μερῶν λέγεις γεγονὸς ἕν τι εἶδος ἕτερον τῶν πάντων μερῶν (204a), whereas *idea* (ἰδέα) is characterized as 'indivisible': ἀμέριστος (205c), as an integral entity that "has no parts": μέρη μὴ ἔχει καὶ μία ἐστὶν ἰδέα (205d). This corresponds to a fundamental distinction, which has taken on many shapes throughout the history of logic, i.e. the distinction between *class-as-many*, distributed to its elements by means of prediction, and *class-as-one*, standing as an "individual in an extended sense" and capable of being an element of a further "class-as-many" (Russell), or between the notions of *distributive* and *collective class* (Leśniewski), or between *extension* and *intension* (Carnap), or between *singular* and *general terms* (Quine).

Regardless of the language used, the underlying idea is quite clear. The entities named by *beautiful* and *the beautiful* are in this sense 'associated', but they are not equivalent in terms of logical function. Plato seems to be aware of this distinction when in *Hipp. Maj.* Socrates corrects Hippias and makes clear to him that what they are seeking for is the definition of the second entity, but not the first: "not what is *beautiful*, but what *the beautiful* is": Οὐ τί ἐστι καλόν, ἀλλ' ὅ τί ἐστι τὸ καλόν (287d).

In *Phaedr.* the contrast between *eidos* and *idea* once more strikingly manifests itself in connection with the problem of "divisibility". Plato's suggestion is "to 'divide' entities 'in kinds', and to 'comprise' each single [kind of them] by [some] one '*idea*': κατ' εἴδη τε διαιρεῖσθαι τὰ ὄντα καὶ μιᾷ ἰδέᾳ ... καθ'ἕν ἕκαστον περιλαμβάνειν (273e). Division, which is the main task of dialectics, is always carried out by Plato 'in kinds' (κατ' εἴδη), but never 'into *ideas*' (κατ' ἰδέας): e.g. *Phaedr.* 265d-e, 273e, 277b-c; *Soph.* 264c. On the other hand, the term περιλαμβάνειν ('comprise') which is associated with the *idea* bears a quite clear collective shade of meaning, which is apparent elsewhere too, by the use of such terminology like ἄγειν ("bringing-together"), ἐπιφέρειν ("bringing-on"), etc. For example, in *Phaedr.* 265d he suggests "'bringing-together' a plurality of scattered particulars to one *idea*": εἰς μίαν τε ἰδέαν συνορῶντα ἄγειν τὰ πολλαχῇ διεσπαρμένα, while in *Resp.* a plurality of things is 'brought-on' to a common name: εἰς κοινὸν ὄνομα ἐπιφέρομεν (596a). In general, each of these terms – *eidos* and *idea* – is associated with definite grammatical constructions and linguistic phenomena and there is no instance of their confusion by Plato.[4]

This is connected with a distinction between two types of predication which correspond to "participation" (μέθεξις), that is to "membership", and to "communion of *eidōn*" (κοινωνία τῶν εἰδῶν), that is to "inclusion", respectively, although both are expressed by the same verb "to be" (ἐστιν).

Consequently, Plato's theory of Forms appears to have a straightforward three levelled ontology, with an implicit three level semantics. Without accepting this three level semantic picture it would be difficult to make sense of many otherwise conflicting propositions of Plato and, in particular, it seems scarcely possible to understand the Third Man Paradox.

This paradox is obtained, speaking in modern terms, as follows:[5] at first, out of *all* the things of an initial domain of particulars to which the property '... is large' *(idea)* applies is made a class (referred to by the term "the large" – τὸ μέγα), i.e. Parmenides makes, in fact, use of an intensional method to define "class", since he uses the concept of *idea* which bears a rather intensional shade of meaning. Then *this* class (αὐτὸ τὸ μέγα) is added to the initial domain of individuals and the scope of the universal quantifier "all" (πάντα) is extended over it, taken for individual. The construction results in the impredicative generation of a (potentially) infinite sequence of new classes and, correspondingly, in a constantly expanding ontology (131e-132b).

The source of the paradox lies in the double role that the intermediate entity αὐτὸ τὸ μέγα plays: on the one hand, it plays the role of "class-as-many", distributed to the particulars which participate in it by means of predication (in virtue of an *idea*), and, on the other hand, it plays the role of "class-as-one", admitted to quantification, which gives rise to the impredicative generation of a new "class-as-many". Under this *confusion* between "class-as-many" and "class-as-one" the individuation of *eidōn*, which is expressed in Plato by phrases like ἕν εἶδος (a *single* kind), ἕκαστον τῶν εἰδῶν (each of the kinds *singly*) etc., ceases to wind down.

After an unsuccessful attempt to solve the paradox (132c-d) in terms of a semantics that partly echoes the historical Parmenides, Plato puts the following solution into Parmenides' mouth.

The new semantic picture is introduced by Socrates in 132c-d. The *eidos* is defined as a *paradigm* (παράδειγμα) which expresses the form of instances (ὁμοιώματα –"copies", "images") of the *eidos*, considered as a singular thing 'found' in nature (ἐν τῇ φύσει). In other words, the *eidos* here can be considered as singular abstract term denoting an abstract entity which is admitted to quantifiers, since it is 'found' in nature, and the "copies" as instances of this entity. Moreover, "participation" in an *eidos* is identified with "instantiation" (εἰκασθῆναι) of the *eidos*. This does not prevent self-reference yet, since, intensionally considered, the paradigm can, in principle, become an instance of a further paradigm. This possibility is excluded below.

Further, the *eidos* is compared with a fixed instance of it (τὸ εἰκασθέν), that is a two-place predicate "__ is similar to ..." (ὅμοιον εἶναι) is introduced and the following question is posed: can we conclude that *an eidos is similar to an instance of it* (ὅμοιον τῷ εἰκασθέντι) *on the basis that the latter is an "instantiation" of the eidos* (καθ' ὅσον αὐτῷ ἀφωμοιώθη)? (132d-e).

Parmenides elicits Socrates' naive agreement which, in fact, does not prevent self-reference. However, Parmenides corrects Socrates and suggests another treatment of similarity that leads to a negative answer to the above question. He defines that entities are similar to each other if and only if they participate in one and the same *eidos* (132e). In this way, Parmenides takes, in essence, into account what we today call *domain* of the class (that is the domain which is formed by the "participants" of the *eidos* or, equivalently, by the instances of the paradigm). This "domain", obviously, consists of homogeneous things ("similar" to each other). This amounts to a restriction on the formation of *eide* (now their domain should be taken into account), from which immediately follows the next conclusion of Parmenides that neither the thing is similar (homogeneous) to the *eidos*, nor the *eidos* to another thing that participates in it (ἄλλῳ).

The last conclusion is demonstrated by *reductio ad absurdum* (132e-133a).

Therefore, Parmenides makes a quite clear demarcation between two kinds of homogeneous entities: the level of particulars and the level of *eidōn*, and the confusion between them is removed *ad hoc*. The lower level is defined by the possibility of establishment of the relation of similarity (dissimilarity) among its individuals, while the relation of the individuals of the lower level to those of the higher level, i.e. to the *eide*, is now defined by "participation" (133a), that corresponds to the modern relation of "belonging". Moreover, Plato uses elsewhere too a language which suggests that the *eide* exist separately (χωριστά) from and have higher ontological status than the particulars. This stratification and the corresponding restriction on the formation of *arbitrary* Forms (*eide*) removes the paradox from the Platonic ontology. Furthermore, the fact that the above solution is suggested by Parmenides, who has actually posed the question gives us grounds to believe that the matter does not concern any critics of Plato's theory of Forms, but rather a dialectical inquiry over various semantic conceptions of antiquity.

It is known that the distinction in types between classes and their elements was suggested by B. Russell as a solution to his famous paradox about the class of all classes which are members of themselves, that seemingly was discorvered in 1899 [9] 33–46.

Schröder's introduction of level distinctions in 1890, and Peano's distinction between membership and inclusion showed that the idea of stratification of objects in levels was not entirely original with Russell.

Peano's distinction, however, was but a reification of an old scholastic tradition, as is explicitely stated by himself. In fact there was a late scholastic theory of types, as L. Hickman recently demonstrated [10], which was connected with problems of simply nondescending predicates. The scholastics had elaborated an informal technique of elimination of paradoxes involving self-reference (the so called paradoxes of descent), which was known to Peano and was based on the distinction in types of predication corresponding to two senses of the verb "to be": one *in sensu compositi* and another *in sensu divisi*.[8]

Hickman's examination of the scholastic variety of theory of types has led him to the conclusion that the "simple theory of types was in the offing from the time of Plato and Aristotle" ([10] 180). The historical problem of the formation of type-theoretic notions in ancient Greece is a quite sophisticated one and is discussed in certain detail in [15]. Their appearance has apparently been determined by a whole spectrum of scientific and cultural factors. In particular, it seems to be connected with logical problems inherent in Parmenides' metaphysical conception, with questions of definitions which were, probably, first posed by Socrates and further discussed in the Socratic school and Plato's Academy, with the sophists' inquiries on language, its grammar and semantics and their development in Plato's Academy, as well as with foundational problems concerning infinite processes and their ontological status, and the structure of the continuum which have been arisen out in mathematics by the discovery of incommesurability behind of which lies an infinite process of seeking for the common measure of the diagonal and the side of the square.

Summing up, we could not but agree with M. A. Bulgakov's declaration that "the manuscripts do not burn" which has served as a motto in our paper. Neither are the live ideas exposed in the classical works of the past, but their language grow antiquated and needs refreshment in order to revive them.

*University of Moscow*

## NOTES

[1] It seems no mere chance that such discussions, on the one or the other form, have been taking place within all the historiographic traditions. We are reminded, for example, of the ideas of Bachelard and Canguilhem on the question of breaks and succession in history of science and the critics by the latter of the "virus of predecessorship", the dispute between I.-G. Zeuthen and M. Cantor on geometric algebra, an issue that has revived in our days, or the controversy between the "antiquarists" and the "modernists" in the Russian historiography – hence our terminology.

[2] R. Bombelli found the solution of this equation when $a = 4$, $b = 3$.

[3] R. Bombelli in the solution of the same problem takes 25 for the given square. From the solution it is clear that it is possible to take *any* square of rational number.

[4] The distinction between *eidos* and *idea* in Plato seems to have passed almost completely unnoticed. For the first time it was suggested in 1930 by the Russian philosopher and keen student of the Greek terminology and culture A. F. Losev (1893–1988) in a work devoted to the systematic examination of these terms throughout the Platonic corpus and the detailed classification of their various meanings [12].

Ten years later P. Brommer, seemingly independently from Losev, was also led to analogous conclusions [7] – see also Losev's review of Brommer's book in ([13] Vol. 4, 741–45).

[5] A detailed logical reconstruction of the argument is carried out in [16].

[6] In fact, these two terms have a much wider range of meanings in medieval logic than these, which are mentioned by Peano.

## REFERENCES

1. Bashmakova I. G., 1981, Arithmetic of algebraic curves from Diophantus to Poincaré. *Historia Mathematica* **8**, 393–416.
2. Bashmakova I. G., and Slavutin E. I., 1984, *History of Diophantine Analysis from Diophantus to Fermat.* Moscow: Nauka. [In Russian].
3. Bashmakova I. G., 1984, Indeterminate equations and their role in the development of algebra. *Voprosy Istorii Estestvoznanija i Tekhniki* **2**, 43–56. [In Russian].
4. Bashmakova I. G., 1986, The main stages in the development of algebra. *Istorija i metodologija estestvennykh nauk* **32**, 50–65. [In Russian].
5. Bashmakova I. G., 1987, Diophantine equations and the evolution of algebra. *Intern. Congr. of Math., Berkely, California,* pp. 1612–1629. [In Russian]; English Transl. by A. Shenitzer and H. Grant: *Amer. Math. Soc. Transl.* (2) Vol. **147**, 1990, pp. 85–100.
6. Becker, O. and Hoffman J. E., 1957, *Geschichte der Mathematik.* Bonn.
7. Brommer, P., 1940, *Eidos et idea. Etude sémantique et chronologique des oeuvres de Platon.* Assen.
8. Burnet J. Ed., 1899–1906, *Platonis Opera.* 5 Vols Oxford.
9. Drucker T. Ed., 1991, *Perspectives on the History of Mathematical Logic.* Boston: Birkhäuser.
10. Hickman L., 1980, *Modern Theories of Higher Level Predicates. Second Intentions in the Neuzeit.* München: Philos. Verlag.
11. Kneale W. & M., 1984, *The Development of Logic.* Oxford: Clarendon Press.
12. Losev A. F., 1930, *Essays of Ancient Symbolism and Mythology.* Moscow. [In Russian].
13. Losev A. F., 1963–88, *History of Ancient Aesthetics.* 7 Vols, Moscow: Iskusstvo [In Russian].
14. Rashed R. Ed., 1984, *Diophante, Les Arithmétiques,* t. 3, 4. Paris: Les Belles Lettres.
15. Vandoulakis J. M., 1991, *On the Formation of the Mathematical Science in Greek Antiquity.* Ph. D. Thesis (Moscow University) [In Russian].
16. Vandoulakis J. M., 1982, Plato's Anticipation of the Simple Theory of Types. *Istoriko-matematicheskie issledovanija* **35** (in print) [In Russian].
17. Vernadsky V. I., 1981, *Selected Works on the History of Science.* Moscow: Nauka. [In Russian].
18. Veselovskij I. N. (Tr.) and Bashmakova I. G., (Intr. & Comm.), 1974, Diophantus' *Arithmetica,* Moscow: Nauka. [In Russian].
19. Vlastos G., 1954, The Third Man Argument in the "Parmenides", *Philosophical Review* **63**, 319–49.
20. Weil A., 1983, *Number Theory: An Approach Through History. From Hammurapi to Legendre.* Boston: Birkhäuser.

EBERHARD KNOBLOCH

# THE INFINITE IN LEIBNIZ'S MATHEMATICS – THE HISTORIOGRAPHICAL METHOD OF COMPREHENSION IN CONTEXT

G. W. Leibniz published only a very small part of his papers and manuscripts. Up until now we know only 15% of his mathematical writings; their publication will probably fill 35 volumes of the Academy edition of his works and letters. The first volume, edited by my collaborator W. S. Contro and myself, appeared three years ago (Leibniz 1990), after many years of rational printing planning.

Among Leibniz's hitherto unpublished writings is a voluminous and extremely interesting treatise on the arithmetical quadrature of conic sections. The transcription of this consists of more than 200 pages, a critical edition of which I have prepared for the "Abhandlungen" of the Göttingen Academy (Leibniz 1993).

In it, Leibniz deals especially with the infinite. Below I shall rely mainly on this treatise. The subject is important and interesting, but at the same time very difficult. Leibniz wrote in his *New essays on human understanding* (Leibniz 1962, 377): "All this makes clear that the human spirit considers such strange questions, especially if the infinite is implied that we cannot be astonished if it turns out to be difficult to come to a conclusion."

He describes his own difficulties when he discusses the infinite. But the same applies also to historiography concerning his explanations. We come across affirmations that say quite the opposite of what he had in mind or that introduced notions which he did not use. In short: it is a matter of how we have to read and try to understand the scientific texts, the basis of our research. In order to avoid contradictions we have to try to understand the explanations of an author in their contexts. I would like to try to demonstrate that Leibniz dealt with the infinite in a consistent manner, although the contrary seems to be the case. Let us consider three examples:

1. The asymptote of the hyperbola.
   If $d$ is the asymptote of the curve we find in Leibniz



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 265-278.

| | | |
|---|---|---|
| | either $d \cap c = \emptyset$ | The curve and the asymptote concur nowhere. The asymptote is the "unbounded" infinite. |
| $d$ asymptote | | |
| | or $d \cap c \neq \emptyset$ | The asymptote is the "bounded" infinite |

2. The infinitely small is not equal to zero. It is not finite, while the extension of a point is zero. Nevertheless, Leibniz says that

   AA is an infinitely small straight line of a point.
   The positive powers of zero are infinitely small or equal to zero.
   The logarithm of zero or of the infinitely small.

3. Leibniz denies the existence of an infinite number which one could call the greatest number. Nevertheless, he calculates with the sum of the divergent harmonic series.

   I would like to try to throw light on these contradictions by discussing the following four points:

1. The mathematical notions or objects with regard to the infinite
2. The reasoning
3. The calculation with the infinite
4. The geometry of indivisibles according to Leibniz

## 1. THE MATHEMATICAL NOTIONS OR OBJECTS WITH REGARD TO THE INFINITE

Leibniz uses the notions of "bounded" (terminata) and "unbounded quantity" (quantitas interminata) in order to explain the notion of "infinitely great" and "infinitely small". In his correspondence with Pierre Varignon of 2 February, 1702, he says that, strictly speaking, the infinite must have its source in the unbounded; however he does not see any possibility in finding a suitable foundation to distinguish it from the finite without using the unbounded (Leibniz 1859, 91). Beyond that there is a close relationship between the notion of infinitely great and infinitely small, and the notions of divisible and indivisible. We must not mix these notions with each other. For that reason, it is necessary to begin the discussion by explaining the relations between the notions "bounded" and "unbounded".

An infinite quantity is a bounded or unbounded quantity which is greater than any assignable quantity or any quantity which can be represented by numbers:

infinite quantity

| | |
|---|---|
| bounded or immense quantity (fiction) | unbounded quantity with last point at least on one side in case of lines. Nothing can be added. |

1. An unbounded infinite quantity is a quantity where we cannot choose a last point at least on one side.
2. A bounded infinite quantity is a fictitious quantity on which we rely if we measure infinitely long but finite spaces.

This distinction explains the different descriptions of the asymptote.

Leibniz does not discuss the question of existence; the same applies to infinitely small quantities. It is the task of the metaphysician to examine whether the nature of things admits such quantities. The geometer is content demonstrating what follows from the presuppositions. This is apparently a rather modern and pragmatic standpoint.

Leibniz explicitly says that an unbounded line (the asymptote of a hyperbola for example) is as little the subject of geometrical consideration as the magnitude of a point. We add or subtract infinitely many points in vain. A bounded, arbitrarily often repeated line can no more compose or exhaust an unbounded line. By way of contrast a bounded line is made up of an arbitrary set of finite lines, even if the number of elements of that set should exceed any given number, that is:

A bounded infinite line is composed of finite lines.

A finite line is composed of infinitely small, but nevertheless divisible, lines.

While we can use the infinite and infinitely small in this sense in geometry, this is not true with regard to the minimum (the smallest line or point) and to the maximum (the greatest line or unbounded line).

The bounded infinite, but not the unbounded infinite is the object of mathematics. Leibniz never changed this attitude. He mentioned it in his treatise in 1675/76, and in his correspondence with Pierre Varignon in 1702. All the more surprising is the fact that a contrary remark is found in the secondary literature which appeared very recently (Breger 1990, 65): "Leibniz seems to have called in question the utility of the bounded infinite already during his stay in Paris. Later on he seems to have abandoned the bounded infinite in mathematics."

The contrary is correct. Let us consider, for example, the cited text for the first affirmation. It is a matter of a manuscript entitled "Extensio interminata" (unbounded extension). There Leibniz says that two unbounded straight lines, which are not parallel, have a common point. This does not apply to bounded straight lines. But they can be produced up to such a point of intersection. Leibniz even adds that straight lines, which are unbounded by themselves, are bounded by us or by bodies (Leibniz 1980, N. 66). There is not the slightest indication that Leibniz has called in question bounded infinite quantities.

By contrast, Leibniz underlines the uselessness of unbounded quantities, of unbounded lines. While he at first tries to characterize the bounded lines by saying that they are "in a certain way" the mean between the smallest line (linea minima) and the greatest line – the truly unbounded line – he completely rejects this first solution and explains:

The finite line $l_f$ is the mean proportional between a certain infinitely small line $l_{is}$ and a certain infinite line $l_i$, and that truly and exactly, but not "in a certain way."

$$l_i : l_f = l_f : l_{is} (*)$$

Hence he accepts different orders of infinite lines and different infinitely small lines. If we choose an arbitrary finite line $l_f$, we find an infinitely small line $l_{is}$ and an infinite line $l_i$, so that we obtain the equation (*) or

$$l_f{}^2 = l_i \bullet l_{is}.$$

Such an equation is not generally true but only in some special cases, for example, in the case of the conic hyperbola.

The mathematician applies the bounded infinite and the infinitely small which is divisible; these are the "infinitesimae parts" (infinitesimae partes) of a straight line. He certainly does not apply points which are truly indivisible. Hence, we obtain the scheme:

infinitely small quantity
smaller than any assignable quantity

| | |
|---|---|
| unequal to zero | zero (practically) |
| divisible | indivisible |
| (fiction) | (way of speaking) |
| (in proofs) | (in calculations) |
| infinitely small | infinitesimal point |

It is very interesting to see that Leibniz again takes up the question of whether an absolutely unbounded space can be reduced to a finite space. First, I would like to explain briefly the presuppositions. It is a question of the so-called "figure of angles", that is, of the circle and of the curve constructed according to the transmutation theorem (Figure 1).

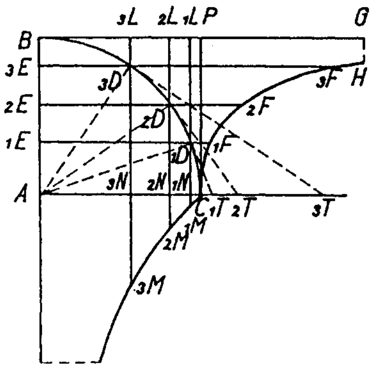



Fig. 1. *EF* is equal to *LM*.

We construct the figure in the following way:

1. We choose a point D on the circle ABC.
2. We draw the tangent in D.
3. The tangent cuts the axis in D.
4. Let the point of intersection between the parallel through D and the perpendicular in T be F.
5. All points F form a new curve.

Let us now consider an arbitrary space CAEFC (Figure 2). In ccordance with his theorem 7, Leibniz knows that in such a figure the following equation holds:

> The space CA$n$E$n$FC (limited by four lines) = 2. sector $n$DAC$n$D

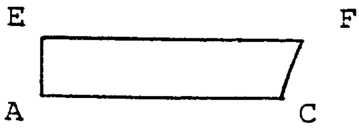



Fig. 2.

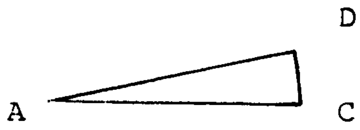



Fig. 3.

(see Figure 3).

Hence we obtain

$$\frac{2.\ {}_1DAC_1D}{2.\ {}_2DAC_2D} = \frac{{}_1DAC_1D}{{}_2DAC_2D} = \frac{{}_1\widehat{DC}}{{}_2\widehat{DC}} = \frac{\sphericalangle CA_1D}{\sphericalangle CA_2D}$$

Hence, we obtain the following corollary:
The space of the figure of angles
(the length of which is infinite)

$$\text{CABG} \ldots \text{HFC} : \text{a finite part CAEFC}$$
$$= \sphericalangle\ \text{BAC (right angle)} : \sphericalangle\ \text{DAC (acute angle).}$$

Leibniz remarks that the right angle seems to correspond with the absolutely unbounded space. But he dares not pretend, because of this fact, that this space is reduced to a finite space. He continues, "There can be no doubt that the right angle does not correspond with any space of the figure of angles save with the absolutely unbounded space."

## 2. THE REASONING

Leibniz underlines that the method of infinite quantities is dangerous. We have to apply the thread of an indirect proof if we have in mind to rely on it. Why? Because the infinite is characterized in a privative way. In reality, there is no last point, no end, we have only relied on a fiction in order to admit it in mathematics.

In his treatise Leibniz avows that he does not know any method by means of which he can demonstrate a single quadrature without indirect reasoning. He has reasons why he fears that one cannot attain this aim, because of the nature of things without fictitious quantities, that is, without the infinite and the infinitely small quantities.

He repeats this opinion again in 1698 in his correspondence with John Bernoulli (Leibniz 1856, 524). Let us consider the following example (Leibniz 1993, theorem 20):

Let there be given three quantities X, Z, V. Let V + X have a finite ratio to V + Z which is unequal to one:

$$\frac{V + X}{V + Z} \neq 1.$$

1. If X, Z are finite, V will also be finite.
2. If X or Z is infinite, V will also be infinite.

*Proof:*
1. Let X, Z be finite. This implies that V is also finite.
   Proof: Let X, Z be finite, but suppose that V is infinite.
   → V + X is infinite, V + Z is infinite (that is, Leibniz calculates: finite + infinite = infinite)
   → (V + X) − (V + Z) = infinite.
   For if we subtract a "greater" infinite from a "smaller" infinite, where their ratio is ≠ 1, the rest is infinite.
   Thus we obtain a contradiction, because (V + X) − (V + Z) = X − Z is finite according to the presupposition.
2. Let X or Z be infinite. This implies that V is also infinite.
   Proof: Let Z be infinite, X finite (without restriction of generality), but suppose that V is finite. As a consequence, V + X is finite, V + Z is infinite. Thus we obtain a contradiction, because V + X has a finite ratio to V + Z, which is unequal to number one according to the presupposition. But we have to state that the ratio finite: infinite is not finite, it is infinitely small. The contradiction consists in the fact that an infinitely small quantity is not a finite quantity, while the ratio (V + X)/(V + Z) is infinitely small.

### 3. THE CALCULATION WITH THE INFINITE (THE ARITHMETIC OF THE INFINITE)

Leibniz admits, that we can calculate with the infinite in mathematics. But then, how can we calculate with it?

I would like to explain the practical arithmetic by means of two examples and an enumeration of rules which I have taken from the treatise. Leibniz implicitly uses these rules without having demonstrated them, except the sixth and the eighth laws. He believed such proofs to be superfluous, because of the definition of these quantities. He would have been able to give such indirect proof.

*First Example:*

Leibniz proves, by means of geometry, that the sum of the harmonic series

$$1 + 1/2 + 1/3 + \ldots = \mathrm{A}$$

is an infinite quantity (Leibniz 1993, Theorem 45), which he designates by 1/0 or $A$. Thus he proves that the series is divergent. Nevertheless, we are allowed to calculate with $A$ in the same measure as the nature of the series is known as using "useful fiction", while convergent series are "true quantities". The letter $A$ does not at all imply the existence of the designated quantity:

series

| divergent | convergent |
|---|---|
| useful fiction | true quantity |

Let us calculate the sum of the reciprocal triangular numbers which is 2. Leibniz calculates in the following manner:

Let $1 + 1/3 + 1/6 + 1/10$ etc. be equal to 2 B.
$\rightarrow 1/2 + 1/6 + 1/12 + 1/20$ etc. $= \mathrm{B}$
$\rightarrow \mathrm{A} - \mathrm{B} = 1/2 + 1/3 + 1/4$ etc. $= \mathrm{A} - 1$
$\rightarrow \mathrm{B} = 1 \rightarrow 2\,\mathrm{B} = 2$

Before we consider the second example let us have a look at the list of the above-mentioned twelve rules.

THE TWELVE RULES

1. finite + infinite = infinite

2.1 finite ± infinitely small = finite

2.2 $x,y$ finite, $x = y$ + infinitely small $\rightarrow x - y \approx 0$
(not assignable difference)

3. $\text{infinite}_1 - \text{infinite}_2 = \text{infinite}_3$, if $\text{infinite}_1 > \text{infinite}_2$
(or $\text{infinite}_1 : \text{infinite}_2 \neq 1$)

4. infinite ± infinitely small = infinite

5. finite $x$ infinitely small = infinitely small

6. infinite $x$ infinitely small $\left\langle \begin{array}{l} \text{infinite or} \\ \text{finite or} \\ \text{infinitely small} \end{array} \right.$
(proof is needed)

7.1 infinite $x$ infinite = infinite

7.2 $x^n$ infinite $\rightarrow x$ infinite

8. infinite : infinite $\left\langle \begin{array}{l} \text{finite} \\ \text{infinite} \end{array} \right.$
(proof is needed)

9. $x$ infinitely small, $y > 0$, $y < x \rightarrow y$ infinitely small

10. finite : infinitely small = infinite : finite = infinite
(greater than any assignable ratio)
*Corollary:* finite : infinitely small = $x$ : finite $\rightarrow x$ infinite

11. infinitely small : finite = finite : infinite = infinitely small (infinitesimal)
*Corollary:* finite : infinite = $x$ : finite $\rightarrow x$ infinitely small

12. $x : y = (x + \text{infinitely small}_1) : (y + \text{infinitely small}_2)$

The laws 10 and 11 are particularly important, because they permit the demonstration that a quantity is infinite or infinitely small. Leibniz discusses the solution of the sixth law by means of these two theorems. We are looking for the value of a product, which is a rectangle. Thus the area of the rectangle is the quantity $x$, the value of which is sought. Leibniz deduced a proportionality: one side of it is a known ratio, while the ratio of the other side contains the quantity $x$.

The second example uses law 6:
*Theorem* 21 (Leibniz 1993). Let the curve $oC_1C_2C$ be a hyperboloid (a hyperbola of an arbitrary degree)

$$x^n \cdot y^m = a$$

The rectangle under the infinitely small abscissa AoB and the infinitely great ordinate oBoC is
1. an infinite quantity, if $n > m$,
2. an infinitely small quantity, if $n < m$,
3. a finite quantity, if $n = m$.

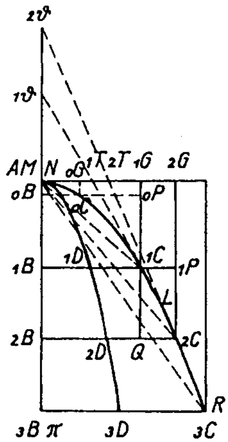



Fig. 4.

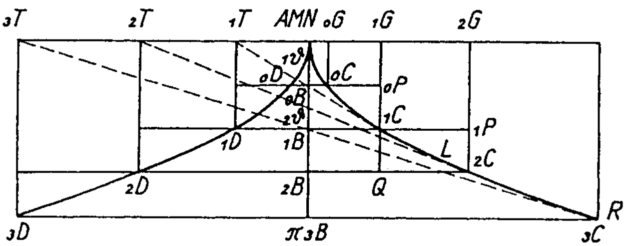



Fig. 5.

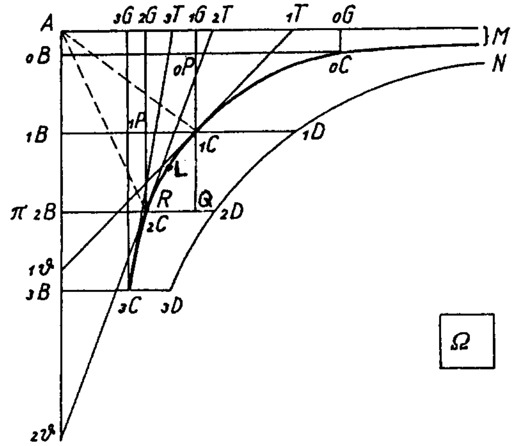



Fig. 6.

*Proof:*

1. If $n = m$ then the curve is the conic hyperbola. Hence, the rectangle oCoGAoB is equal to the square A2BR2G in agreement with the nature of the hyperbola. The square is finite. Hence, the rectangle oCoGAoB formed by the infinitely small segment AoB and the infinitely great straight line oBoC is finite.
2. $n > m$, for example $n = 2$, $m = 1$ or $x^2y = a$.

$$\rightarrow \frac{\mathrm{oBoC}}{\mathrm{1B1C}} = \frac{\mathrm{A1B}^2}{\mathrm{AoB}^2}$$

$$\frac{\text{rectangle AoBoCoG}}{\text{rectangle A1B1C1G}} = \frac{\mathrm{AoB}}{\mathrm{A1B}} \cdot \frac{\mathrm{oBoC}}{\mathrm{1B1C}} = \frac{\mathrm{AoB}}{\mathrm{A1B}} \cdot \frac{\mathrm{A1B}^2}{\mathrm{AoB}^2} = \frac{\mathrm{A1B}}{\mathrm{AoB}}$$

$$\frac{\text{rectangle AoBoCoG (unknown)}}{\text{rectangle A1B1C1G (finite)}} = \frac{\text{A1B (finite)}}{\text{AoB (infinitely small).}}$$

$$\text{If } \frac{x}{\text{finite}} = \frac{\text{finite}}{\text{infinitely small}}\text{, then } x \text{ is infinite (by law 10)}$$

→ the rectangle AoBoCoG is infinite. The proof has to be generalized, if generally $n > m$.

Leibniz adds another elucidating illustration of this result, because powers of zero are concerned. Let us consider the rectangle $xy$ under the first abscissa and ordinate under the relationship

$x^n y^m = 1.$

This rectangle $xy$ can be expressed by $x^{(m-n):m}$, for if $y^m = x^{-n}$ or $y = x^{-n:m}$, then $yx = x^{(m-n):m}$.

If $x = 0$ and $m > n$, the quantity $m-n : m$ will be positive,
$m < n$, the quantity $m-n : m$ will be negative.

A positive power of zero is infinitely small, a negative power of zero is infinite. Hence, we obtain in another way the same result as before.

3. The third case, $n < m$ corresponds to a replacement of the variable $x$ by $y$ and vice versa.

## 4. THE GEOMETRY OF INDIVISIBLES ACCORDING TO LEIBNIZ

As we know, indivisibles are, strictly speaking, points – zeros which are useless in geometry according to Leibniz. Hence, he has to interpret them in a way which does not follow the original idea of Cavalieri. Leibniz understands by "the sum of all straight lines" the sum of all rectangles under the same united straight line after having supposed a constant, always equal, and indefinitely small interval (indefinitae parvitatis). According to Leibniz, this is the correct form of Cavalieri's geometry of indivisibles, which can be otherwise misleading. We state that he interprets Cavalieri's theory in conformity with his own philosophy of mathematics. His famous example is the paradox of the hyperbola, discussed in his treatise on the arithmetical quadrature of conic sections in 1676 (scholium of theorem 22), presented by John Bernoulli some 20 years later in 1698 (Leibniz 1856, 522–524).

The example presupposes the following theorem: In a simple analytical relationship

$$y^n \bullet x^m = a \quad \text{or} \quad bx^m = y^n$$

the ratio of the zone between two ordinates, the arc of the curve and the axis of the so-called "conjugated zone" between the two corresponding abscissas, the same arc of the curve and the conjugated axis is the same as the ratio of the exponent of the power of the ordinate (the variable $y$) to the exponent of the power of the abscissa (the variable $x$), that is, as $n : m$. In the case of the conic hyperbola, the hatched zones are equal, because n $= m = 1$. This applies to two arbitrary conjugated zones. Hence, all horizontal zones up to $A$ fill the area 2C2BAM2C, all vertical corresponding zones fill only the area 2C2GM2C.

Hence, it follows that a part equals the whole. This is absurd, because Leibniz always presupposes the validity of the axiom:

The whole is greater than its parts.

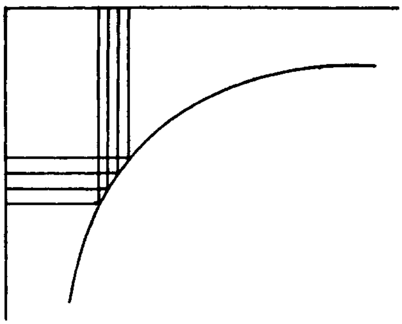


Fig. 7.

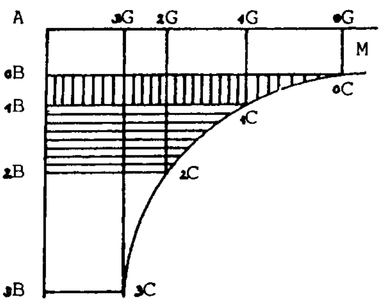



Fig. 8.

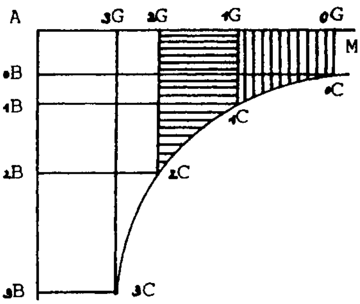



Fig. 9.

This is his fundamental axiom of a science of quantities, that is, of mathematics. Where is the error? The error consists in identifying the indivisible or zero with the infinitely small, though there is a decisive difference between these two notions. The last abscissa (ultima abscissa) AoB does not equal zero. Strictly speaking, there is no such last abscissa. This is only a manner of speaking.

*Technical University of Berlin*

## REFERENCES

Breger, H., 1990, Das Kontinuum bei Leibniz. In: *L'infinito in Leibniz, Problemi e terminologia*, Simposio Internazionale Roma, 6–8 novembre 1986, a cura di Antonio Lamarra, pp. 53–67. Roma: Edizioni dell' Ateneo.

Knobloch, E., 1990, L'infini dans les mathématiques de Leibniz. In: *L'infinito in Leibniz, Problemi e terminologia*, Simposio Internazionale Roma, 6–8 novembre 1986, a cura di Antonio Lamarra, pp. 33–51. Roma: Edizioni dell' Ateneo.

Leibniz, G. W., 1856, *Mathematische Schriften*, hrsg. von C. I. Gerhardt, volume III. Halle: H. W. Schmidt. Reprint Hildesheim: Olms, 1962.

Leibniz, G. W., 1859, *Mathematische Schriften*, hrsg. von C. I. Gerhardt, volume IV. Halle: H. W. Schmidt. Reprint Hildesheim: Olms, 1962.

Leibniz, G. W., 1962, Nouveaux Essais sur l'entendement humain. In: *G. W. Leibniz, Sämtliche Schriften und Briefe*, hrsg. von der Preußischen (später Deutschen) Akademie der Wissenschaften zu Berlin. Reihe VI, Bd. 6, bearbeitet von H. Schepers und A. Robinet. Berlin: Akademie-Verlag.

Leibniz, G. W., 1980, *Sämtliche Schriften und Briefe*, hrsg. von der Preußischen (später Deutschen) Akademie der Wissenschaften zu Berlin. Reihe VI, Bd. 3, bearbeitet von H. Schepers, W. Schneiders, W. Kabitz. Berlin: Akademie-Verlag.

Leibniz, G. W., 1990, *Sämtliche Schriften und Briefe*, hrsg. von der Preußischen (später Deutschen) Akademie der Wissenschaften zu Berlin. Reihe VII, Bd. 1, bearbeitet von E. Knobloch und W. S. Contro. Berlin: Akademie-Verlag.

Leibniz, G. W., 1993, *De quadratura arithmetica circuli ellipseos et hyperbolae cujus corollarium est trigonometria sine tabulis*. Kritisch herausgegeben und kommentiert von Eberhard Knobloch. Göttingen: Vandenhoeck & Ruprecht.

CHRISTINE PHILI

# JOHN LANDEN: FIRST ATTEMPT FOR THE ALGEBRIZATION OF INFINITESIMAL CALCULUS

The conceptualization and codification of infinitesimal calculus during the 17th century were not able to elucidate their fundamental notions. The criticisms and controversies surrounding the foundation of calculus finally led to important scientific activity which gave rise to the rigorous foundation of analysis. Nevertheless, before the final period of its arithmetization in the 19th century, there was a turning point in the 18th century when algebrization transformed the metaphysical face of analysis.

The precarious situation of new calculus – on account of the prestige of algebra – acquires the ability to integrate exactness into the mechanism of reasoning. The "vanishing" quantities subject to the algebraic homogeneity, confirm the intrinsic relation of the constitutive operations of analysis to the properly algebraic operations in the Lagrangian conception: "To connect this calculus to the rest of the algebra so as to make a single method of all the means which we have used to explain the differential calculus and to demonstrate its rules"[1] indicates Lagrange, principal representative of the algebrization of analysis.[2]

According to the Lagrangian doctrine, infinitesimal calculus, absorbed into the generalized form of the Taylor series development participates in the universality of the algebra "the calculus of functions – lagrangian expression for his theory – moreover, serves to link directly the differential calculus with algebra, which we could say was until now a separate science."[3]

Meanwhile, in England, the expulsion of the infinitely small quantities, of fluxions, of vanishing quantities for the benefit of algebraic operations, preceded the publication of the Lagrangian work of 1797.[4] The movement of the algebrization of analysis and, similarly, the tendency to elevate algebra to a "universal" science, reflects the generality of algebra and the certainty which it generates in the system of relations as well as in its logical extension. In his lectures in Cambridge,[5] Newton emphasized the role of algebra, stressing that in algebra every conclusion is a real theorem.[6]

In this atmosphere, where the difficulties and theoretical uncertainties did not approach the laws and necessary operations of algebra, there



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 279-293.

appears the first attempt for the algebrization of analysis, that of John Landen (1719–1770).[7] He was a true amateur of mathematics who studies, above all, the works of his compatriots and, in particular, the work of Simpson, Stirling, and Cotes. Thanks to his ardour for science he was able to enter the Royal Society of London in 1766. He owes his mathematical reputation mainly to his works about the rectification of an hyperbolic arc and the elliptic integrals.[8,9]

Landen, who was not in touch with the Continent, who was not aware of the European scientific reality and who was not a fanatic reader of the scientific publications of foreign academies, was the first to reduce the differential calculus to algebraic processes and was, in a certain way, the precursor of the Lagrangian doctrine.[10]

His conviction about the preponderant role of algebra appears in his first critique concerning the Newtonian method: "Finding from velocities, the spaces passed over and *vice versa* may be managed by common *Algebra*, without the least obscurity. The Business had always been better considered in the Light, without ever making use of the term *Fluxions*, as if a new kind of *Analysis*, tho' in *Fact*, only the *Doctrine of Motion* improved and applied to Purposes before unthought of."[11]

This assimilation of calculus with algebraic truth was crystallized around the new analysis. Residual Analysis was exposed in 1758 in his book *A Discourse concerning the Residual Analysis: A new branch of the algebraic art, of very extensive use both in Pure Mathematics and Natural Philosophy*.[12]

The regard for rigor – which, in the first years of the formation of the infinitesimal analysis remained distant – led Landen (i) to formulate his critiques concerning the introduction of the notion of movement into calculus and (ii) to go on to the conquest of the rigorous foundation of infinitesimal calculus with the subtle and strong arms of algebra.

We may indeed very naturally conceive a line to be generated by motion; but there are quantities... which we cannot conceive to be so generated. It is only in a figurative sense, that an algebraic quantity can be said to increase or decrease with some velocity. Fluxions therefore are not immediately applicable to algebraic quantities... It therefore, to me, seems more proper, in the investigation of propositions by algebra, to proceed upon the *anciently received* principles of that art... That the borrowing principles from the doctrine of motion, with a view to improve the analytic art, was done, not only without any necessity, but even without any peculiar advantage, will appear by showing, that whatever can be done by the method of computation, which is founded on those borrowed principles, may be done as well by another method founded entirely on the *ancient received* principles of algebra... It is by means of the following theorem

$$\frac{x^{m/n}-u^{m/n}}{x-u} = x^{(m/n)-1}\;\frac{1+\frac{u}{x}+\left(\frac{u}{x}\right)^{2}+\left(\frac{u}{x}\right)^{3} \quad (m)}{1+\frac{u}{x}+\left(\frac{u}{x}\right)^{2m/n}+\left(\frac{u}{x}\right)^{3m/n} \quad (n)}$$

where $m$ and $n$ are integers that we are enable to perform all the principal operations in our said analysis.[13]

The inventor of the new calculus tries to defend his own method and to prove its superiority in relation to the Newtonian method:

Fluxionists, in determining the limit of the ratio of the increments of $x$ and $x^{m/n}$, commonly have recourse to the binomial theorem (which is much more difficult to investigate than the limit they are seeking): But how easily may that limit be found, without the help of that theorem, by the equation exhibited on page 5! Thus, the increment of $x$ being denoted by $x'$, the increment of $x^{m/n}$ is $\overline{x+x}'^{\,m/n}-x^{m/n}$, and the ratio of those increments is:

$$\frac{\overline{x+x'}^{\,m/n}-x^{m/n}}{x'} = \frac{\overline{x+x'}^{\,m/n}-x^{m/n}}{\overline{x+x'}-x} =$$

$$= \frac{\overline{x+x'}^{\,(m/n)-1}\;1+\frac{x}{xx'}+\left(\frac{x}{x+x'}\right)^{2}+\left(\frac{x}{x+x'}\right)^{3} \quad (m)}{1+\left(\frac{x}{x+x'}\right)^{m/n}+\left(\frac{x}{x+x'}\right)^{2m/n}+\left(\frac{x}{x+x'}\right)^{3m/n} \quad (n)}$$

which, when $x'$ vanishes, is manifestly equal to $\frac{m}{n}\,x^{(m/n)-1}$, the limit of the said ratio.[14]

Without making use of the limit of D'Alembert,[15] Landen proceeds with pure algebraic process to find the derivative or the residual quotient of $x^{m/n}$.

The birth of this new analysis in the land of fluxions was not long in provoking objections. Landen's book was an anonymously attacked[16] in the *Monthly Review*.[17] The writer asserted that the Residual Analysis "is no other than Sir Isaac Newton's method of differences; and it is well known, that if the differences are diminished so as to vanish, their vanishing ratio becomes that of fluxions... that his pretended Residual Analysis renders the investigations more tedious and obscure than any other." The originality of this "able English geometer"[18] who was proposed an homogeneous algebraic process was not be appreciated by his harsh critic.[19] So the introduction of algebraic tools in calculus[20] had not appeared as a consequence of necessary evolution. In the Eighteenth century, England remained attached to the Newtonian tradition.

The scientific knowledge of calculus based on the plurality of systems (fluxions, infinitely small quantities, vanishing quantities, limits, etc.) remained for Landen the inevitable cause of confusion and complications. Convinced of the efficacy of the algebraic system, many years later he resumed[21] his attempt to present the algebraic foundation of calculus. The book title itself reveals his purposes: *The Residual Analysis, a new branch of the Algebraic Art, of very extensive use, both in Pure Mathematics and Natural Philosophy.*[22] The following passage is characteristic in this respect:

> The principles of common Algebra and Geometry having been though insufficient to enable the Analyst to pursue his speculations in certain branches of sciences; new principles very different from those before made use of, have through a supposed necessity, been introduced into Analytics. The fluxionists following Sir Isaac Newton introduce imaginary notion...
>
> Mr. Leibnitz and his followers, to avoid the supposition of motion, consider quantities as composed of infinitesimal elements and reject certain parts of the infinitely small increments of quantities as infinitely less than other parts. In the Residual Analysis, (admitting no principles but such as were anciently received in Algebra and Geometry) we neither have recourse to infinitesimals, nor to the principles of motion.[23–26]

For Landen – as well as for Lagrange – the binomial theorem became an attractive theorem for the analogy which hides inside the differentials of all orders and the powers of the binomial of the same order. In the foreword of his book, Landen informs us that he has worked on a binomial theorem with a new and easy method[27] and this way has led him to "a new method of calculation which might be acceptable to the mathematical world."[28,29]

For this new branch of the science, Landen proposed the term "Residual Analysis", "because, in all the enquiries wherein it is made use of, the chief means whereby we obtain the desired conclusions are such quantities and algebraic expressions, as by Mathematicians are denominated residuals."[30] The principal tool for Landen's calculus is not the quotient of fluxions or of the one of differentials but "the value of the quotient of one residual divided by another."[31]

After the definition of a determinate or invariable quantity and of an indeterminate or variable quantity,[32] Landen gives the following definition for a function: "an algebraic expression, composed in any manner, of any power of any variable quantity, with any invariable coefficient, is called a function of that quantity, for instance,

$$a + bx^m + cx^n \text{ and } \frac{(ex^r + \sqrt{f^2 + x^2})}{gx^s}$$

are functions of the variable $x$".[33,34] Echoes of J. Bernoulli[35] and of L. Euler [36] definitions are obvious.

A residual is an expression which is formed by the subtraction of $y$ function from $x$ and $y_1$ his similar function of $x$[37] "the value of the quotient of $y-y_1$ divided by $x-x_1$ in the particular case where $x_1$ is equal to $x$ is called the special value of that quotient; $x$ and $y$ are respectively named the prime member of the divisor and the prime member of the dividend."[38,39] The manipulation of these indeterminate[40] forms leads Landen to "state precisely" the derivatives and to present their applications in geometry.[41]

Residual Analysis "which is founded (as I conceive) on the genuine principles of Analytics"[42] is not a firework in the scientific sky of England; Landen has worked seriously for the elaboration of his new calculus and his rigorous foundation of the rules that prevail there. The second chapter of his book is dedicated to "The Invention of Rules necessary to facilitate Computations in this Analysis."

So Landen proceeds as follows to find the residual quotient of $x^{m/n}$; "it is by means of the following theorem... that we are enabled to perform all the principal operations to our said Analysis".[43]

$$\frac{u^{m/r}-\omega^{m/r}}{u-\omega} = \frac{u^{m-1}+u^{m-2}\omega+u^{m-3}\omega^2+u^{m-4}\omega^3 \qquad (m)}{u^{m-m/2}+u^{m-2m/r}\omega^{m/r}+u^{m-3m/r}\omega^{2m/r}+u^{m-4m/r}\omega^{3m/r} \qquad (r)}$$

$$= u^{(m/r)-1}\,\frac{1+\frac{\omega}{u}+\left(\frac{\omega}{u}\right)^2+\left(\frac{\omega}{u}\right)^3 \qquad (m)}{1+\left(\frac{\omega}{u}\right)^{m/r}+\left(\frac{\omega}{u}\right)^{2m/r}+\left(\frac{\omega}{u}\right)^{3m/r} \qquad (r)}\,, \quad {}^{44-47}$$

where $m$ and $n$ are positive integers. His demonstration[48] is based on the classical rules of arithmetic; taking $\omega=u$, he finds the value of the quotient $u^{m/r}$, he finds the value of the quotient $u^{m/r}$.

In reality, Landen manipulated the expressions which are infinite series and at the end, equaled the two components $\omega$ and $u$ resulting in the special value.[49]

Landen proceeded in the same way to find $v^{-(m/r)}$. Ordinary examples complete his first theorem, thus the applications for $\frac{m}{r} = \frac{4}{3}$

$$u^{4/3}-\omega^{4/3}:\overline{u-\omega} = u^{1/3}\,\frac{1+\frac{\omega}{u}+\overline{\frac{\omega}{u}}\Big|^{2}+\overline{\frac{\omega}{u}}\Big|^{3}}{1+\overline{\frac{\omega}{u}}\Big|^{4/3}+\overline{\frac{\omega}{u}}\Big|^{8/3}}$$

when $\omega = u$ the residual quotient of $u^{4/3}$ become $u^{4/3}-\omega^{4/3}$ divided by $u-\omega$ equals $\frac{4}{3}u^{1/3}$ .[50]

As corollary, he exposed the general case where $p=\frac{m}{r}$ .

$x = u$, $x_1 = \omega$ so the special value is:

$$\overline{x^p - x_1^p} - \overline{x-x_1} = px^{p-1}$$

This formula of Landen is nothing less than the derivative of the function $x^p$, nevertheless, we must quote that the English mathematician obtains the derivative by purely algebraic processes[51] – in the eighteenth century the expansion of function in series is an algebraic operation. The questions concerning the development's existence did not belong to the framework of this period. Many years later, this method received the approval of Lagrange who, in fact, in 1797 had developed his original ideas of 1772 and also in Landen's way. Lagrange wrote:

(Landen's method is) analogous to the differential method, but instead of employing infinitely small differences or zeros of variable quantities, he uses, first, different values of these variable quantities which he sets equal to zero after having removed, by division, the quantity which would make the equality invalid. By this means one avoids infinitely small and vanishing quantities. But this procedure and the applications are encumbering and hardly natural, and one ought to agree that this way of making the differential calculus more rigorous in its principles loses its main advantage, the simplicity of the method and the facility of its operations.[52]

Many years later Clairaut in his *Elements of Algebra* (2 Vols Paris, 1797) points out this important contribution of Landen, i.e., his demonstration concerning the binomial theorem:

The second volume contains... a demonstration of Newton's binomial formula based on a theorem of the English geometer Landen and which deserves to be known for its elegance and its usefulness which can contribute in series' expansion[53]... I must refer here to this demonstration which we owe to the English geometer Landen, because it is founded on a very important analytic remark which can be extended to many other subjects.[54]

Another favourable opinion, that of Lacroix and Thevenau, appears in the additions they wrote for the 5th ed. (1800) and for the 6th one of Clairaut's book. They consider Landen's proof to be better than Clairaut's."... The way adopted by Clairaut is not among the more satisfactory, because in fact it offers only a verification of some first terms. We can conclude those which follow only with the help of analogy's laws, laws which aren't so evident".[55]

Certainly Landen's Residual Analysis – in spite of his attempt at algebrization – remains a method in which evanescent fractions appear. Consequently, it raises controversies: "Landen, one of those men so frequent in England whose talents surmount their narrow education, produced in 1758, a new form of the Fluxionary Calculus, under the title of Residual Analysis, which, though framed with little elegance, may be deemed, on the whole, an improvement on the method of ultimate ratios."[56]

Following the prototypes of the contemporary books, Landen dedicated the third chapter of *Residual Analysis* to exponentials[57] and logarithms. For the exponential function $n^x$ and, also, for the logarithmic one, Landen uses the form of power series, resonance of eulerian conception.

$$\text{So } n^x = 1+Ax+Bx^2+C^3+\ldots \tag{i}$$

and he undertakes finding, in terms of $n$, the coefficients $A$, $B$, $C$..

$$n_1^x = 1+Ax+Bx^2+Cx^3+\ldots \tag{ii}$$

He subtracts from formula (i), formula (ii) and he obtains

$$n^x-n_1^x = \overline{Ax-x_1}+ B\overline{x^2-x_1^2} + C\overline{x^3-x_1^3}+\ldots \tag{iii}$$

after dividing formula (iii) by $x-x_1$

$$\frac{n^x-n_1^x}{x-x_1} = A+B\overline{x+x_1}+ C\,\overline{x^2+xx_1+x_1^2}+\ldots$$

The quotient of $n^x-n_1^x$ divided by $x-x_1$ equals $g\times n^x$ when $x=x_1$, $g\times n^x$ or $g+gAx+gBx^2+gCx^3+\ldots$ is equal to

$$A+2Bx+3Cx+4Dx^2+\ldots$$

From the comparison of the homologous term, he obtains:

$$A=g,\quad B=\frac{gA}{2},\quad C=\frac{gB}{3},\quad D=\frac{gC}{4}\,.$$ [58]

Therefore

$$n^x = 1 + gx + \frac{g^2x^2}{2} + \frac{g^3x^3}{2.3} + \frac{g^4x^4}{2.3.4} + \ldots$$

In this part of the third chapter, Landen gives the definition of logarithms as a solution of the system of functional equations $(xy) = f(x) + f(y)$ and $f(x^a) = af(x)$:

Logarithmus being an artificial Set of numbers corresponding to another Set of numbers in such a manner, that the sum of the logarithmus of any two numbers is equal to the logarithmus of the product of those two numbers; and the sum of the exponents of any two powers of any given quantity being equal to the exponent of that power (of the same given quantity) which is produced by multiplying those two powers together.[59]

The absence of the concept of continuity makes it difficult for Landen to prove the interdependence of these two functional equations.

In the same chapter Landen is concerned with the problem of convergence.[60] However, his verbalism – consistent with the written language of his era – could not succeed in producing the elaboration of the concept of convergence. Cauchy, with the precise and clear notion of limit, provided the solid groundwork on which the analysis would be renewed and reorganized in the nineteenth century.

A different theorem together with the aforementioned one theorem form the protagonists of Residual Analysis. This appears in the following pages of his book:

Suppose $E$ to be an algebraic expression composed of $x$ and other quantities; and suppose, that how near soever $x$ be taken to some certain quantity $g$, $E$ is positive when $x$ is less than $g$, and negative when $x$ is greater than $g$; or positive when $x$ is greater than $g$, and negative when $x$ is less than $g$; then shall $E$, or its reciprocal, be equal to zero when $x$ is equal to $g$.[61]

In modern mathematical language this theorem can be written as follows. Every explicit algebraic function studied on an interval containing $g$ which for $x<g$ is $>0$ and for $x>g$ is $<0$, can have one of the following forms:

(i) an explicit algebraic expression from which the denominator isn't equal to zero for $x=g$.

(ii) an explicit algebraic expression from which the denominator is equal to zero when $x=g$.

a) in the first case the function is continuous so after Bolzano's theorem is equal to zero on $g$.

b) if the denominator does not cancel $g$, the preceding remark is applied to $\frac{1}{E}$

Landeu's demonstration is completely verbal and not very rigorous. Nevertheless we must stress the originality of the mathematical thought of Landen that with the help of algebraic expression, studies the property of a continuous function and becomes Bolzano's precursor. Certainly the lack of a rigorous notion of limit delayed the restructuring of analysis in this era.

Chapters concerning tangents and other geometric applications complete Landen's book.[62] To find the derivative at the point $P(x_1,y_1)$ he proceeds as follows: Let the tangent line l intersect the $x$-axis at $N$. Let the ordinate line, drawn at $x_1+\Delta x_1$, intersect the tangent line l at $p$ and the parabola at $q$. Then $x_1-N$ is the length of the subtangent, and will be denoted by $s$.

Therefore

$$pq = a(x_1+\Delta x_1)^{m/r} - ax_1^{m/r} - \frac{ax_1^{m/r}}{s}\,\Delta x_1$$

by the hypothesis concerning concavity, this expression is greater than zero for all values of $\Delta x_1 \neq 0$

$$\frac{pq}{-\Delta x_1} \gtrless 0$$

according to $\Delta x_1 > 0$ or $< 0$.

From the aforementioned theorem[63]

$$\frac{pq}{-\Delta x_1} = \frac{ax_1^{m/r}}{s} - a\left[\frac{x_1^{m/r}-(x_1+\Delta x_1)^{m/r}}{x_1-(x_1+\Delta x_1)}\right] =$$

$$= \frac{ax_1^{m/r}}{s} - ax_1^{(m/r)-1}\left[\frac{1+\left\{\frac{x_1+\Delta x_1}{x_1}\right\}+\ldots+\left\{\frac{x_1+Dx_1}{x_1}\right\}^{m-1}}{1+\left\{\frac{x_1+\Delta x_1}{x_1}\right\}^{m/r}+\ldots+\left\{\frac{x_1+\Delta x_1}{x_1}\right\}^{(r-1)m/r}}\right]$$

From the other principal theorem[64] Landen asserts that this expression equals zero when $\Delta x_1 = 0$.

Therefore

$$\frac{ax_1^{m/r}}{s} - ax_1^{(m/r)-1}\left[\frac{1+1+1+\ldots(m)}{1+1+1+\ldots(r)}\right] = \frac{ax_1^{m/r}}{s} - \frac{m}{r}\,ax_1^{m/r-1} = 0$$

since

$$\frac{ax_1^{m/r}}{s} = \frac{y_1}{s} = \frac{\mathrm{d}y}{\mathrm{d}x}$$

it results that

$$\frac{dy}{dx} = \frac{m}{r}\ ax_{1}^{(m/r)-1} \quad [65]$$

From this example we can easily establish that Landen's process does not differ essentially from the method used by other scientists who consider calculus to be some kind of expendient for evaluating the quotient of two zeros. This calculus on zeros, which had provoked the reasoned criticisms of Berkeley,[66] reflected the scientific framework[67] of the era: "The infinitesimal calculus is a calculus on the zeros, but on the zeros which kept the trace of their origin, if we can say so."[68]

Landen, however, convinced that this first theorem[69] formed the "secret bridge" that joined the analysis with algebra "demonstrated" the autonomy of calculus and, at the same time, shattered the framework of the metaphysics. His theorem – which plays a role quite analogous to Taylor's theorem in Lagrangian theory – constituted for Landen a kind of revelation. "It is a matter of surprise to me that the algebraists have not before observed it, and shown its singular use in analysis."[70]

Landen's book, with its homogenous character, reflects his own personal vision of Infinitesimal calculus. Landen, who did mathematics for his own amusement[71] did not closely follow the currents of scientific thought. Indeed, as we have already established from studying his Residual Analysis, he naturally knew the works of Newton, Leibniz, and those of Bernoulli, d'Alembert etc., but he followed his own way, concentrating on the simple notion of algebra so as to free infinitesimal analysis from obscure and uncertain notions.

Landen's method was criticized[72] but the absence of necessary tools has not permitted a thorough examination of his doctrine. Meanwhile the criticisms, the absence of approval from his colleagues and perhaps a kind of strict self-criticism probably made Landen change his mind, completely abandon his residual analysis and turn to the Newtonian method.[73,74]

Our reservations concerning his difficult and obscure notation and complicated and long[75] calculus do not diminish his contribution to the algebrization of analysis. His determination to found the analysis on algebra, to disengage calculus from heterogeneous notions, like those of infinitely small fluxions,[76] and to give his work a sense of unity[77] Landen requires that in his doctrine everything is deduced from his principal theorem $(u^{m/r}-\omega^{m/r})/(u-\omega)$ surely constitute an important advantage for his book. Nevertheless, it made no impression[78] until the publication, in 1797, of the Lagrangian theory which would eradicate every difficulty "in relating directly this calculus (infinitesimal) to the algebra."[79]

*National Technical University, Athens, Greece*

## NOTES

[1] J. L. Lagrange, "Manuscrit" of his lectures (1797–1798). Notes by the citizen Le Gentil. M.S.S. 1323. Bibliothèque de l'École des Ponts et Chaussées, Paris.
[2] Cf. his monumental work: *Theory of analytic functions, containing the principles of the differential calculus, free from any consideration of infinitely small quantities or evanescents, of limits or of fluxions and reduced to the algebraic analysis of finite quantities*, Paris 1797.
[3] J. L. Lagrange, *Oeuvres*, t. X, p. 7.
[4] The framework of the lagrangian theory appeared in 1772, cf. his mémoire: "Sur une nouvelle espèce de calcul relatif à la différentiation et à l'intégration des quantités variables", *Nouveaux Mémoires de l'Académie Royale des Sciences et Belles Lettres de Berlin*, pp. 185–221.
[5] Newton gave lectures from 1673 to 1683. His lectures were published in 1707, cf. *Arithmetica Universalis, sive de compositione et resolutione arithmetica liber*, Cantabrigiae. Typis Academicis & Londini. Benjamin Tooke.
[6] I. Newton, *op. cit.*, p. 1.
[7] For more details, cf. the paper of H. G. Green and H. J. J. Winter: "John Landen F.R.S. (1719–1790), mathematician", *ISIS* **35**, (1944) pp. 6–10.
[8] "An investigation of a general theorem for finding the length of any arc, of any conic hyperbola, by means of two elliptic arcs, with some other new and useful theorems deduced therefrom", *Phil. Trans* **1775**, pp. 283–289.
[9] Cf. J. E. Montucla: *Histoire des Mathématiques*, Vol. III, Paris. A. Blanchard, 2e éd. 1968, p. 240.; Cantor, *Vorlesungen über Geschichte der Mathematik*, Vol. III, Leipsig 1894, pp. 842–847. For more details, cf. G. Mittag-Leffler, "An Introduction to the theory of the elliptic functions", *Annals of Mathematics* 1922–23; Ch. Houzel, "Fonctions elliptiques et intégrales abeliennes", *Abrégé d'Histoire des Mathématiques*, Paris, Hermann 1978, pp. 1–113.
[10] G. Vivanti, *Il Concetto d'infinitessimo et la sua applicazione dela matematica*, Mantova 1894, p. 57.
[11] Animadversions on Mr. Simpson Fluxions by Z. Tertius. Truth Triumphant: or, Fluxions for the Ladies. Shewing the cause before the Effect, and different from it; that space is not speed, nor Magnitude Motion, with a Philosophic Vision, most humbly dedicated to his illustrious Public, by $X$, $Y$, and $Z$ who are not of the family of $x$, $y$, $z$ but near relations of $x'$, $y'$ and $z'$ ... London printed for W. Owen, 1752.
[12] London, J. Nource.
[13] J. Landen, *op. cit.*, pp. 2–5.
[14] *Idem*, pp. 6–7.
[15] C. B. Boyer, *The History of the Calculus and Its Conceptual Development*, New York, Dover 1959, p. 237.
[16] It is only an anonymous letter signed with his initials A.B., addressed to the authors of the

*Monthly Review.*
[17] June 1759.
[18] J. L. Lagrange, *Theory of analytic functions*... Paris 1797, p. 4.
[19] "First, the title of *Residual Analysis* is no more than a new term given to Sir Isaac Newton's method of differences, and therefore is no new branch of *algebraic art*: since it has been known, and treated of by many, in a more easy and familiar manner than by Mr. Landen; especially, besides the inventor, by Brook Taylor; by Cotes in his *Harmonia Mensurarum*; by Stirling in his book called *Methodus differentialis*; and occasionally by many others.

But Mr. Landen surely dares not say this theorem is of his own invention, or that it was taken notice of before. He may, perhaps, imagine he has so disguised it by a new form, as to make it pass for his own, amongst credulous and ignorant readers. But to show that this curious invention is no new one, Mr. Laurin says, in his Algebra, page 109, art. 118. Generally, if you multiply $a^m-b^m$ by $a^{n-m}+a^{n-2m}b^m+a^{n-4m}b^{3m}+$ etc. continued till the terms is in number equal to $n/m$, the product will be $a^n-b^n$.

The author cannot sure plead ignorance, and say he has not read Mr Laurin's work. This would look ridiculous for one who cites in his works, L'Hospital, Bernoulli, Act. Erud. Leips. Archimedes etc. to pretend not to know the authors of his own country, and in his mother tongue", *Monthly Review*, p. 560.
[20] Landen in his reply in the July issue does not hide his annoyance of this "scientist": "He objects to *prime number, function* etc. as terms never heard before. Alas! how egregiously does he betray his ignorance".
[21] After the publication of his Discourse Landen who continued to work on Residual Analysis published only a paper about the summation of series. Cf. "A new method of computing the sum of certain series", *Phil. Trans.* **1760**, pp. 67–118.
[22] London, L. Hawes, W. Clarke and R. Collins, 1764.
[23] John Landen, *op. cit.* pp. 3–4.
[24] However, many years later, Landen uses the method of fluxions, cf. "A disquisition concerning certain fluents, which are assignable by the arcs of the conic sections; where are investigated some new and useful theorems of computing the fluents", *Phil. Trans.* **LXI**, 1771, pp. 298–309.
[25] Cf. the first pages of the Theory of analytic functions...
[26] Cf. also the book of Maseres: *Scriptores logarithmici*, 1791, Vol. 2, p. 170.
[27] John Landen, *op. cit.*, p. 2.
[28] *Idem.*
[29] "By the method of computation which is bounded on these borrowed principles, may be done, as well by another method founded on the *anciently-received* principles of algebra", John Landen: *A Discourse concerning Residual Analysis: A new branch of the Algebraic Art*, London 1758, pp. 4–5.
[30] John Landen, *The Residual Analysis, a new branch of the Algebraic Art, of very extensive use, both in Pure Mathematics and Natural Philosophy,* London 1764, p. 3.
[31] *Idem*, p. 4.
[32] *Idem*, p. 5.
[33] *Idem*, p. 6.
[34] For more details about function's concept, cf. Ch. Phili's conference at the Ecole Normale Supérieure, Seminar of Philosophy and Mathematics, Paris 1974, pp. 1–24.

[35] *Mém. de l'Academie des Science*, Paris 1718, p. 100, *Opera omnia*, t. II, Lausannae et Genevae 1742.
[36] *Introductio in Analysin Infinitorum*, Lausanne 1748, liv. II, ch. I.
[37] With the sign $x_1$ Landen expresses the similar function "if, in any given expression or function of $x$, wherein $x$ is not concerned, $x_1$ be substituted instead of $x$, the given expression and that which results from such substitution are called *similar functions* of $x$ and $x_1$ respectively", J. Landen, *op. cit.*, p. 5.
[38] *Idem.*
[39] "The said quotient, which algebraists commonly denote by $\frac{(y-y_1)}{(x-x_1)}$ or $\overline{y-y_1} \div \overline{x-x_1}$ we shall for brevity sake, sometimes denote by $[x/y]$; and the special value thereof we shall express by $[x \div y]$" *op. cit.*, p. 5.
[40] C. B. Boyer, *The History of the Calculus and Its Conceptual Development*, New York, Dover, 1949, p. 236.
[41] Lagrange continued in the same way. The second part of his Theory of analytic functions contains geometrical applications.
[42] J. Landen, *op. cit.*, p. 2.
[43] *Idem*, p. 6.
[44] We follow the original notation of the book, naturally these "brackets" play the role of parentheses.
[45] *Idem*, p. 6.
[46] Several commentators at the end of the eighteenth century called this formula "Landen's formula", cf. Maseres, *op. cit.*, p. 170.
[47] Cf. note 70.
[48] About Landen's demonstration, cf. chapter VIII "John Landen et les démonstrations algébriques" of M. Pensivy Thesis: *Jalons historiques pour une epistémologie de la serie infinie du binôme*. Thèse de 3ème cycle, Université de Nantes, Nantes 1986.
[49] J. Landen, *op. cit.*, p. 10.
[50] *Idem*, p. II.
[51] De Morgan says that in Landen's residual quotient exists the notion of the limit of D'Alembert: "It is the limit of D'Alembert supposed to be attained instead of being a terminus which can be attained as near we please. A little difference of algebraic suppositions makes a fallacious difference of forme: and though the residual analysis draws less upon the disputable part of algebra than the method of Lagrange, the sole reason of this is that the former does not go so far into the subject as the latter". *Penny Cyclopaedia Art*: Differential Calculus.
[52] J. L. Lagrange, *Theory of Analytic functions...*, Paris 1797, pp. 4–5.
[53] A. Clairaut, *Eléments d'Algèbre*, Paris 1797, Tom. I, p. 9.
[54] A. Clairaut, *op. cit.*, Tom. II, p. 90.
[55] A. Clairaut, *Eléments d'Algèbre*, 6th ed. Paris 1801, p. 325.
[56] Cf. J. Leslie's *Dissertation* on the progress of mathematical and physical science; John Leslie, Dissertation Fourth. *Encyclopedia Britannica*, 7th ed. vol. i. 1842, p. 601.
[57] "Therefore, having shewn how $x^n - x_1^n$ may be divided by $x - x_1$, it will now be proper to shew how $n^x - n_1^x$ may also be divided by the same divisor $(x - x_1)$. In doing we shall first assign the value of $n_1$ in a certain series of terms of $n$ and $x$, wherein the exponents of the several powers of these quantities shall be invariable: by the help of which series we shall be enable readily to obtain the desired quotient of $n^x - n_1^x$ divided by $x - x_1$". J. Landen, *op. cit.*, p. 30.

[58] Landen's method is the same as the one which Euler uses in his *Introductio in Analysin Infinitorum.*
[59] J. Landen, *op. cit.*, p. 31.
[60] "It may sometimes be of use to observe, that, when $v$ is very large number, the Log of $\overline{1+\frac{1}{v}}$ will be $= \frac{1}{v}$ nearly the value of the series $\frac{v^{-2}}{2} - \frac{v^{-3}}{3} + \frac{v^{-4}}{4}$etc. being there so very small that is may be neglected". *Idem* p. 33.
[61] *Idem*, p. 46.
[62] The titles of the ten chapters of his book are the following: "Terms and Characters explained"; "Of the inventions of Rules necessary to facilitate computations in this analysis"; "Of exponentials and logarithms"; "Of the properties of certain algebraic expressions"; "Of the tangents of curve lines"; "Of the investigations of useful theorems, by finding the nature of a curve from a given property of its tangents"; "Of the evolution and curvature of lines, with some inferences relating to the focuses of reflected and refracted rays and the curves called caustics"; "Of the greatest and least ordinates, the points of contrary flexions, and reflexion, and the double and triple etc. points of curve lines"; "Of the asympotes of curve lines"; "Of the diameters and centers of curve lines".
[63] J. Landen, *op. cit.*, p. 9.
[64] *Idem*, p. 46.
[65] *Idem*, pp. 50–51.
[66] "Certainly when we suppose the increments to vanish, we must suppose their proportions, their expressions, and everything else derived from the supposition of their existence, to vanish with them", G. Berkeley, *The Analyst, or a Discource addressed to an Infidel Mathematician. Wherein it is examined whether the Object, Principles, and Inferences of the Modern Analysis are more distinctly conceived, or more evidently deduced, than religious Mysteries and Points of Faith.* London 1734, § 13.
[67] "Infinite parve concipimus, non ut nihila simpliciter et absolute, sed ut nihila respectiva..., id est ut evanescentia quidem in nihilum, retinentia tamen characterem ejus quod evanescit", G. W. Leibniz, Lettre à Grandi, 6 Sept. 1713, M. IV, p. 218.; cf. also: "une quantité infiniment petite n'est rien d'autre qu'une quantité évanouissante, et c'est pourquoi en réalité elle sera égale à 0." L. Euler, *Institutiones Calculi differentialis*, St. Pétersbourg 1755, p. 77.
[68] P. Mansion: *Resumé du cours d'analyse infinitésimale à l'Université de Gand*, 1887, p. 213. For more details, cf. G. Vivanti, *op. cit.*
[69] J. Landen, *op. cit.*, p. 10.
[70] *Idem*, p. 3.
[71] Landen's profession was land surveyor.
[72] Cf. the anonymous letter in Monthly Review, notes 14 & 15.
[73] Cf. his paper on fluxions, cf. note 23; moreover, from his former publications we can establish that Landen never returned to his theory. Cf. *Mathematical Memoirs*, Vol. I, II, London 1780.
[74] Lagrange also had abandoned his own theory, seeing that his paper in 1772, made no impression and he had followed Leibniz's conception in his *Mécanique Analitique.*
[75] "The Residual Analysis rests on a process purely algebraical: but the want of simplicity... is a very great objection to it", *Monthly Review* Vol XXVIII, 1799. Appendix Review of Lagrange's *Theory of functions...*
[76] Cf. Landen's foreword; cf also "I can also mention a method which Landen gave in 1758 to avoid consideration of infinity, of motion or of fluxions... he is perhaps the only english

mathematician, who has acknowledged in inconvenience of the method of fluxions", Review of Lacroix's *Calcul différentiel. Monthly Review*, London 1800, p. 497.
[77] "Landen, I believe, first considered and proposed to the treat of fluxionary calculus merely as a branch of Algebra." R. Woodhouse, *Principles of Analytical Calculation*, Cambridge 1803, p. xviii. His homogeneous method was also appreciated by the philosophers Compte, Hegel, Cohen etc. Marx in his Mathematical Manuscrits stresses the importance of Landen's uniformity, cf. *The Mathematical Manuscrits of Karl Marx*, New Park Publication, New York 1983, pp. 33, 75, 113, 139.
[78] Moritz Cantor in his *History of Mathematics* criticizes severly Landen's Residual Analysis: "die residual Division, geschieht aber selbstverständlich durch Grenzübergang, so dass die Landenschen Methode nichts wesentlich Neues darbietet", M. Cantor, *Vorlesungen über Geschichte der Mathematik*, Band III, Leipsig, 1894, p. 661.
[79] J. L. Lagrange, "Discours sur l'objet de la théorie des fonctions analytiques", *Journal de l'Ecole Polytechnique*, 6[e] Cahier, t. II, Thermidor (Juillet-Août) An VII, p. 232.

MICHAEL OTTE

# HISTORIOGRAPHICAL TRENDS IN THE SOCIAL HISTORY OF MATHEMATICS AND SCIENCE

In this paper I will attempt to consider the connection between epistemology and the sociology of knowledge by sketching the historical development and I want to show how both come to bear upon, in different ways, similar issues in mathematical and scientific knowledge. Mathematics in fact becomes a touchstone for these attempts as a constructivist epistemology that is pointedly expressed in mathematical cognition, substitutes traditional empiricism, nominalism and behaviorism thereby providing a basis for understanding that connection.

Cassirer's and Durkheim's considerations on the concept of concept are used to support this claim. The issues will be discussed in the light of two "complementarities" regarding:
–the individual versus social determinism of human knowledge,
–the formal versus functional approaches to characterizing theoretical knowledge.
The overriding thesis defended here is that a socio-historical account of mathematical and scientific theory required the adaption of a complementarist approach toward the issues in a way that avoids the horns of each of the above two false dilemmas. In order for this to be achieved it is necessary that the diverse perspectives have common theoretical elements that permit their coordinated interpretation. Various such elements are discussed: activity, concept, consciousness and communication.

I start with the work of Karl Mannheim (1893–1947) whose book *Ideologie und Utopie*, already translated into English in 1936, represents still today a classical and scholarly introduction to the problems of the sociology of knowledge. Mannheim, however, excludes mathematics and logic from his theory. He writes: "The existential determination of thought may be regarded as a demonstrated fact in those realms of thought in which we can show ... that the process of knowing does not actually develop historically in accordance with immanent laws, that it does not follow only from the 'nature of things' or from 'pure logical possibilities', and that it is not driven by an 'inner dialectic'. On the contrary, the emergence and the crystallization of actual thought is influenced in many



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 295-315.

decisive points by extra-theoretical factors of the most diverse sort (1929, 239)".

This is commented by D. Bloor by saying that such an equating of social causes with 'extra-theoretical' factors confines the sociology of knowledge to the "sociology of error" (Bloor 1976, 8), as if the social explanation of true beliefs would automatically result from their truth. He is right, and the attitude he observes corresponds to our untheoretical everyday understanding which assumes that communicative contacts are impaired only where the truth of what is communicated does no longer seem certain. In this way, the very opinions we exchange daily and which are generally accepted are deemed to be true. The sociology of knowledge confines itself either to adopt the general consensus or to criticize it as false consciousness. It does not get far with this, however, as it cannot escape the suspicion that it itself only presents an opinion based on false consciousness. The sociology of knowledge thus must distinguish between knowledge and truth of knowledge on the one hand, and it must operationalize this distinction on the other, while refraining from considering it to be a quasi absolute and purely objective fact. Truth is a concept belonging to meta-knowledge. But knowledge and meta-knowledge are within the social context no more then two different operations which do not establish a fixed hierarchy. It is the distinction itself that matters theoretically.

One may observe that followers of the so-called "strong programme" (Bloor 1976, 5) of the Edinburgh School tend in fact themselves to contrast the logical with the social and to more or less explicitly identify social with irrational. They generally try to show that mathematics is not as rational as it is supposed to be. Processes of mathematical "proof" are accordingly a favorite object of their investigation. I also believe that D. Bloor and the followers of the strong programme are misguided in aiming merely at a *causal* model of social history of science and in refusing any functional or teleological explanations in this context. Concentrating on causality leads to the rather simplistic idea as if the social perspective were confined to the fact that humans are, in all what they do, engaged in social interactions with others. This views lead to social behaviorism (Bloor 1983).

As mathematicians shy away from behaviorism and as they concentrate on the demarcation problem between creative and mechanical behaviour, an interest much aggravated by the advent of the computer, a mentalist attitude prevails in most reflections about science and mathematics. It is the very point where the sociology of knowledge steps in, by critizising this mentalism. On the average this criticism confines itself, however, to labelling it again merely as false consciousness.

Brannigan, for instance, devoted much consideration to theories of discovery that had a mentalist or psychological cast (Brannigan 1981). He believes the sociological domain to be by far the most interesting, because it is the one in which one can search for clues which explain how problem solvers become interested in certain problems, how they represent the problems etc., etc.

But he is more interested again in social processes by which somebody may come to be called a genius or a problem solution is accepted as a scientific discovery. And he is not interested in the nature of knowledge which would explain its essentially social character.

Now coming back to Mannheim's book one may realize that it does not only lend itself as a starting point for criticism but also provides some very fundamental historical observations to which we now turn our attention. Mannheim introduces his book with considerations on the origin of modern epistemological, psychological, and sociological modes of reflection: "Epistemology was the first important result of the collapse of the unified world view which heralded modern times... All epistemological speculation is oriented towards the polarity of object and subject. Either it begins with the objective world, which it somehow dogmatically assumes as something familiar to all, explaining the subject's position in this world order by deriving the latter's powers of cognition from this world order; or it starts out with the subject as something immediately and unquestionably given, striving to derive the possibility of valid knowledge from this subject" (Mannheim 1929, 13).

The individual's introspection and the philosophy of consciousness were the very basis of epistemology. From Descartes to Kant and classical German philosophy, the effort was to represent cognition as a result of mental activity and to regard intuition as this activity's supreme expression. In the course of social differentiation, however, it became increasingly difficult to

> overlook that the subject formed by no means, as had been assumed before, a secure starting point from which to arrive at a new conception of the world. While internal experience is given in a certain sense immediately, and the inner connection of experiences can be grasped with more certainty if one is capable, among other things, of sympathetically understanding the motivations producing certain acts, it is clear that the risks of an ontology could not be totally avoided. Psyche, too, with all its internally immediately perceptible 'experiences', forms but one element of reality. And knowing of these experiences requires a theory of

reality, i.e. an ontology. As doubtful ontology may have become with regard to the external world, as dubious is it with respect to psychical reality, too (Mannheim 1929, 16).

Almost at the same time as Kant's foundation of epistemology, and in some way continuing it, sociology, and with it the social perspective on science, was created by A. Comte (1798–1857). John Stuart Mill described Comte's special achievement versus the psychologism and biologism of previous social theories as follows: "Since the phaenomena of man in society result from his nature as an individual being, it might be thought that the proper mode of constructing a positive Social Science must be by deducing it from the general laws of human nature, using the facts of history merely for verification. Such, accordingly, has been the conception of social science by many of those who have endeavoured to render it positive, particularly by the school of Bentham. M. Comte considers this as an error. ... As society proceeds in its development, its phaenomena are determined, more and more, not by the simple tendencies of universal human nature, but by the accumulated influence of past generations over the present. The human beings themselves, on the laws of whose nature the facts of history depend, are not abstract or universal but historical human beings, already shaped, and made what they are, by human society" (Mill 1865, 307).

Comte's achievement thus was the insight that societies, at a specific historical stage, had attained a degree of complexity that made it necessary to consider social phenomena to be relatively independent of the psychological or biological facts, just as psychology considers human consciousness to a certain extent independent of the facts of brain anatomy or of its genetic blueprint.

We can summarize the parallels in these problems with a diagram that should help to indicate that there is a structural analogy of the problems in the relation between brain and consciousness on the one hand and the relation between consciousness and society on the other:

"brain(matter)–consciousness–society."

Epistemologists would view the first relationship as that between object and subject while the sociologists of knowledge would discuss the second relationship as that between consciousness and communication. One essential point is that each case involves structural couplings between entities that principally differ in logical type. Neither matter nor the social can be understood as an all-encompassing context. In this sense, there is no causal relationship between social collective activity and individual consciousness and no direct interference. Instead one may speak more appropriately of different aspects of humanity's active encounters with reality. The fundamental concept when dealing either with subject object relations or with the problem of the interaction of individual and social, is the concept of "activity". The social and psychological laws can reveal themselves only in activity and through it. Sociologists of knowledge would concentrate on communication as an activity and philosophers or psychologists on the activity of concrete or symbolic constructions. We may describe these different systems of activities but we cannot subordinate them to one another. They are complementary to each other in the sense of the complementary aspects of the concept of concept as described by Cassirer (1910) and Durkheim (1912).

Mannheim considers that the orientation toward the example of mathematics, as found, for instance, in Kant, conceals the "active element contained in cognition" and thereby simultaneously covers up the superior character of the sociology of knowledge compared to epistemology (Mannheim 1929, 253). However, historically, the active character of human cognition has always particularly expressed itself in mathematics. It is nonetheless true that the essence or the basis of the dynamics of cognition as an activity has increasingly shifted from individual intuitions to social cooperation and communication. Kant's philosophy marks a certain turning point in this historical trend. Since Kant, there exists both an epistemological as well as empirical orientation in work on the theory of knowledge and a certain tension between epistemology and the sociology of knowledge. As was said above the new sociology of knowledge concentrates on this tension and makes the relation between epistemology and social science the focal point of its interests.

There exists two quite different attitudes with respect to the problem of the nature of mathematical statements like $2 + 2 = 4$, which correspond to the intent to stress either the features specific for mathematics in contrast to all other knowledge areas or to exhibit those aspects of cognition and knowledge that are common to all those areas.

So we either start by asking what distinguishes mathematical knowledge or activity from all other fields of thought, what is peculiar to it alone; or we set ourselves the task of characterizing mathematics as a part of human thought in general.

The first of these attitudes, being quite common among pure mathematicians, leads to conceiving mathematics as a formal science or syntactical game that is not at all influenced by social factors and has itself no social relevance either. As the positivists in the Vienna circle used to say, mathematics as well as logic is just one great cascade of tautological transformations. Nowadays mathematicians are more inclined to see themselves as sorts of artists, because they are afraid of being mixed up with computers. Or they follow some sort of Platonism. Platonism is more en vogue at present when asked to explain what mathematics is all about (Davis/Hersh 1981, Kitcher 1984).

The other approach mentioned would perhaps not deny that mathematics is a formal science as well as that proofs or activities of proof analysis constitute the greater part of it, but would understand the formal character of knowledge as an expression or a sign of a high degree of Vergesellschaftung (societal interdependency). Highly abstract and formalized knowledge makes sense only when there exist high degrees of social division of labor and cooperation. An algorithm is "socialized" operation in the same manner as the tool is individual human operation. A traffic light is part of a formal knowledge system as soon as there exists a social context in which it functions as part of such a system.

Worshipping the moon is more complex than Apollo 12 because believing is more complex than calculating; the social development and organization that enabled humans to set foot on the moon was, however, much more complex than that of an ancient traditional society worshipping the moon. With respect to logic and mathematics, it becomes obvious that individual human beings participate in knowledge communities but are not the location of that knowledge, because purely formal knowledge is nothing apart from its common approriation and apart from the social functions it fulfills in the service of social men. Machines, as is commonly said, cannot be fundamental explanations of anything; the concept machine is defined by its function (Pattee 1979, 131). Formal systems or algorithms or machines of all sorts have to fulfill a function but are not to be identified with any one particular function, although, as was said they are not to be conceived as being completely separated from the totality of all of their possible functions either. This implies a *relative* independance of structure and function. If a piece of formal knowledge, a written text, for instance, were defined just as a communicative function, this would not only lead to a psychologistic conception of meaning, which insists on the unanswerable question of "what did author really mean?", but would also exclude the seemingly strange and unexpected interpretation of a text as well as the possibility of radically novel insight associated with such an original and unanticipated application of the text (Collins 1990, 11). To appropriate a meaning is to appropriate an operation (Leontiev) and not a subjective inner experience or psychic constellation. Every piece of art, literature, or music is formal in the sense that it cannot be exhausted by any particular interpretation, and it is functional as it is nothing apart from its reception and interpretation by humans. It is this complementarity of form or representation and function or process that seems to be essential.

A purely functionalist conception, identifying knowledge with its use or a text with a *particular* interpretation, would narrow down functionality itself. This seems paradoxical at first sight but it follows readily from what was said. Following too closely one specific function excludes all other possible functions of a particular design. With respect to the history of science the reaction to this situation is mirrored in "Bernal's paradox" characterizing the evolution and transformation of science during the early 19th century. "At the time when science should have been most obviously connected with the development of the machine age arose the idea of pure science" (Bernal 1963, 29).

All knowledge entails intention. Knowledge is never to be reified. The map is not the territory. Theories are realities sui generis, that are, however, not without purpose or function, although they must not be identified or confused with any particular application or with any specific purpose either.

We may again use the issue of printed text as an illustrative example (cf. Otte 1986). Text as communicative function or subjective interpretation vs. text as materialized structure, as a representation of knowledge provide two complementary aspects of cognition that save us from the paradox that meaning on the one hand is a whole not to be reduced to the individual meanings of particular words and that on the other hand texts, as well as reality in general, are intelligible only if they are accessible to segmented explanation. The most important prerequisite of knowing and cognition is having the simultaneous experience of knowledge and its use. Principally speaking this simultaneous experience of structure and function is provided by social cooperation. We have already mentioned this when dealing with the "social character" of mathematics.

These claims may be generalized in that a characteristic of modern abstract science is the multitude of perspectives used by different researchers. The significant differences between research fields lie less often in what it described than in how it is described. One important way of developing a science is to apply new perspectives to a part of reality, thereby highlighting new features of this reality. Perspectives determine what data are seen, what theories are developed, and what kinds of result turn up.

But remember: Scientific knowledge is a second order phenomenon resulting from the observation of observation. Communication has for instance to rely on meta-communication, as Bateson has repeatedly stressed (Bateson 1972). Effective communication depends on the distinction of the communicative activity from other behaviour. From this results a distinction between information and communication the plurality of perspectives provides just the grounds for the process of the recursivity of the acts of observation. It is therefore irrelevant whether we attribute the pluralism to the manifoldness of reality or to the multiplicity of perspectives. It is the multiplicity of distinctions, which as a connected network are the instruments of cognition, that counts, that keeps the cognitive process going. As I see it, recent sociology has developed along these lines. It does not intend to be a sociology of error and strives for an alternative to a mere historization and relativation of viewpoint. The sociological approaches refer to mathematical activity in a specific way. From mathematics, they take an extremely formal concept of operating and of observing conceived of as a distinguishing description. Observing is defined as unity of distinguishing and designating and thus really is a basic operation of mathematics. Acting and observing are, however, complementary to each other because acting does not designate itself as what it is. It needs a second observation to constitute an act as an observation. "An operation that uses distinctions in order to designate something we will call observation ... An observation leads to knowledge only insofar as it leads to re-usable results in the system. ... The first distinction is the observation itself, distinguished by another observation..." (Luhmann 1990b, 69–73).

The basis of operative mathematics is indeed a distinction without "reason". The concept of the element in mathematical ontology is "the particular absolute, without any real contents,... for it is simply determined as something different without any real contents being defined with regard to that from which it is different", as Grassmann wrote in 1844. Imagine a geometrical point in the sense of Euclid's description. It has no length, no width, it has no properties at all. According to Leibniz' principle of the identity of indiscernables two such points cannot be distinguished and therefore do not really exist at all (Leibniz' principle consists in the thesis that there are no two substances which resemble each other entirely but only differ numerically). Grassmann on the contrary believed that geometrical points can be taken as a perfect instance to illustrate the act of distinguishing as reflecting will with no motive or further reason. On such a constructivism the so-called hypothetico-deductive structure of mathematics relies and it is a conception of the kind on which its set theoretical ontology is based. An abstract set in the sense of Cantor is a collection of objects that are without any particular further properties. Set theory studies the properties these collections have by virtue of being collections, no matter what they are collections of. Thus, the operation of distinguishing is indeed that which can be observed first. In other terms: all mathematics begins by drawing a distinction which simultaneously involves a designation. If one distinguishes the "red things" from the rest of the world one usually designates this set by, let us say "$x$". Mathematical thinking starts by an observation of the first observation. The concept of equation $a = b$ or the concept of set are expressions of this basic situation. In an equation, something first distinguished is set as something equal.

To acknowledge the fact that the multitude of perspectives and individualism of outlook are a product of the social, a result of cultural history and not of nature, leads to the claim that science, despite all its structural diversity and incoherence, exhibits another type of identity, namely *functional* identity as well as the fact that this identity must be represented.

This idea of a functional or social identity of science is, as has already been said, a result of the experience of the beginning industrial Revolution. It was clearly expressed, for instance, by A. Comte in his "Discours sur l'esprit positif". Because of the "inevitable diversité entre les phenomenes fondamentaux", we should objectively, that is, understanding our theories "comme exacte representations du monde reel", search for no other unity than that of method, said Comte. Matters are, however, quite different from the subjective aspect, where we consider theories "comme des resultats naturels de notre evolution mentale ... destinés à la satisfaction normale de nos propres besoins quelconques. ... On ne doit plus alors concevoir, au fond, qu'une seule science, la science humaine on plus exactement sociale ..." (49).

From the social perspective knowledge and experience have an interest only in terms of functional equivalence. Their functionality is related to the complexity of the social life itself that knowledge serves, and correspondingly, the functionality of knowledge can also have a very complex social character such that it may, for instance, not be represented by a coherent formal system. As we don't share Comte's idea that there is only one context, the social context and only one science, sociology, function and structure of knowledge are really complementary to each other and are not to be subsumed to one another. It is, however, not at all clear how to conceptually fix this complementarity. The following remarks represent only a first and preliminary approach to the problem.

The question arises: is it possibly true that what is constant in the process of evolution of science are some of its functions in human culture? ... Knowledge fulfills two major roles in human society: a practical one and a philosophical one. ... it seems that science came into being with the requirement of ... coherence and that one of the functions it performs permanently in human culture consists in unifying ... practical skills and cosmological beliefs, the *episteme* and the *techne*. ... despite all changes that science might have undergone, this is its permanent and specific function which differentiates it from other products of human intellectual activity (Amsterdamski 1975, 41–44).

Society lends its own complexity to give birth to science, and science conversely functions to produce generality and coherence where there were none before. Is this the formula that may be used to found a social theory of abstract science? Does it really work smoothly and "globally"? Not quite.

It has to be remembered that the concept of functional identity permits very different things for the various forms of knowledge and scientific disciplines. Functional equivalence is only a very weak constraint and it is one that can sometimes precisely stimulate the plurality and variety of phenomenal forms rather than restrict it. Thus one additionally has to take into account that science and technology not only produce order from chaos, but are at the same time themselves a source of disorder and chaotic complexity. The fact that no point of absolute and static balance can be found, that coherence and consistency are possible only locally, is what makes a social theory of knowledge possible and necessary. In the light of Gödel's incompleteness theorem, we know that even a mathematical system cannot escape incompleteness or internal contradiction. Gödel's own reaction to this problem consisted in Platonism. Platonism expresses an attitude that views of mathematical cognition as a process transforming chaos into order by ever clearer insights and recognitions of the yet undiscovered and unknown. Every piece of knowledge increases the clarity of our view of the Platonist universe. We have, however, to accept that human creativity lies exactly in the movement from chaos to order and back from order to chaos again, in the mind's capability to produce *and* destroy order. Mind is, according to Charles S. Peirce (1839–1914), "to be regarded as a chemical degree a habit of taking and laying aside habits" (6. 101).

Every scientific achievement removes some uncertainties and produces order from chaos, at least for the human being who has found the solution to a problem or the structure that helps to organize his/her perspectives. For everybody else however, the situation is quite different. As the new information or the new idea increases the number of possible alternatives (from which to choose, to move, to view reality, etc. ...), indeterminateness and uncertainty also grow. Formalization or formal explicitness in particular is a source of possible bifurcation. One should remember the history of the axiomatic method, for instance, and the role it played in the discovery of non-Euclidean geometry or in the development of set theory, with or without continuum hypothesis, to mention but two examples (Otte 1992).

The fact pointed at by Mannheim, namely that epistemology as well as social theory have a common origin, directly relates to this idea of a functional identity of science. More specifically stated, problems in an understanding of the function of theoretical concepts are of fundamental importance for the philosophy as well as the sociology of science. I want to document this to some extent by referring to the work of Cassirer on the one hand and Durkheim on the other.

Both Cassirer (1910) as well as Durkheim (1912) start off their argumentation by addressing some problems related to the Aristotelian view of the concept. It is remarkable that we encounter a focus on the same set of problems both in Kant's succession and in that of Comte.

The problem on which they both base their arguments can be illustrated in the following question: How can the theoretical concept, as it is produced by mere abstraction, that is, by neglecting and discarding all specific characteristics of the real phenomena, help us to orient ourselves within the world of these phenomena?

Ernst Cassirer used a conception that was first expressed explicitly by Kant to describe the growth of modern science as a transition from thinking in substances to thinking in relations. "In opposition to the logic of the generic concept, which ... represents the point of view and influence of the concept of substance, there now appears the *logic of the mathematical concept of function*" (Cassirer 1953, 21). For Cassirer, the significance of this new understanding of the theoretical concept lies in the fact that it retains the determinations of the special cases that become lost if the concept is conceived as an abstraction. He wrote:

In this criticism of the logic of the Wolffian school, Lambert pointed out that it was the exclusive merit of mathematical 'general concepts' not to cancel the determinations of the special cases, but in all strictness fully to retain them. When a mathematician makes his formula more general, this means not only that he is *to retain* all the more special cases, but also be able *to deduce* them from the universal formula (Cassirer, 1953, 20).

It follows from this perspective that the theoretical concept cannot just be defined in terms of reference; namely by extension, that is, by counting all objects that belong to it, or by criteria that decide whether an object belongs to it or not. The meaning of a concept is not just that to which it refers. It is rather that the concept also expresses a constructive, active element of knowledge and an intentionality that requires the concept to be defined "not by counting that which falls under it, but purely intensionally by giving a specific propositional function" (Cassirer 1974, vol. III, 352). This understanding also changes the relation between concept and object. The abstractive concept initially calls for the precondition that objects have objective being and are fixed data. The object must first be fixed as an objective substance through relations of equality and difference before it can be attributed properties or conceived in relations. If, in contrast, the concept is given as a characterization function or propositional function as such, in relative independence from objects, then "the determinateness, that we can ascribe to the 'matter' of knowledge, belongs to it only relatively to some *possible order* and thus to a formal *serial concept*. ... Matter *is* only with reference to form, while form, on the other hand, *is valid* only in relation to matter" (Cassirer 1953, 310/311).

While Cassirer sees the generality of the concept virtually in the fact that a certain conceptual mode of construction can be used multifunctionally, and while his theory of the concept is more concerned with the logico-psychological context of the creation and development of concepts. Durkheim's focus is on the social generality of the concept and on the latter's social function. Social function here means that the most diverse individuals having had most different experiences may have interpretative access to one and the same conceptual form. Durkheim writes in 1912:

Logical thought is made up of concepts. Seeking how society can have played a role in the genesis of logical thought thus reduces itself to seeking how it can have taken a part in the formation of concepts.

If, as is ordinarily the case, we see in the concept only a general idea, the problem appears insoluble. By his own power, the individual can compare his conceptions and images, disengage that which they have in common, and thus, in a word, generalize. Then it is hard to see why this generalization should be possible only in and through society. But in the first place, it is inadmissible that logical thought is characterized only by the greater extension of the conceptions of which it is made up. If particular ideas have nothing logical about them, why should it be different with general ones? The general exists only in the particular; it is the particular simplified and impoverished. Then the first could have no virtues or privileges which the second has not.

Therefore the concept must be defined by other characteristics. It is opposed to sensual representations of every order – sensations, perceptions or images – by the following properties. Sensual representations are in a perpetual flux; they come after each other like the waves of a river, and even during the time that they last, they do not remain the same thing. ... On the contrary, the concept is, as it were, outside of time and change. It is a manner of thinking that, at every moment of time, is fixed and crystallized. ...

And at the same time that it is relatively immutable, the concept is universal, or at least capable of becoming so. A concept is not my concept; I hold it in common with other men, or, in any case, can communicate it to them. It is impossible for me to make a sensation pass from my consciousness into that of another; ... On the other hand, conversation and all intellectual communication between men is an exchange of concepts. The concept is an essentially impersonal representation; it is through it that human intelligences communicate.

If the concept is common to all, it is the work of the community. Concepts

are the work of society and are enriched by its experience. That is what makes conceptual thought so valuable for us. If concepts were only general ideas, they would not enrich knowledge a great deal, ... But if before all else they are collective, representations, they add to that which we can learn by our personal experience all that wisdom and science which the group has accumulated in the course of centuries. ...

Thinking conceptually is not simply isolating and grouping together the common characteristics of a certain number of objects; it is relating the variable to the permanent, the individual to the social (Durkheim 1912, 432–439).

In Durkheim's theory the concept has lost its intentionality and individual consciousness has become something derived from societal consciousness.

In summary we could say that Cassirer's and Durkheim's views together lead us back to our diagram "matter(brain)–consciousness–society" or rather to a somewhat transformed version of it, namely

"operation–concept–communication".

By summing up Cassirer's and Durkheim's contributions towards clarifying the problem of the concept, and by synthesizing these in a diagram, we obtain a presentations which systematizes the means of cognitive activity in a mode analogous to our first diagram. First, we have the intentional element of the concept which dissociates it from a conception identifying knowing with communication, i.e. with statements about what is actually present (see the theoretician's dilemma, Tuomela 1981, 6).

This intentionality of the theoretical concept is due to the fact that its possible objective applications are not all specified or even known in advance. The social perspective on knowledge does not encompass the individual intuitions connected with the objects, but it has a certain degree to coordinate the individuals actions and conceptual moves. It seems not clear at all how this occurs. It requires that order is not in contrast to complexity but is an aspect of it. This, however seems best to be accomplished in the case of mathematics because the latter is based on a few peculiarly sharp concepts. To the individual mind, representation in language appears to be an impoverishment and falsification of the wealth of its mental experiences and intuitions. From the perspective of language and communication, however, it can be seen that no thought can be exhausted by an individual interpretation and experience and thus is infinitely more complex and manifold than can be realized by an actualization or interpretation.

From the perspective of the individual consciousness, the development of the history of science is dependent on which individual intuition or experience finds its way into the explicit formulations of a theoretical proposition. In this sense, Bateson stresses that our theory of evolution would have been very different if Wallace had anticipated Darwin. He writes:

> It is argued for instance that in 1859, the occidental world was ready and ripe (perhaps overripe) to create and receive a theory of evolution that could reflect and justify the ethics of the Industrial Revolution. From that point of view, Charles Darwin himself could be made to appear unimportant. If he had not put out his theory, somebody else would have put out a similar theory within the next five years. Indeed, the parallelism between Alfred Russel Wallace's theory and that of Darwin would seem at first sight to support this view. ... But, of course, it does matter who starts the trend. If it had been Wallace instead of Darwin, we would have had a very different theory of evolution to-day. The whole cybernetics movement might have occurred 100 years earlier as a result of Wallace's comparison between the steam engine with a governor and the process of natural selection. ... It is, I claim, nonsense to say that it does not matter which individual man acted as the nucleus for the change. It is precisely this that makes history unpredictable into the future. This error is a simple blunder in logical typing, a confusion of individual with class (Bateson 1979, 46–48).

Nonetheless, without the potentials and problems given by the sociocultural context, even the greatest of talents would be powerless. The process of cognition and scientific development is inconceivable without its social functions and can therefore also not be conceived independently from its social context. The development of knowledge does not take place within the framework of natural evolution but within the frameworks of sociocultural development.

It seems plausible that the debate on the social character of formal knowledge would gain a great deal of stimulation from the development of the computer. Much debate, for instance, has arisen from the thesis that a machine cannot do anything without having a rule telling it to do so. To this argument, Hofstadter answers: "... both machines and people are made of hardware which runs all by itself, according to the laws of physics. There is no need to rely on 'rules that permit you to apply the rules', because the 'lowest' – level rules – those without any 'metas' in front – are embedded in the hardware, and they run without permission" (Hofstadter 1979, 685).

I agree and I disagree with Hofstadter and others from the area of AI. I agree with his refuting the "fallacy of the rule" (Bourdieu). It is completely unplausible to attribute all concrete or cognitive activities of humans to conscious decision and explicit prescription. But to base the explanation on a reductionism to the laws of physics is no less mystifying than to worship an illusionary total awareness. So we are led to accept lawful structures not only on the physical level but also on the level of the social system to the effect that these structures are objective for the cognizing subject. One might be tempted to say: Society also runs without permission, "all by itself". With respect to the determination of consciousness, we may claim that even a person's most intimate experience is a conscious attitude only in so far as her dealings with it involve explicit knowledge from the pool of symbolic means of reflection. Knowledge is necessarily social knowledge. However, only knowledge of a personal kind can be true, because there is no such thing as consciousness as such or in general (cf. Popper's World 3). Knowledge advances neither by unmixed immediacy nor by unmixed mediation, as Hegel had written, and no complete determination by either the lower material context or by the higher social context exists.

Over the endless dark centuries of its evolution, the human nervous system has become so complex that it is now able to affect its own states, making it to a certain extent functionally independent of its genetic blueprint and of the objective environment. A person can make himself happy, or miserable, regardless of what is actually happening 'outside', just by changing the contents of consciousness (Csikszentmihalyi 1990, 24).

This argumentation runs completely parallel to the justification of the social level of knowledge as a sphere, that is to a certain extent functionally independent of the nature of individual consciousness (cf. the discussion on Comte above).

We must view consciousness at the same time with regard to the social determinations, and we must relate the two boundaries to one another. This double complementarity of consciousness with regard to brain on the one hand and to society on the other is also already found with the classical authors Kant or Comte.

On the basis of these considerations, it does not seem surprising that social theories of knowledge are stimulated by the computer and the problems of artificial intelligence. A lively debate is, for instance, to be found in Vol. 19 (1989) of *Social Studies of Science*, and, even more recently, one of the scholars most actively engaged in that debate has published a book on *Artificial Experts* (H. M. Collins, 1990).

The controversial debate was triggered by Peter Slezak's claim that certain computer programs "constitute a pure or socially uncontaminated instance of inductive inference and are capable of antonomously deriving classical scientific laws from raw observational data" (563). These developments, according to Slezak, represent the kind of "mentalism" that is considered as incompatible with a social explanation of science by people like Bloor (1976) or Brannigan (1981). This provides a refutation of the so-called "strong programme" in the sociology of scientific knowledge (SSK).

It is interesting to note that that which constituted a decisive contrast in the writings of philosophers like Lucas, Searle, or Dreyfus, to name but a few, is not put together in opposition to sociologism or social determinism. Philosophical criticism of the Al claims based on a difference in type between consciousness and brain (be it natural or artifical). This philosophical criticism of computer science is copied as sociologism is now identified with a sort of behaviorism. Slezak develops his criticism from Chomsky's challenge to Skinner's behaviorist program. The central argument is based on the fact that we have from the outside no direct access to the human mind and that we cannot directly control its proper functioning by external means. In pedagogy this fact is expressed by saying: You can lead a horse to water but it has to drink by itself.

The notion that theories and beliefs might have external social causes is, according to Slezak, essentially a commitment to the stimulus control of behavior.

> The claim that there could be a science which objectively and independently identified the causes of behaviour was thereby vitiated. Given the profound complexities of the phenomenon, Skinner's claim to have discovered its causes was pure hybris. Chomsky demonstrated that Skinner's 'astonishing claims' of causal correlations between external stimuli and behaviour were entirely without foundation. In this latter respect, the parallel with the claims of SSK is striking – though it is no accident, and no mere parallel. In SSK we are dealing with essentially the same phenomenon (complex human behaviour and attempts to identify its causes). In claiming to identify the external (social) causes of complex human behaviour and belief, SSK has always been clearly committed to a behaviourist methodology, and this has recently become explicit. The strong programme, in seeking external, social causes for beliefs and rejecting internal, psychological factors is, I believe, simply a restatement of the Skinnerian stimulus-response methodology – as the later endorsements by Bloor and Shapin have confirmed (Slezak 1989, 589/590).

Slezak then proposes to consider the two approaches, the psychological one and the sociological, as *complementary*. "The apparent contest between conflicting doctrines is largely an artifact arising from the sociologists' confusions about their own project. Thus, they have explicitly attacked 'psychologism' and traditional epistemological approaches, though these may not be as pertinent to their fundamental concerns as they have imagined" (576).

Slezak does not explain how the term *complementarity* might be understood. Complementarity is more then mere division of labor because it involves the claim that only both views together can provide an appropriate description. Nevertheless, it seems that complementarity of consciousness and communication is fundamental for the recent sociology of knowledge (cf. Luhmann 1990a).

Several conclusions should be drawn from this.

1. Theories and entities of their own kind. The map is not the territory. The map of the map is also to be distinguished from the map but in this second distinction there is no definite hierarchical order involved. Or to take a mathematical example: One can by means of measurement assign a length value to the table but cannot speak about the table of the length. Similarily one talks of the logarithm of the temperature but not of the temperature of the logarithm. However, within mathematical theory numbers, functions and other entities exist as operators as well as objects. One has to distinguish these complementary aspects as only together they characterize mathematical concepts but there is no absolute hierarchy involved.

Mathematics always entails meta-mathematics. This fact is presented in the complementarity of form, structure or object on the one hand and function, algorithm or operator on the other. Stated in different terms: Contrary to the traditional foundationalist interests which attribute to meta-mathematics a direct philosophical relevance while considering mathematics just as a certain type of technology, meta-mathematics is mathematics and vice versa: There is no universal awareness available in advance and there is therefore no other way than to proceed constructively.

All cognition begins by drawing a distinction which simultaneously involves a designation and by observing this first observational act as in an equation $a = b$. All calculation, for instance, is based on the distinction between 1 and 0, as Leibniz has already observed. What Leibniz did not see as clearly was the difference in categoricity or logical type between concepts and real objects. The 0–1 distinction belongs to the map and not to the territory, it is part of the description. As long as one has just the description itself in mind and does not want to communicate the knowledge or to draw conclusions from it, this confusion of types, or this reifying of conceptual constructs, may not lead to much harm. If, however, equations $a = b$ occur, problems may arise. That means that the situation is changed as soon as the same object turns up in different symbolizations such as during communicative activities, when both partners operate on the basis of their own respective observations and their own related descriptive characterizations.

The problems of equations or axioms seem unavoidable as soon as we ask for a map that is less complex than the territory, that means, as soon as we ask for a description that is functionally co-determined. In this manner, the social system of communication brings about the idea that theories and other knowledge systems are realities *sui generis* and are principally to be distinguished from the object fields to which they refer.

2. Secondly, we conclude that questions of meaning and truth of knowledge cannot be treated in the frame of just one closed coherent theory. One might be tempted to say the theory is not to be identified with the theoretician or with science itself. Mathematics is not the same as the attitudes that mathematicians bring to it. It is the activity as a process in time that matters from a social point of view. The paper started with a question of how to develop a sociology of knowledge and truth as well as to an operative conceptualisation of this distinction as essential prerequisites. This brought the act of observation, defined as a unity of distinction and

designation, into focus. The recursive network of observational acts provides the conceptual starting ground. This separates the work of art or of theory not only from the object field but also from the individual human subject, be it the author or the observer. (I have elsewhere (Otte 1991) argued that this double distinction occurred during the industrial revolution.)

We therefore cannot conceive of the social context as an all-encompassing box in which everything is contained. Individual consciousness is not embedded into social intercourse. And collective solipsism seems as conceivable as individual solipsism. The social rather seems to be a sort of layer to be principally differentiated on arguments of type or categoricity from other levels of cognition and existence, but is structurally analogous to those. It is also ruled by the complementarity of representations and process. This implies that strictly causal models of social determination of knowledge, as outlined in the so-called strong programme by David Bloor, seem not feasible. Philosophy of science searches the solution to the problem of objectifying knowledge by presupposing that objectivation requires alternative paths of access to the object (Feher 1981).

In analogy to that, it is plausible that the sociological perspective is confronted with the pluralism of social groups and the incommensurability of paradigms and theories. If the sociological perspective intends to add something of its own to this, it must leave the context of science altogether and subsume the practice of everyday knowledge, of technology, industry, mystical wisdoms, art, and so forth under "alternative approaches to the object". A social theory of science cannot be a causal theory of explicit theoretical knowledge but has to place its conceptualizations within the context of all the different aspects of humanity's encounters with reality.

3. Related to the above is another problem. How can we separate foundation – considered mainly the business of formal logic – and growth of mathematical knowledge, taking into account the fact that there is no innate universal cognitive aptitude from which this knowledge arises quasi automatically (if there were such universals of human cognition scientific knowledge and ordinary knowledge would be of one kind and would arise in the same quasi automatic manner).

To explain a statement like 2+2=4, one first argues like in discourse on ordinary knowledge, saying that this proposition expresses a simple fact of reality to be easily verified (which however is in itself independent of such verification). After a while one goes on, completely as in the case of science, to try and give an *explanation* of this fact. In this endeavour one uses the general and abstract to explain the particular and concrete, or seemingly concrete, in exactly the same manner in which Newton's laws are used to explain simple mechanical phenomena, or Ohm's law is used to explain the facts of electricity. The general, as used in scientific explanations of such kind, in our case for instance the associative law of algebra, is less sure and less positive than the concrete facts to be founded on it. The less certain is used to explain the more certain. Such a strategy makes sense only if it is employed exploratively and predictively.

In this manner the foundation and the growth of mathematical and scientific knowledge become intertwined. Foundational problems of knowledge and questions concerning its development cannot be treated independently from each other, and we end up with a complementarity of epistemological and sociological considerations of mathematics, which is based both on connection and distinction. Up to this century science and mathematics were conceived as founded on firm and unchanging ground (be it commonsense, or logic, or scientific method) and the claims of an interrelationship between epistemology and social theory appeared as rather artificial. Only when abstraction and theorization became predominant and the interdependance of the problems of the genesis and the foundation of knowledge gradually imposed itself on our thinking, did things change and the idea of complementarity gained some importance.

This complementarity of form and function or of representation and process can be interpreted in many different contexts. Comparing Cassirer's and Durkheim's discussion of the concept of concept we gain two such different interpretations of this complementarity which essentially connect epistemology and social theory of science. Complementarity in particular expresses the social nature of mathematics as it characterizes mathematical epistemology.

*University of Bielefeld, Germany*

## REFERENCES


Amsterdamski, A., 1975, *Between Experience and Metaphysics,* Kluwer, Dordrecht.

Bateson, G., 1972, *Steps to an Ecology of Mind,* New York.

Bateson, G., 1979, *Mind and Nature,* New York.

Bernal, J. D., 1963, *The social function of science,* Cambridge/USA.

Bloor, D., 1976, *Knowledge and Social Imagery,* London.

Bloor, D., 1983, *Wittgenstein – A Social Theory of Knowledge,* London.

Brannigan, A., 1981, *The Social Basis of Scientific Discoveries,* Cambridge.

Cassirer, E., 1910, *Substance and Function,* Engl. edition, Dover N.Y. 1953.

Cassirer, E., 1974, *Das Erkenntnisproblem in der Philosophie und Wissenschaft der neueren Zeit,* vol. II, reprint, Darmstadt.

Collins, H. M., 1990, *Artificial Experts,* Cambridge/USA.

Csikszentmihalyi, M., 1990, *Flow,* New York.

Durkheim, E., 1912, *The Elementary Forms of the Religious Life,* English translation. Glencoe/III., 1915, reprint 1926.

Feher, M., 1981, "Some Remarks on Meaning Invariance and Incommensurability", *Science of Science* No. 3-4 (7-8) vol. 2, pp. 431/432.

Hofstadter, H., 1979, *Gödel, Escher, Bach,* Hassocks, U.K.

Luhmann, N., 1990a, *Die Wissenschaft der Gesellschaft,* Frankfurt/M.

Luhmann, N., 1990b, "The Cognitive Program of Constructivism and a Reality that Remains Unknown", *Selforganization,* W. Krohn, G. Küppers and H. Nowotny (eds.), Kluwer, Dordrecht, pp. 64–85.

Lukes, S., 1973, *Emile Durkheim,* Penguin Books.

Mannheim, K., 1929, *Ideology und Utopie,* Frankfurt/M., reprint 1965.

Mehlberg, J. J., 1962, "A Classification of Mathematics Concepts", *Synthese* **14**, pp. 78–86.

Mill, J. S., 1865, "Auguste Comte and Positivism", *Essays on Ethics, Religion and Society,* Coll. Works, vol. X, Toronto, reprint 1969.

Mormann, Th., 1991, "Neuraths Enzyklopädismus: Entwurf eines radikalen Empirismus", *Journal for General Philosophy of Science* **22**, pp. 73–100.

Otte, M., 1986, "What is a Text?", *Perspectives on Mathematics Education,* B. Christiansen, A. G. Howson and M. Otte (eds.), Kluwer, Dordrecht, pp. 173–203.

Otte, M., 1990, "Intuition and Logic", *For the Learning of Mathematics,* vol. 10, pp. 37–43.

Otte, M., 1991, "Style as a Historical Category", *Science in Context* 4 (2), pp. 233–264.

Otte, M., 1993, "Towards a Social Theory of Formal Knowledge", *Learning from Computers – Mathematics Education and Technology,* W. Dörfler, K. Keitel, and K. Ruthven (eds.), Springer, Heidelberg.

Pattee, H. H., 1979, "Discrete and Continuous Processes in Computers and Brains", *Lecture Notes in Biomathematics,* Heidelberg, pp. 128–148.

Peirce, Ch. S., 1965, *Collected Papers,* 6 vols, Cambridge/Mass.

Slezak, P., 1989, "Scientific Discovery by Computer as Empirical Refutation of the Strong Programme", *Social Studies of Science* **19**, pp. 563–600.

VASSILI Y. PERMINOV

# THE CONCEPTION OF THE SCIENTIFIC RESEARCH PROGRAMS AND THE REAL HISTORY OF MATHEMATICS

1. Modern studies in the history of science have two aspects. On the one hand, they are directed to the elucidation of facts and connections in the real history of science and, on the other hand, to the discovery of the general logic of history, and its general schemes and laws. This second aspect is essentially different from the first, and having a philosophical rather than historical-scientific character. We have sufficient grounds to assume that each one is indispensable to the other, and that the philosophical image of science is essentially involved in concrete historical studies.

2. The conception of scientific research programs proposed by I. Lakatos in the sixties is one example of such general philosophical schemes of the history of science. Its basic concepts are well known: the hard core of the theory, empirical content of the theory, its protective belt, logical and heuristic counterexamples, progressive and regressive shifts of problems, etc. Without a doubt, Lakatos's conceptual system is good enough to describe the logic of development of empirical theories. However, Lakatos had pretended to a much wider sphere of application for his conception. In his view, this same scheme is also appropriate for mathematical theories and their development, and he has spent a great deal of effort to justify this thesis.

The examples from the history of mathematics that Lakatos considers in his works are very deep and interesting. In my opinion, however, Lakatos has never provided solid ground for his main thesis. An abundance of facts and reasons demonstrates that mathematics follows an essentially different logic of development in comparison with the empirical sciences. Let me discuss some central issues underlying the specifics of mathematical cognition and its history.

3. First, it should be noted that mature mathematical theories, in contradistinction to empirical ones, have only a hard core, but no protective belt in the proper sense. The protective belt in empirical theories is formed by particular propositions which may vary greatly, without affecting the basic premises of the theory. In the empirical



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 317-325.

sciences the general principles do not determine, in a deductively unique way, the system of particular propositions, but only set the frames of their variance. In mature mathematical theory, on the contrary, no change of any particular proposition is possible without loss of at least one axiom.

In the case of an empirical science, a contradiction which may appear between its principles and the facts can be removed by the adoption of an *ad hoc* hypothesis. On the contrary, in a mature mathematical theory this is impossible. Removal of a contradiction or paradox from a mature mathematical theory can be done only by correction of the axioms in the body of the theory or by the rejection of definitions of a certain kind – that is, only by carrying out some essential modifications in the very foundation of the theory. The cause is quite clear: a mature mathematical theory proceeds from a strictly established system of initial principles and definitions, while new principles are not admitted until they are provable on the basis of the accepted system of initial principles.

Ad hoc hypotheses are used in mathematics, but only in an initial stage of development of theory or in some young branches of mathematics, when the objects involved in the theory have not yet been logically defined, but are assumed intuitively clear. A counterexample can be removed in this case by mere reinterpretation of the initial definitions, like in Lakatos's treatment of Euler's theorem for star-polyhedra. Since any reinterpretation is actually an implicit modification (of the definition) of the concept this is possible only in the formative stage of the conceptual system of the theory. In mature mathematical theory such a variance of initial concepts and their definitions is not admissible. Ad hoc hypotheses at this stage are not acceptable and, consequently, the idea of the protective belt of a theory becomes, in this case, meaningless.

Lakatos has not made any clear distinction between the initial and mature stages of mathematical theory, which plays a decisive role in the considered case. All his counterexamples concern concepts which are not yet strictly defined, for example, the concept of "polyhedron" in his *Proofs and Refutations.*

In the initial stage of development of a mathematical theory various implicit definitions of its basic concepts are possible, which may admit counterexamples. Their elimination is then possible by the introduction of a new definition of the considered object without rejection of the theorems involved. On this level of development a mathematical theory appears to be similar to an empirical theory and possessive of a protective belt, too. However, when its initial concepts are strictly defined this similarity disappears. Appearance of counterexamples at the mature stage of a mathematical theory inevitably entails the collapse of the body of the theory and, thereby, leads to the rejection of some of the basic principles of the theory. Mathematical reasoning that involves "reductio ad absurdum" is based on just this quality of mature mathematical theory.

Consequently, the difference between mathematical and empirical theory in this respect is quite clear. Contrary to an empirical theory that is essentially connected with ad hoc hypotheses at every stage of its development, for a mathematical theory such hypotheses are immaterial. In a mature mathematical theory there cannot be any change, save of the core of the theory, by means of a change in the protective belt.

4. It should be stressed that mathematical theory has a completely different character and logic of justification, than empirical theory. In general, for the applicability of a mathematical theory, or of any science, primarily responsible is its structural adequacy to the described sphere of phenomena. The elaboration and wide application of physical theory counts, at the same time, for justification of the theory. On the contrary, the wide range of application of a mathematical theory establishes only its structural adequacy to some sphere of reality and the presence of some consistent fragments in it, but says nothing about its foundations, the strictness of its proofs or the consistency of its principles. The foundations of a mathematical theory is the subject of special logical and methodological research and is not solved in the process of its application. In physics, there is not and cannot be any special theoretical branch of which the subject would be the establishment of the truth of physical theories. No analogy with empirical knowledge has any place here; such an analogy would be only misleading.

5. Lakatos's empirical conception is essentially based on Gödel's Second Incompleteness Theorem which, according to Lakatos, has delivered a decisive blow to the program of formalistic metamathematics. He notes that a proof of consistency should be complex enough and, consequently, the problem of consistency of the theory in which it is conducted should remain doubtful.

Although Lakatos is right on this point, it is, however, illegitimate to conclude that the consistency proofs now lose their significance and that justification of mathematics is meaningful only in the empirical scheme, as verification of hypotheses in the light of a series of facts. Here it is taken for granted that the logical justification of a mathematical theory is acceptable only when it is complete and final, that is, the possibility of its gradual perfection by restricted means is excluded. However, although we have no means for the complete justification of the consistency of a theory, e.g., of arithmetic, we are able to prove the consistency of separate fragments of it; we can deliberately protect a theory against one or another class of paradoxes, by means of refinement of our logic. Gödel's theorems prove the impossibility of a consistency proof of sufficiently complex theories by means of a finitary metatheory, but this by no means also rejects consistency as an aim of constant improvement of the means of the inner, purely logical, justification of mathematics.

The renunciation of absolute truth as an aim in the empirical sciences is not sufficient grounds for the renunciation of the ideal of consistency in mathematics as well.

6. It seems reasonable to accept that completely different criteria of approval of new theories are peculiar to mathematics in comparison with the empirical sciences. According to Lakatos the grounds for the approval of a new theory lie in its effectiveness, in the possibility of the theory widening the empirical content of science and widening the class of problems which it solves. Such a criterion is, apparently, acceptable, for an empirical theory while it is completely unacceptable for mathematics, at least as an actual principle of preference. Although successful applicability is necessary for the approval of an empirical theory, only the belief in its logical consistency is what is actually required for a mathematical theory's adoption.

Mathematics as a science also tends toward effectiveness and needs success in the field of applications like any other empirical theory. It is from just this point of view that we can understand what the mathematicians actually study in the various epochs. Contrary to the sphere of empirical knowledge, however, the general aim of science is not a criterion of preference for the various theories. Mathematics tends to widen the sphere of its objects not only for reasons of actual effectiveness but, also, so that the system of mathematical knowledge, as a whole, attains a more perfect, inner logical structure. By subjecting mathematical science, along with empirical science, to a unified scheme of growth of knowledge, Lakatos has lost that essential peculiarity of mathematical knowledge, which G. Cantor expressed in the famous quotation: "the essence of mathematics consists in its freedom."

7. The principal difficulty for the Lakatian interpretation of the history of mathematics stems from the concept of "fact" within it. The conception of scientific research programs represents the historical development of a

scientific theory as an interaction between the principles of the theory and some essentially independent "facts" from these principles. But what is a fact in mathematics? Undoubtedly, here we are not concerned with facts in the empirical sense of the word, like phenomena in space and time. Mathematical facts by themselves represent some indisputable connections on the level of idealizations. In other words, mathematical facts are only pure mental constructions. But, in this case, we lose the possibility of any clear demarcation between fact and theory, which quite clearly manifests itself in empirical knowledge and has become the cornerstone for understanding the general scheme of its development. The interaction between theory and fact, on which the whole conception of scientific-research programs is based, becomes meaningless in the case of mathematical knowledge, since we are unable to make this clear-cut demarcation.

We can define the sphere of mathematical facts as a system of truths, which are given to us with apodictic evidence. According to this understanding, we can distinguish the following groups of mathematical facts: the arithmetical truths of the type $2 \times 2 = 4$, elementary assertions about geometrical figures and constructions, effectively verifiable assertions of the form "the equation $x^2 - 1 = 0$ has a root equal to 1." Also included in the class of mathematical facts can be all those mathematical proofs which are finally reducible to such apodictic evidences. Such a definition of mathematical fact seems to make plausible the thesis that mathematics, like any other science, also develops according to the Lakatian scheme of interaction between facts and theory.

However, complete analogy is not actually attained here. In empirical theory, we have to do with interaction between theory and fact in which each of the poles exerts an influence on the other: the facts lead to the appearance of new theories, while new theories lead, finally, to the correction of facts. Both poles are flexible and historically relative. The definition of mathematical facts as apodictic evidences – it seems that there is no other way to define them – compels us to accept that a mathematical theory is based on facts of a peculiar nature, which are not subject to correction, either on the basis of experiment, or by means of theoretical analysis. In other words, we are compelled to accept that mathematics, contrary to experimental knowledge, has, as a certain absolute basis, a system of apodictic evidences, which are not subjected to correction by the theory. This means that the inner logic of the development of mathematics has an essentially different character than that of empirical theories.

8. We cannot ignore a certain degree of cumulativeness in the development of mathematics. A mathematical theorem can be generalised, can be expressed in another language, or can obtain a new interpretation, but once it has been proved it can never be refuted and the sphere of its validity can never be restricted. Every new era of development of mathematics changes its language, widens the system of its concepts, but never rejects the previous advances. This difference stems from the specific place of mathematical science in the system of knowledge, from the role it plays with respect to contentual knowledge as a specific language or sign model. A new mathematical theory from the viewpoint of function appears simply as a wider model which does not eliminate directly the values of the old – its adequacy in application to certain fields of knowledge. As a consequence, if not everything, almost everything which was included in a previous strict system of deductions is preserved in the new theory. In his attempt to reject the prejudices of cumulativeness in mathematics, Lakatos ignores that they have objective grounds which are connected with the very nature of this science.

Attempts to prove the opposite thesis do not seem conclusive. The examples of refutation of theorems which Lakatos suggests, concern the stage of their formation, when still connected with ill-defined concepts. In this case, the use of implicit assumptions and the appearance of counterexamples are possible. However, both the history of mathematics and the everyday practice of the mathematician show that this stage of proof has been overcome very quickly and a stabilization has been attained where it is impossible to challenge the correctness of any result. Of course, in the periphery of mathematics, like in any other science, doubtful points may occur in assertions or modes of reasoning. However, contrary to the empirical sciences, mathematics, possess an absolutely stable centre, an indestructible core, that can only expand in virtue of the establishment of new truths, but never can be rejected by some counterexample or by logical analysis.

9. The aforementioned shows that the conception of scientific research programs does not adequately reflect the historical development of mathematics or the specific features of this development. Of course, there are common features in the development of mathematics and empirical knowledge. Both spheres of knowledge develop from particular theories to more general ones, both strive to attain greater degrees of effectiveness,

both tend to widen the class of problems which they solve to an internal harmony, that is, to the consistency of their principles. The interaction between pure and applied mathematics is similar, in many respects, to the interaction between the theoretical and applied parts of the empirical sciences. But such similarities do not reflect the specifics of mathematical knowledge and cannot provide sufficient grounds to understand its real history.

10. The empirical philosophy of mathematics, which attempts to understand the methodology and history of mathematics by proceeding from the idea of the unity of mathematics and the empirical sciences, should be abandoned. However strange it may seem, today we ought to turn our minds to the ideas of an apriori philosophy of mathematics once again, but in an up-to-date and improved form. This is necessary to understand the nature of logical norms, the specifics of mathematical evidence, the stability of mathematical proofs, and the specifics of cumulativeness in the history of mathematics. This turn is inevitable and has been prepared, partly, by the works of Lakatos himself and his followers – in particular, by those difficulties that appear in the attempts of justification of their ideas, appealing to the history of mathematics.

The contours of this philosophy can be outlined proceeding from the analysis of the concepts of mathematical fact and apodictic evidence. These two concepts are, in my view, the most important not only for the modern philosophy of mathematics in general but, also, for the understanding of the history of mathematics. It stands to reason that we ought not to give up the ideas of the formalistic philosophy of mathematics that also stress the specifics of mathematical knowledge. The weak point of Lakatos's conception is that, motivated by the idea of the methodological unity of mathematics and empirical knowledge, these specifics are concealed.

11. What can all these abstract discussions on the logic of the development of mathematical knowledge do for the history of mathematics? In my view, the philosophy of mathematics can provide a general orientation, sometimes exclusively important, for concrete historical analysis.

In examining the facts of pre-Greek mathematics, for example, we usually seek some empirical grounds responsible for the formation of the first mathematical concepts, while very often it is asserted that pre-Greek mathematics had no proofs and developed as a purely empirical science. In my view, this is just a manifestation of the empirical philosophical conception of mathematics. In fact, primary mathematical concepts have never been abstracted from experience and mathematics has never been developed according to the same laws as the empirical sciences.

Primary mathematical concepts are intersubjective intuitions of a specific kind which are formed by man's activity and reflect the structure of this activity. These intuitions are not merely inductive generalizations or conventions. The possibilities for abstraction from real restrictions underlying the mathematical idealizations are themselves incomparably wider than those admitted in the formation of concepts in the empirical sciences.

The appearance of the primary concepts is, at the same time, and appearance of pure deduction, because the mathematical idealizations have appeared not in isolation, but as a system, in mutual relationship. The concepts of "point", "straight line", and "plane" were not separated in their genesis. We have grounds to suppose that deduction, as a necessary transition from some notions to others, also existed in pre-Greek mathematics, but only as a simple mental process without linguistic articulation of the consequence of the separate steps. The Greeks did not introduce deduction into mathematics; deduction as such is included in the very essence of mathematical idealizations. The Greeks only shaped deduction linguistically and converted it into a systematic method.

Summing up, the Lakatian conception of the history of mathematical knowledge based, in fact, on the analogy between mathematics and the empirical sciences in an attempt to bring together mathematics and contentual knowledge, dispersing the myth of exclusiveness and the inerrable, complete, and cumulative character of mathematical knowledge – is far from being faultless. Lakatos swings, in fact, from one extreme to another: He comes to the dissolution of mathematical knowledge into the contentual one and the complete ignorance of the formal aspects of mathematical knowledge.

Consequently, it allows many aspects of the genesis and succession of mathematical theories to slip. It picks out in mathematics only a very abstract moment of its development, namely, the simple fact that mathematics, like the other sciences, develops through the solution of certain problems. However, the specifics of mathematics are not reducible to the specific character of its counterfacts, to which Lakatos turns his attention, but affect the very logic of succession, development, and justification of mathematical knowledge. These specifics stem from the place of mathematics in the system of human knowledge and the peculiarities of its functions.

Only a more profound treatment of the philosophy of mathematics will open the right perspectives for the study of pre-Greek mathematics. The same, of course, also concerns the other stages of development of mathematical knowledge. Their historical analysis is greatly mediated and determined by philosophy.

*University of Moscow*

MICHAEL FORTUN AND SYLVAN S. SCHWEBER

# SCIENTISTS AND THE STATE: THE LEGACY OF WORLD WAR II

## 1. INTRODUCTION

A convincing case can be made that the principal events shaping the twentieth century until the nineteen nineties have been wars: World War I, World War II, Korea, Vietnam. The two world wars changed the practice of science. Among many other things, both wars highlighted the value to the state of scientists and scientific institutions. But in contrast to the first world war, the second altered the character of science in a fundamental and irreversible way.[1] The importance and magnitude of the contribution to the war effort of engineers and scientists, particularly physicists, changed the relationship between scientists and the state. Already during the war, and with ever greater emphasis after the war with the onset of the Cold War, the armed forces in the United States, particularly the Navy and the Army Air Force, realizing that the future security of the nation and its dominance as a world power depended on the creativity of its scientific communities and the strength of its institutions of higher education, invested heavily in their support and expansion. From the mid-forties to the mid-fifties a close relationship was cemented between scientists and the military. Physicists played a key role in these developments and our paper was an outgrowth of an inquiry into the special skills and characteristics that made their contributions so central until the early sixties.

The obvious traits the physicists possessed were unusual versatility, outstanding mathematical abilities, and remarkable experimental and technical skills. These attributes enabled the best among them to invent quantum mechanics, construct cyclotrons, develop atomic beams apparatus, design rhubatrons and klystrons – and thus acquire mastery over the atomic and nuclear domains. Their particular achievement in the period from 1925 to 1939 was the apperception of the (quasi) stable ontology – electrons, neutrons, protons – the building blocks of nuclei, atoms and molecules – and the formulation of the (quasi) stable theory – quantum mechanics – that described the interactions and dynamics of these entities. Their unique powers were their ability to translate their understanding of microscopic phenomena, i.e. of nuclear, atomic and



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science,* 327-354.

molecular structure and dynamics, into explanations and predictions of the macroscopic properties of matter, and moreover to translate that mastery into the design of macroscopic devices.

The generation of physicists that matured with the birth of quantum mechanics was a "population" in possession of traits that allowed it not only to adapt to existing conditions but in fact to dominate and transform the scientific and technological environment. The depression exerted a rigid selection pressure on the community: only the very best survived and these were a remarkable lot. World War II gave them the opportunity to display their powers and in the process they acquired some measure of power. The physicists working on radar devices, on proximity fuses, on missiles, and the atomic bombs had unlimited funds and equipment for experimentation, so that the technical knowledge and skills of the community were greatly increased during the war. Being a hardy, self-confident, wide-ranging and resourceful new species physicists occupied new niches after the war.

World War II is an example of a Hacking type of scientific revolution.[2] It will be recalled that two of the characteristics of a Kuhnian revolutions are that:

(i) the transformation consists essentially of changes *within* a discipline and

(ii) the transformation is discontinuous in that the pieces come together in a new way and is primarily an *epistemological* revolution.[3]

The genesis of quantum mechanics culminating with the 1925–27 Heisenberg–Dirac–Schrodinger–Born–Bohr formulation of the theory is an illustration of such a revolution.

In contrast, Hacking type revolutions are "big" revolutions: "they are embedded in, pervade, and transform a wide range of cultural practices and institutions":

(i) They are *inter or cross* disciplinary. In a Hacking type revolution something happens in more than one discipline; a wide range of practices and institutions are transformed.

(ii) New institutions are formed that epitomize the new directions.

(iii) They are linked with substantial social change. After a Hacking type revolution there is a different feel to the world, there is a marked change in the texture of the world.

(iv) There can be no complete, definitive history of a Hacking type revolution.

That the first scientific revolution satisfies all these criteria is clear. The

*Royal Society* and the *Académie des Sciences* are some of the new institutions it created. Hacking has pointed to the rise of the bourgeoisie as indicative of the substantial social change associated with that revolution. The end of the nineteenth century witnessed another Hacking-type revolution connected with the microscopic representation of matter and the establishment of science-based technologies, particularly those grounded on the new understanding of electricity. The Physikalische-Technische-Reichsanstalt, and its replicas – the National Physical Laboratory in Teddington, England and the Bureau of Standards in Washington D.C. are some of the new institutions spawned by that revolution. The new emphasis on graduate education, the post-doctoral research positions in chemistry, Felix Klein's transformation of Göttingen, and his *Göttinger Vereinigung zur Förderung der andgewandten Mathematik und Physik*, are all indicative of the new partnership between universities, industry and the state.

World War II was responsible for yet another Hacking type revolution. The revolution was brought about by the plethora of novel devices and instruments that were developed principally by physicists: oscilloscopes; microwaves generators and receivers; rockets; computers; the myriad of new vacuum tubes and circuits; nuclear reactors;[4] the many new particle detectors. Many of these devices had been introduced before the war but in a relatively primitive state and on an individual and piecemeal basis. It is the scale on which these devices and instruments become available that refurbishes the stage. Thus the experimental practice of physics was transformed by the ready availability of oscilloscopes, magnetrons and klystrons, low noise amplifiers, pulse analyzers, nuclear piles, etc,; chemistry by NMR and by microwave spectroscopy, instrumentation that the developments of radar had made possible. Similarly, radio telescopes transformed astronomy; rockets, meteorology; and similar transformations occurred in geology, oceanography and in other branches of the physical sciences. That there is new feel to doing science is indicated by the transformation of the laboratory effected by the new instrumentation and their ready availability commercially; by the new sources of support for research and the new laboratories that are established on many campuses in the United States (e.g., the Laboratories of Nuclear Studies at Cornell, Chicago, M.I.T. and elsewhere; the Microwave Laboratory at Stanford; the Laboratory for Electronics at M.I.T., ...); by the new relation of physicists to applied scientists and engineers and that of chemists to physicists.

World War II also initiated a revolution in management science, risk assessment and military planning – and military planning has dominated much of world affairs since World War II. Operations research and its methods were in part responsible for instigating these changes. World War II also helped consolidate an engineering approach that became known as systems engineering. Systems engineering had its origins in developments within AT&T, the communications monopoly that until the mid-eighties oversaw the Bell system, the 24 operating companies and the Long Lines department that delivered telephone service to the United States. More particularly, it was a by-product of the partnership between the Bell Laboratories, the research and development component of AT&T, and the Western Electric Company, its manufacturing and supply arm. What characterized the Bell system was "technical integration," the concept introduced by Hendrik Bode, to describe "the ease with which technical information could flow back and forth between Western Electric, the operating Companies and the Bell Labs."[5]

These developments after WW II were deeply influenced by the research and development activities carried out during the hostilities, activities in which physicists played a dominant role. The pattern of interaction that were developed by the OR groups between scientists and the military in the United States became the model for the post-war committees that advised the armed forces on weapons development and on tactical and strategic matters. In the US, World War II was responsible for a new framework for expert advice to the armed services by civilian scientists – principally physicists or former physicists – as indicated by the proliferation of advisory committees. The structure and mission of these committees were patterned after the NACA, the committee that had been set up before the war to advise the government on military and civilian aviation, the NCRD – the *civilian* agency that Bush headed during the war that was responsible for weapons development – and the many OR groups that had advised military commands during the war. This quantitative increase translated itself into something dramatically new in the relation between science and the military. Similarly, the consolidation of the systems approach and of system engineering owes much to the activities of physicists at the Radiation Laboratory at M.I.T., the Applied Physics at Johns Hopkins, and the Jet Propulsion Lab at Cal Tech, – facilities that could rightly be described as systems facilities – and to their post war activities in such summer projects as Lexington, Hartwell, Charles and Troy and Vista.[6] These summer projects were depicted as systems studies because they explicitly dealt with "the called upon to cooperate in forming a system."[7] They were also representative of projects wherein *collective* creativity – rather than individual creativity – determined the success of the enterprise.[8]

Although it is clear why the contribution of the physicists were central in the successes of radar, sonar, proximity fuses and the atomic bombs it is not so clear what they contributed to operations research (OR) and systems engineering and which of their skills were particularly relevant. Our paper is an attempt to answer that question for OR. In the process we briefly look at the history of OR and of Taylorism, and compare the developments in the United States and Great Britain. We also discuss the relation of OR to systems engineering. Our paper can be considered as a study in how authority is appropriated and how the differing contexts in the US and Great Britain shaped the differing outcome in these two countries.

## 2. OPERATIONS RESEARCH[9]

The principles of operations research were first formulated in Great Britain. The original conception of teams of scientists working at the operational level in military commands is attributed to the physicists E. C. Williams and P. M. S. Blackett. The appellation "Operational Research" derived from the fact that their main field of activity was "the analysis of actual operations using as data the material to be found in an operations room, e.g. all signals, track charts, combat reports, meteorological information, etc. ...[T]hese data are not, and on secrecy grounds cannot, in general, be made available to the (service) technical establishments. Thus such scientific analysis, if done at all, must be done in or near operations rooms."[10] "Blackett's circus," as the original group was called, included three physiologists, two mathematical physicists, one astrophysicist, one Army officer, one surveyor, one general physicist, and two mathematicians. This mixed-team aspect of group – in fact, both the team element and the interdisciplinary nature of the effort – were to become distinguishing characteristics of operations research and systems engineering.[11]

During the early phase of WW II the problems addressed by operational methods were concerned with the allocation of scarce radar stations, search techniques and convoy strategies in anti-submarine warfare. OR rapidly expanded to the study of all aspects of the functioning of complex organizations and operations – including personnel and machines – encountered in the different branches of the armed forces. Eventually it became an integral element in the planning of the major campaigns – including strategy and logistics, the training and management of manpower, the cost effectiveness of weapons, and the allocation of resources. The specific primary purpose of operations research during WW II was to discover means for making the best use of the military forces and weapons then currently available. OR studied the operation or the process while it was going on and usually its recommendations could be implemented very quickly.[12] One of its essential features was the comparative freedom of the analysts to seek out their own problems, and the direct coupling of the analysis to the possibilities of executive action.

By the end of the war several hundred people were involved in Great Britain in analyzing both tactical and strategic problems. The Admiralty, the Chiefs of Staff, the RAF Coastal Command, the SE Asia Command, the Fleet Air Arm, Bomber Command all had operational units attached to them.[13]

The first U.S. OR groups were formed in 1942. One was attached to the Naval Ordnance Laboratory and was under the supervision of Ellis Johnson, who was then on leave from Carnegie Tech. The NOL group worked on the problem of mine warfare. Their analysis led to the enormous aircraft mining blockade of the Inland Sea of Japan in 1945. Another, that later became known as the Anti-Submarine Warfare Operations Research Group, worked at the Rad Lab at M.I.T. It began under the stimulus of Phillip Morse, and reported initially to both the Army and the Navy. This group – which included Morse, George Kimball, Francis Bitter, and John von Neumann as a frequent consultant – was later expanded into the Operations Research Group that was attached to the staff of the Commander-in-Chief, U.S. Navy, Admiral King; it dealt with problems connected with submarine warfare, aircraft and amphibious operations, anti-aircraft and new weapons analysis.[14] From September 1942 to July 1943 Von Neumann was closely connected with the Morse Navy Group – the Mine Warfare Section of the Navy Bureau of Ordnance – and worked on the physical and statistical aspect of mine warfare and countermeasures to it; these activities took him to England during the first half of 1943.

Von Neumann's involvement was particularly influential in that it brought "game theory" into operations research. In 1928 he had proved his famous minimax theorem for (finite) zero-sum games that states that for such games there exists an optimal strategy that is given in terms of a unique number that represents the minimum gain and maximum loss each player can expect. Both in 1932 and again in 1937 Von Neumann had lectured in Princeton on mathematical economics, and when Oskar Morgenstern came there in 1938 he collaborated with him on a monograph on the subject. The manuscript of their influential *The Theory of Games and Economic Behavior* was completed in January 1943, just prior to Von Neumann's departure for England. By the end of the war the new game theoretic methods that had been developed by Von Neumann and Morgenstern were added to the toolkit and mathematical techniques that OR scientists deployed.[15] These proved very valuable, and game theoretic approaches took on great importance after the war, particularly after the introduction of fast digital computers.

As in Great Britain, each of the services set up during the war a functioning operations research group at various command level. As in Great Britain, the earliest U.S. efforts had short term aims. But it was soon established that OR techniques could usefully be applied throughout the whole military effort. A particularly intensive effort went into analyzing the tactical problems of the air war against Germany.

The importance and influence of OR activities is indicated by the fact that operations research organizations were maintained in all the branches of the armed forces after the war. The Navy (through a 1947 ONR contract) supported an Operations Evaluation Group (OEG) at MIT, with Jacinto Steinhardt as its director. The Air Force had an Operation Analysis Group (OAG) both at the Office of the Deputy Chief of Staff/Operations and at the various Air Force Commands. In 1946 it contracted with Douglas Aircraft to manage Project RAND (*R*esearch *an*d *D*evelopment) to deal with long term assessments. In 1948 the RAND Corporation was established, initially financed by the Ford Foundation to demonstrate its supposed independence. After the passage of the National Security Act of 1947 that created the National Military Establishment, the Weapons Systems Evaluation Group (WSEG) was established in December 1948 to advise the Joint Chiefs of Staff, with P. M. Morse as its technical director. In 1949 the Operations Research Office (ORO) was created by the Army, and staffed by personnel hired under a contract with the Johns Hopkins University. In each case the OR group reported their analysis and made their recommendations directly to the commanding officer of the branch to which the group was attached. For it had been recognized during the war that the successes of OR efforts were in part due to the fact that the head of the OR team reported directly to an operational military commander who could implement the recommendations.[16]

During World War II the time scale of change in the enemy's environment – the theater of operation, the targets, the weapons and technologies used – was relatively long. It was therefore usually possible to respond fairly effectively to the introduction of new weapons – such as U-boats with snorkels, V-2 rockets – and the new strategies they made possible, by changing one factor in the opposing weaponry and strategy. Hence OR's incremental, additive approach was, for the most part, adequate to meet the challenges.

But during the war it also had been necessary to address such questions as the optimal allocation of limited resources among a variety of competing military demands. Thus the M.I.T. Radiation Laboratory had to decide what kinds of radar sets to develop, how many to manufacture, where to deploy them, and to supervise the manufacture and insure the adaptation of the apparatus to the field conditions. From these efforts developed what became known as "systems analysis", the planning of responses to military an security needs based on *projected* weapons systems and the environment in which they would operate.

Traditional operations research for the most part addressed problems where the objectives were precisely spelled out, and the existing systems and weapons, i.e. the hardware, were considered *fixed and unchangeable*. OR was usually concerned with tactical problems and could be stated quantitatively and mathematically, and the aim of the analysis was "to find more efficient ways to operate, in situations where the meaning of 'more efficient' is fairly clear."[17] OR problems usually admitted a unique solution that represented the optimal allocation of the hardware available to the decision maker.

> Systems analysis, on the other hand, refers to the far more complex problem of choice among alternative future systems, where the degrees of freedom and the uncertainties are large, where the difficulty lies as much in deciding what ought to be done as in how to do it...The total analysis is thus a complex and untidy procedure, often with little emphasis on mathematical models, with no possibility of quantitative optimization over the whole problem, and with necessary great dependence on considered judgments.[18]

Because in the problems that are considered in systems analysis the objectives and criteria cannot be precisely defined, there are usually no unique or sharp optimal or maximal solutions. Nonetheless, by insisting on precision and logical rigor, the systems analysis approach [in the military area] at times evidently could achieve more reliable assessments of the possible choices, their costs and consequences than other means. But very often what was meant by "systems analysis" was the attempt to synthesize the various specialized research techniques into a coherent analytical framework that could be used to confront the problem as a whole. But the systems approach had one feature that became its hallmark: Heterogeneous professionals – physicists, chemists, electrical engineers, generals and admirals, economists, managers – and heterogeneous organizations – universities, manufacturing firms, branches of the armed forces, government laboratories, private foundations, think tanks – were seen as and operated as essential interacting components in a "system." Indeed, "disciplines, persons, and organization [took] on one another's function as if they [were] part of a seamless web."[19]

Already during the war, many members of the various operations research organizations became convinced that the techniques they had developed were not limited to military applications. But how to translate the wartime OR and decision making skills and insights into educational programs was far from clear. Two somewhat separate approaches can be discerned in the immediate post war period. The first – directed primarily at undergraduates – was designed to train students as "scientific generalists" and made statistics the core of the curriculum. The second approach – which became identified with OR programs – was aimed at the graduate level, and hoped to draw students who had been trained in one of the sciences as undergraduates. When writing about OR's accomplishment after the war the definitions of OR that were given usually emphasized its general features and its applicability in non-military contexts. Thus the physicist Charles Kittel characterized OR as "a scientific method for providing executive departments with a quantitative basis for decisions. Its object is, by analysis of past operations, to find means of improving the execution of future operations."[20] One of the classic definitions of OR was given a few years later in Morse and Kimball's *Methods of Operation Research*:

> Operations research is a scientific method of providing executive departments with a quantitative basis for decisions regarding the operations under their control.[21]

This emphasis on "scientific method" reflected the then prevalent philosophical outlook of the physics community in the United States, namely, operationalism and logical positivism with their emphasis on methodology and methodological issues.

By 1952 "Operations Research" was considered to be synonymous with the terms "Operations Analysis", "Operations Evaluation", "Weapons Systems Evaluation." All were considered to employ scientific methods to solve action problems. In these definitions, the scientific method was understood as a systematic method of learning by experience, or more explicitly, "as that combination of observation, experiment and reasoning (both deductive and inductive) which scientists are in the habit of using in their scientific investigations ... [with the] reasoning often [taking] one of the forms of numerical deduction and induction covered by the term statistical analysis."[22]

By the mid-fifties OR had established itself as a new discipline. Graduate programs in OR were operating successfully at several major American and British Universities.[23] In the United States the process had been accelerated by efforts of the National Research Council (NRC). The NRC, "recognizing the great value of operations research, and wishing to further its development and applications outside the armed forces," had in 1949 set up a committee within its Applied Mathematics Division for this purpose. The NRC Committee surveyed existing manpower, formulated guidelines for graduate training program, established several doctoral fellowships, and funded several conferences. The Committee also issued a pamphlet designed to provide executives and administrators and scientists with a brief description of OR. In its concluding section the pamphlet asked the question: "What is new about OR?" Its reply was that the novelty of OR does not lie in the use of certain specialized scientific techniques, but "in the broad scientific approach to the problems of the organization *as a whole*."[24] However, a more thorough historical inquiry indicates that the answer given was open to question. In fact, the OR rhetoric of efficiency and effectiveness, and of system analysis, was the same as the one used by the advocates of Taylorism and scientific management during the first decades of the century. Similarly, both depended on measurements and quantitative studies and both made use of the mathematical language of optimalization. Both were concerned with situations in which alternative courses of action existed. The objective of both scientific management and OR is to clarify the differences between the several courses, indicate their outcome and relate these to the stated objectives of the operations.

Taylorism had promised "an escape from zero-sum conflict" – in which the gain of one party could only be accomplished by the loss of the other. Taylorism in the post World War I period was a vehicle "to validate a new image of class relationships" and as an escape from having to accept the class confrontation and social division that had existed in the prewar period.[25] Taylor's system of scientific management and Ford's assembly line encapsulated the american commitment to technological efficiency and productivity. Taylorism effectively provided both a methodology and a legitimation for linking the rationalization of technology to the rationalization of labor. In the name of efficiency, control of the labor and machinery processes were given over to allegedly "neutral" engineers. Ford's assembly line combined machine tools, time-and-motion studies, electrical power into a rationalized system of efficient and profitable production under engineering control. Both Taylorism and Fordism were enthusiastically embraced in the United States and "pervaded the entire culture." The reception of Taylorism and Fordism in Europe, as evidenced by public discussion and government sponsorship there, was equally enthusiastic but somewhat more selective. The greater sensitivity to bureaucratic control was undoubtedly a factor. Max Weber noted that scientific management, which he defined as "the exercise of control on the basis of knowledge," was the essence of modern bureaucracy.[26]

Henry Le Chatelier, the distinguished physical chemist and metallurgist, did much to promote Taylorism in Europe. He was the first to introduce and expound Taylor's principle in France. As early as 1908 he translated several of Taylor's articles, and thereafter proselytized Taylorism in his extensive writings on the subject. These were collected and issued with a lengthy prefatory assay in an influential volume in 1927.[27] Le Chatelier saw Taylorism as rigidly *deterministic*. By conceiving of management as a *deterministic* process and by employing extensively experimental measurements, Taylor had done for industrial phenomena what Claude Bernard, one of the pioneers of experimental medicine, had done for natural phenomena in general. For Le Chatelier Taylorism consisted primarily in the quantification and the subsequent mathematization of the management of industrial organization and production. Blackett's work in World War II fleshed out and gave an explicit formulation of Le Chatelier's mathematical vision of Taylorism.

An analysis of Taylorism and scientific management and the models that were set up to treat the problems encountered during World War II[28] suggest the following response to the question: "What was new in OR?" A partial answer is that Taylorism with its deterministic conception of the world used deterministic models – where the effect of a given action was assumed to result in a well-defined, determined effect. OR on the other hand dealt principally with stochastic processes and with probabilistic models that explicitly recognize uncertainty as an intrinsic feature of the processes being modelled.

In fact, we would like to suggest that it is the use of probabilitistic models, later to be refined with a host of mathematical techniques – decision theory, game theory, Monte-Carlo methods[29] that is the characteristic feature differentiating OR from "scientific management." Scientific management investigations – like the later OR ones – were interdisciplinary in character. There was nothing novel about the "mixed-team" approach of OR. Industrial psychology grew out of the "scientific management" investigations of the effect of repetitive actions on workers on the assembly line. Nor was it unusual to have physiologists on the team to help in the investigation of fatigue in the workplace. The wholism of Kurt Lewin, one of early pioneer in the field of industrial psychology had its roots in the "systems" approach of Taylorism.

The connection between OR and Taylorism having been established and the differences noted, we are now in a position to indicate why the physicists played such an important role in making OR so highly effective.

## 3. WHY THE PHYSICISTS?

Why were physicists, and more particularly, why were the British and American physicists key contributors to OR? A superficial answer might suggest that physicists trained or resonating with the Anglo-Saxon physics tradition of "empiricism + computability" were ideally suited to carry out these developments. Their pragmatic, instrumental tradition was particularly helpful in shaping the theoretical tools of OR. Bronowski in his review of Kimball and Morse in 1951 hinted at this: "Operations research has done its major work, and it turns out to have been a piece of education – the education of scientists and warriors in a new empiricism."[30] Moreover, in the British and American context, perhaps more so than anywhere else physicists came to embody the notion that:

rationality = computability

Crowther and Whiddington also alluded to this equation in 1947 in their book *Science at War*. In their discussion of OR they noted:

> This is the major conception – the reduction of war to a rational process. It is the contrary of that held by Hitler, who had a romantic view of war. He believed that wars are won by

great inspiration... Systematic scientific work on known weapons paid larger and quicker dividends. It beat Hitler.

Hitler and his generals failed to produce any operational research comparable to the British development. If they had, they probably would have won the submarine campaign and the war. But it was impossible for them to collaborate on the basis of equality with the rational, egalitarian scientists...[31]

Crowther and Whiddington's observation that the hierarchical structure of Nazi Germany prevented the collaboration between the German armed forces and the "rational, egalitarian scientists" points the way to a richer and more sensitive explanation that takes the political context into account; an explanation, incidentally, that also sheds light on Taylorism.

Given that the initial developments of OR stemmed from problems connected with radar – the instrumentation for which was initially designed by physicists, the experiments for which were initially carried out by physicists, and the mathematics and theory for which were initially formulated by physicists – it is not surprising that physicists would be involved in the problems connected with the allocation of the radar sets and their most effective deployment. If we accept that probabilistic modelling was the novelty in OR then the role of the physicists in that enterprise becomes clarified further. Quantum mechanics had familiarized them with probability, and more so than most other scientists they conceptualized their universe of inquiry probabilistically. The models of their world were intrinsically probabilistic.

It is worth noting the parallelism between physicists and engineers in OR and Taylorism. It was an engineer who brought forth Taylorism and scientific management and engineers were to the midwives for the new utopian world these techniques promised. As Herbert Hoover commented in 1922:

[The engineer's] lifelong training in quantitative thought, his intimate experience with industrial life, leading to an objective and detached point of view, his strategic position as a party of the third part with reference to many of the conflicting economic groups, and above all his practical emphasis on construction and production, place upon him the duty to make his point of view effective.[32]

Physicists played much the same role with respect to OR in the post World War II world. Both sets of professionals emphasized their objectivity, their neutrality, and their use of scientific methodology. Blackett saw the relative novelty of OR "not so much in the materials to which the scientific method is applied as in the level at which the work is done, in the comparative freedom of the investigators to seek out their own problems, and in the direct relation of the work to the possibility of executive action." It was important in order to maintain objectivity and neutrality that the operations research scientists not make the decisions – they were doing "science in collaboration with and on behalf of executives."[33]

If for the most part the answer to the question: "Why were the physicists so influential in the development of OR during the war?" is that many of them had a mastery of certain mathematical tools, experience with modelling natural processes using those tools, and that proficiency in the use of these tools was much more prevalent among them than in any other discipline – then to a large measure their centrality is accidental. What physicists did have is self-confidence and this may account for the "anything goes" characteristics of OR. Any approach is legitimate in an operations research study as long as it leads to a better understanding of the problem.[34] If this argument is right then population geneticists, economists, physical chemists in possession of these same tools could and would have done have the same thing. And indeed they did.

In fact, with the evolution of operations research into systems analysis there occurred a corresponding change in the professionals that were in charge of the analysis. Whereas physicists and applied mathematicians were predominantly responsible for the creation and establishment of OR, economists, lawyers and social scientists took over the influential positions in systems analysis. Effective systems analysis must integrate a wide variety of professional skills, including those of the social and political scientists. Economists and lawyers could act as "generalists and integrators"[35] because on the one hand the more technical aspects of subproblems could be relegated to engineers and applied scientists and on the other, the more integrative aspects of the problem – such as the allocation of resources among expensive competitive weapons systems, assessing the cost of different modes of research and development, bargaining among competitors, evaluating the political consequences of certain recommendations – were the traditional areas in which the toolkits and the special skills of economists and lawyers could be effectively employed.

The intellectual tools possessed the physicists may have been necessary conditions for the development of OR. But physics and physicists do not explain why OR was predominantly an anglo-saxon phenomenon. Ted Porter in a recent talk on "Quantification and the Accounting Ideal in Science"[36] has insightfully given the clue for the answer. He pointed out that quantitative knowledge "is especially useful to coordinate the activities of diverse actors, and to lend credibility to forms of belief and action when personal trust is in short supply." Porter stated the matter succinctly and epigrammatically: "Objectivity has prestige in a pluralistic society mainly as a substitute for trust." And the logic of objectivity is most readily expressed by the language of numbers, quantification and mathematics.

We can thus view Taylorism and scientific management, with their emphasis on quantification, as a technology for dealing with mistrust in liberal democratic countries. How were fair wages to be determined? Industry might simply have negotiated a price. But by the end of the nineteenth century an agreement resting on no more authority than managerial judgment lacked credibility. The numbers gleaned from time and motion studies, time sheets, etc., were to be the objective quantitative evidence on which both workers and management could agree. They were to be the basis for setting fair wages for the workers and reasonable returns for the company.

The same for OR. What is at issue is how power is exercised in situations where military commanders and scientists must work together in a political system in which ultimate responsibility lies with civilian leaders elected by the citizenry. In the wartime setting the decisions could simply have been made unilaterally by the military. But such decisions would lack credibility in the democratic and presumed egalitarian setting of the paradigms of twentieth century liberal societies, Great Britain and the United States. So OR was mobilized to provide quantitative evidence on the basis of which decisions were to be made.

Miller and Rose have noted that accounting is wholly typical of the normal way power is exercised in modern liberal societies.[37] Power there "is not mainly an agency of direct, dictatorial control, but a set of methods and implicit standards by which people are judged and by which they judge themselves. Subordinates are left with some autonomy, provided they maintain good numbers. Superiors are not supposed to intervene in the details, nor to impose forcefully their own discretion, but so far as possible to respect the objective measure."[38]

The comparison of OR with accounting is striking. Much of what Miller and Rose say about accounting applies directly to the authority and prestige of OR in the wartime situation, to OR and British Labour party politics in the post World War II period, to OR within american corporations during the fifties and sixties, and for that matter more generally to the authority and prestige of the knowledge and methods of the natural sciences in the political realm.

But there remains an additional question. Physicists did play a dominant role in giving scientific and technical advice to the U.S. government at the highest level well into the sixties. The composition of the advisory committees to the various branches of the Armed Forces, that of PSAC, of the GAC of the AEC, all give proof of that fact.[39] Why? The obvious answer is of course their technical knowledge, particularly in matters nuclear. But there are additional factors to be considered. The first to be noted is that one of the legacies of OR may well have been the consolidation of the advisory mechanism that had been developed during the war whereby civilians could interact with the military. A comparison between the United States and Great Britain helps shed light on the matter. In Great Britain there existed a cadre of high level scientific personnel of the first rank *within the government* – people like Henry Tizard, Edward Appleton, John Cockcroft – whose advice could be sought on *technical* matters *without* recourse to people outside the government.[40] This implied a fairly rigid separation between "insiders" and "outsiders."[41] In the U.S. no such cadre existed. The heads of Los Alamos, Argonne, and Oak Ridge did not have the same standing as Cockcroft – and for that matter these were not government laboratories. They were run for the government by the University of California, the University of Chicago, and by Union Carbide, respectively. The wartime experience gained with NACA, OSRD, the civilian-military OR teams, helped consolidate an advisory system that permitted great permeability, flexibility and efficiency, for which physicists in the U.S. were particularly well-suited.

It should also be noted that the flexibility and efficiency that were the result of this collaboration between physicists, economists, and social scientists, for the most part came to be organized around, and for the sake of, state-sanctioned violence or the threat of violence in the name of "national security." The OR-legacy of World War II was transformed – by men such as Robert McNamara, the assorted RAND analysts he brought to the Pentagon, and others counted among the "best and brightest" toiling for such organizations as the Jason division of the Institute for Defense Analysis – into the paradigmatic method for planning and waging the war in Vietnam.[42] The rationality of systems thinking became the rationalization for an entire war system, as epitomized in the 1967 *Report from Iron Mountain*.[43]

Another point to be noted is that to the extent that physicists's influence can "labelled" it derived much more from their skills as "systems analysts" than as "operational researchers." This is certainly the impression gleaned from a comparison between the post war developments in England and the U.S. OR did not fare well in England after the war, neither in the civilian sector nor in the military sector. In fact, OR reestablished a foothold in Europe only in the late fifties and then by virtue of NATO sponsored activities.[44] Similarly, for all of its successes in the US, OR had its limitations. In December 1950, Philip Morse, who spearheaded the efforts to have OR recognized as an academic and professional discipline of great physicists in an editorial in *Physics Today*[45], asking them "Must we always be gadgeteers?" He indicated that they could contribute to the resolution of the cold was crisis not only by being "gadgeteers" at Los Alamos or the Naval Ordnance Laboratory, but also by helping to make decisions working in any of the operations research groups attached to the armed services. But the tone of the article indicated that as valuable as their contribution might be and no matter what access to the top of the chain of command they might have – they would be "on tap, not on top." The systems analysis and engineering approach and the institutional structures to give advice to the government using systems methods proved more challenging and rewarding to physicists.

The particular skills of the physicists are best observed in connection with their activities in projects such as Hartwell, Charles, Troy and Vista which became known as "systems studies".[46] It was precisely their ability to master the interplay of factors resulting from operating requirements in the field, from research projects in the laboratory, from new apparatus developed from research projects, from engineering and manufacturing constraints that physicists could see how to coordinate and steer the developments into a functioning operating system. Their accumulated *collective* experience – and the ability of the project director to tap into that collective expertise by inviting the various experts to participate in the project – paid rich dividends.[47] These summer studies made clear that systems analysis goes far beyond operational research "in comparing the requirements of the field with the performance possibilities of new developments and coordinating the two."[48] In doing these things it also became apparent that their influence on management in industry and on the military command was much more authoritative than is the case with OR.[49] They were more on top and less on tap.

Finally, it should be noted that eventually physicists suffered the canonical fate: they were coopted by the system via the surest route devised – helping to construct the very system by which they would be coopted.

## APPENDIX A

It is of some interest to trace the genealogy of the institutions and of some the men involved in the mediation. Tables 1, 2, and 3 sketch part of that story. It is worth pointing out that Bush, Conant and K. Compton – the architects of NCRD and OSRD – all served in World War I. Conant as a young Army officer in the chemical corps was in charge of a highly secret plant for the manufacture of mustard gas near Knoxville, Tennessee. People working there were not allowed to leave the premises and their mail was censored. During WW II Conant was second in command at OSRD and had responsibility for the nuclear weapons program. Los Alamos has many similarities to the Knoxville plant. Conant and Rabi were two of the people who most influenced Oppenheimer in how and where Los Alamos was run.

The particular structure of NCRD and OSRD and their contractual arrangements owes much to the developments at MIT during the interwar period. The contract system developed by MIT in their Division of Industrial Cooperation and Research (DICR) to help industry with their technical problems became the model for NCRD/OSRD-university and NCRD/OSRD-industry contracts. Bush was at MIT starting in the early twenties and headed the electrical engineering department there; he became second in command after Compton accepted the presidency in 1930. Slater, Stratton and Morse were all faculty members at MIT during the 30s. The Radiation Laboratory at Berkeley, where Lawrence was building ever bigger cyclotrons is another institution of importance. Building large cyclotrons is a "team" activity in which the "systems approach" is implicit. During the thirties building a cyclotron was de rigueur in the physics department of the best american universities. The men in charge of these efforts became the heads and group leaders at the wartime laboratories, e.g. DuBridge (Rochester, Rad Lab), Bacher (Cornell, Rad Lab and later Los Alamos). Lawrence was one of the prime movers in setting up the Rad Lab at MIT in the fall of 1940. Many of the physicists who first came to work at the Rad Lab at MIT obtained their training in Lawrence's Rad Lab: Alvarez, McMillan,... Rabi was second in command at the Rad Lab and brought with him some of the best

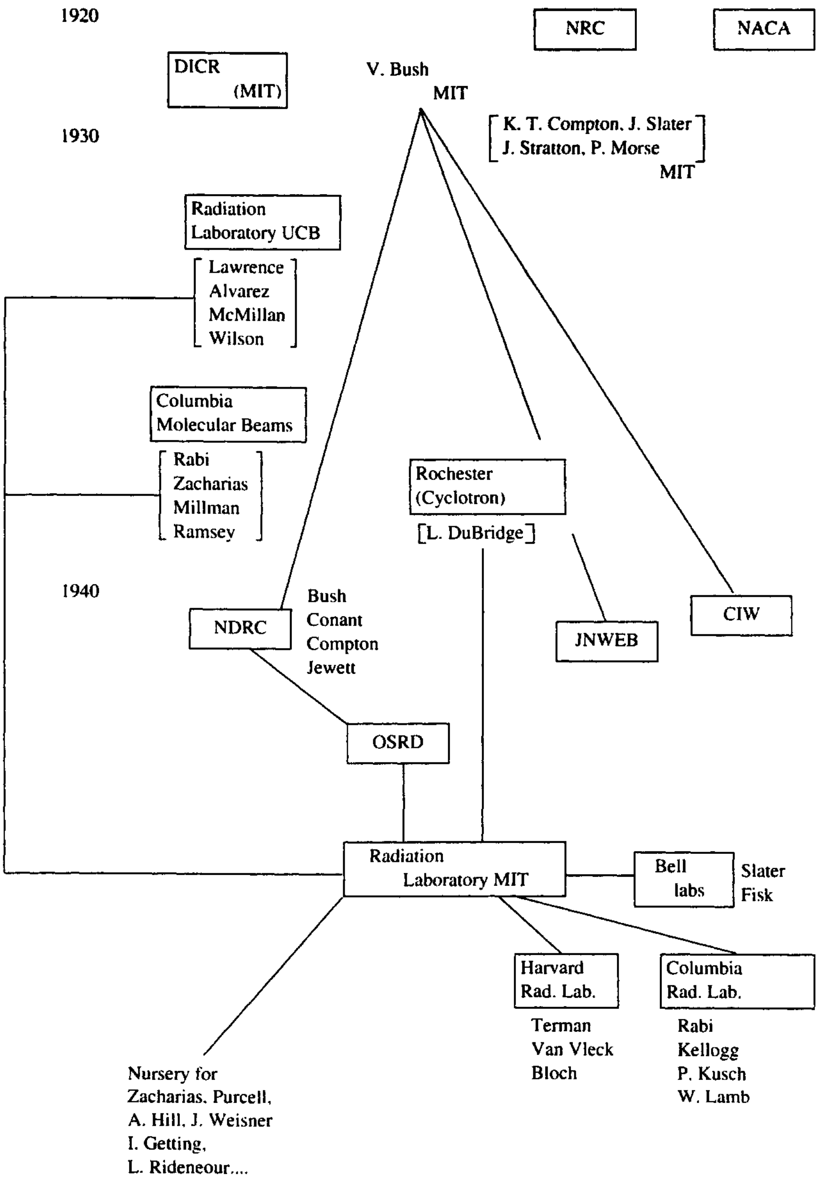

1920
NRC
NACA
DICR (MIT)
V. Bush MIT
1930
K. T. Compton, J. Slater J. Stratton, P. Morse MIT
Radiation Laboratory UCB
Lawrence Alvarez McMillan Wilson
Columbia Molecular Beams
Rabi Zacharias Millman Ramsey
Rochester (Cyclotron)
L. DuBridge
1940
NDRC
Bush Conant Compton Jewett
JNWEB
CIW
OSRD
Radiation Laboratory MIT
Bell labs
Slater Fisk
Harvard Rad. Lab.
Columbia Rad. Lab.
Terman Van Vleck Bloch
Rabi Kellogg P. Kusch W. Lamb
Nursery for Zacharias, Purcell, A. Hill, J. Weisner I. Getting, L. Rideneour....

TABLE 1

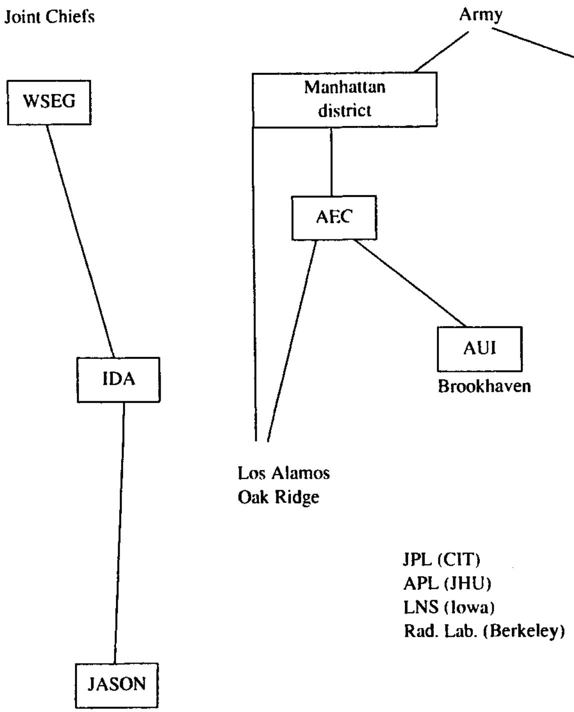
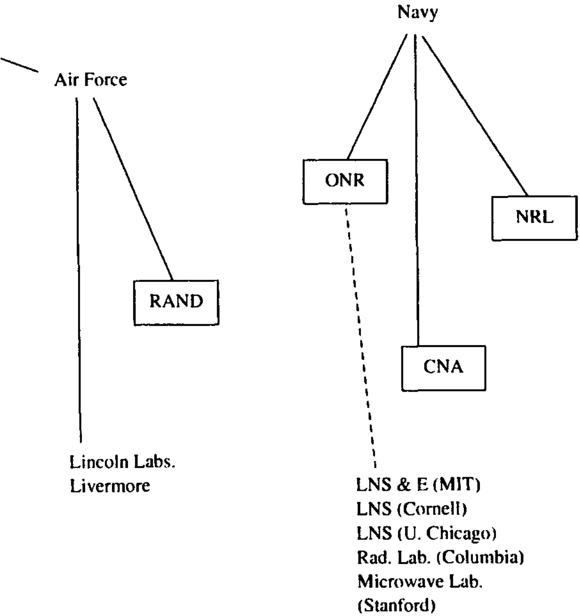

Joint Chiefs
WSEG
IDA
JASON
Army
Manhattan district
AEC
AUI
Brookhaven
Los Alamos
Oak Ridge
JPL (CIT)
APL (JHU)
LNS (Iowa)
Rad. Lab. (Berkeley)

TABLE 2


Navy
Air Force
ONR
NRL
RAND
CNA
Lincoln Labs.
Livermore
LNS & E (MIT)
LNS (Cornell)
LNS (U. Chicago)
Rad. Lab. (Columbia)
Microwave Lab.
(Stanford)

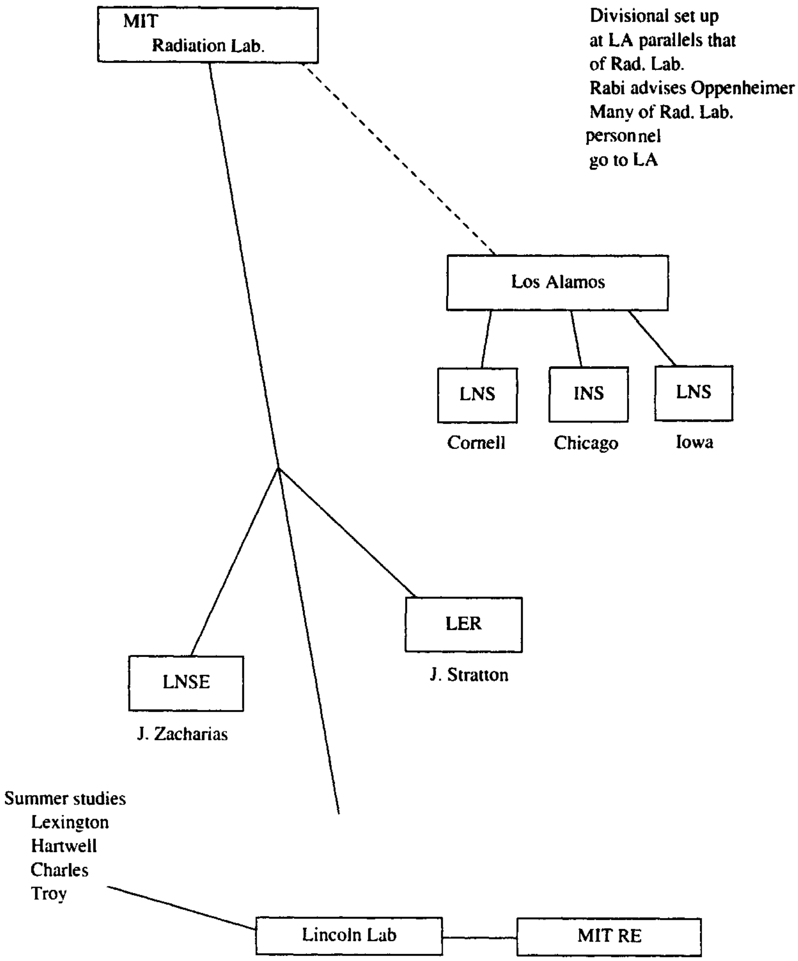

MIT
Radiation Lab.
Divisional set up
at LA parallels that
of Rad. Lab.
Rabi advises Oppenheimer
Many of Rad. Lab.
personnel
go to LA
Los Alamos
LNS
Cornell
INS
Chicago
LNS
Iowa
LER
J. Stratton
LNSE
J. Zacharias
Summer studies
Lexington
Hartwell
Charles
Troy
Lincoln Lab
MIT RE

TABLE 3

students that he had trained at Columbia: Zacharias, Ramsey,... Perhaps more than any one else Rabi set the tone for the interaction between the civilian scientists and their Navy and Army clients. The barriers that operated because information could only be disseminated on a "need-to-know" basis – a basic Navy policy – were broken down. The organizational set up at Los Alamos owes much to the advice that Rabi gave Oppenheimer at the time Groves put him in charge of setting up that laboratory. Many of the people who had worked at the Rad Lab transferred to Los Alamos in the spring of 1943 (e.g. Bethe, Bacher, Alvarez, McMillan,...).

The new laboratories that sprung up after the war and the sources of their funding are schematically illustrated in Tables 2 and 3.

*Harvard University*

## NOTES

[1] One of the characteristic features of the scientists' involvement in technical activities to support the war effort during World War is that they try to solve problems defined by the military and that they do so in government laboratories. Furthermore the information that is available to them is given on a "need to know basis."

D. Pestre and G. Pancaldi have explored the consequences of the first world war on the scientific institutions in France and Italy in papers presented at a conference on "Science, Technology, Institutions in Europe (1900–1920)" that was organized by the Consiglio Nazionale delle Ricerche (CNR) and held in Rome on the 27th and 28th of November 1990. It is interesting to compare the subsequent history of the various national research councils that were formed during the first world war. In France, Great Britain and Italy these government sponsored and supported organizations had a substantial impact on science policy in the interwar period. In the United States, although influential, the National Research Council remained a division of the National Academy of Sciences, and was supported primarily by the Rockefeller Foundation. See Heilbron and Seidel. For a discussion of Great Britain see Gummett.

[2] Ian Hacking, "Was there a Probabilistic Revolution 1800–1930?," in *The Probabilistic Revolution*. L. Kruger, G. Gigerenzer, and M. S. Morgan, eds. (Cambridge, Mass.: The M.I.T. Press, 1987) volume 1, pp. 45–58.

[3] T. S. Kuhn, *The Structure of Scientific Revolutions*. 2nd edition (Chicago: The University of Chicago Press, 1971).

Recall also that a Kuhnian revolution proceeds from a period of normal science with its paradigms and puzzle solving aspect to one of crisis in which the reigning paradigm is unable to address the observed anomalies. The crisis is resolved by the introduction of a new paradigm which ushers in a new period of normal science.

[4] It is of interest to consider the case of nuclear reactors in greater detail. Nuclear chain reactors are devices in which neutron induced fission reactions occur on a large scale and

result in the controlled release of large quantities of energy. The theory of such devices falls into two main areas:

1. the theory of the nuclear processes that occur in the chain reactor. That theory deals with the microscopic phenomena and is quantum mechanical. It is concerned with the reaction rates and the energetics of the relevant nuclear processes (fission, neutron capture, neutron scattering).

2. the theory of neutron diffusion and multiplication. This theory is classical and can be formulated in such a manner that the cross-sections for the nuclear processes enter in the formulation as *phenomenological, experimentally determined* (energy dependent) parameters. Alvin Weinberg and Eugene Wigner in their classic exposition of the theory of nuclear reactors observed that: "Of all modern technologies nuclear technology is unique in having sprung up full blown almost overnight. Only four years separate the date of the discovery of fission (1938) and the date of the first nuclear reactor. (1942). The first moderately powered reactor (Oak Ridge Graphite Reactor) went critical late in 1943; the first really high powered reactors started in 1944 at Hanford. Electricity was first produced from fission at the Experimental Breeder Reactor I in Idaho in 1951."

The presentation in their text was an attempt to resist "what is surely a deplorable trend to substitute a code for a theory, to substitute an oscilloscope display of many curves for a detailed *physical* understanding of the system." Their monograph therefore did "less that justice to the many elegant machine computing techniques which are now (1958) in vogue in many nuclear centers." Wigner and Weinberg also declared: "Let the new generation remember that the first full scale reactors, Hanford, were designed with desk calculators and slide rules." Alvin Weinberg and Eugene Wigner, *The Physical Principles of Nuclear Chain Reactors*, (Chicago: The University of Chicago Press, 1958).

[5] Hendrik W. Bode, *Synergy: Technical Integration and Technological Innovation in the Bell System*. (Murray Hill, New Jersey. Bell Telephone Laboratories, 1971) Bode credits T. N. Vail, who became the president of AT&T and chief executive of the Bell system in 1907, with articulating as early as 1909 the concepts of the "systems approach." His slogan was: "One Policy, One System, Universal Service." Vail reorganized the numerous small autonomous units that constituted AT&T into an integrated and more manageable structure with higher and more uniform standards.

[6] Project Lexington was carried out during the summer of 1948 under M.I.T. auspices. It explored the feasibility of designing and constructing a long range nuclear powered airplane. Interservice rivalry played a large role in launching the project. The project was initiated by the Air Force in order to obtain parity with the other services in matters of funding. The technological issues of an A-plane were considered and answered. The question of whether it is desirable to have such an airplane was not addressed. A program to develop a nuclear powered aircraft continued for a decade. It was canceled in 1961 by Kennedy after over a billion of dollars were spent. J. Tierny, "Take the A-Plane: The $1,000,000,000 Nuclear Bird that Never Flew." *Science* (1982).

In 1950 – shortly after the fall of China and the detonation of Joe 1, the first Russian atomic bomb – the Navy approached M.I.T. to look at the threat posed by Soviet submarines. Jerrold Zacharias – following the precedent that had been set at the Rad Lab by Rabi in dealing with the Navy during World War II – agreed to head the summer study but only on condition that he be allowed to address the larger issue of how to move ships and materials across the oceans in case of war. He had been a member of Project Lexington and had been appalled by the constraints that had been accepted in carrying out that assignment. The Navy accepted Zacharias' demand and his "Report on the security of overseas transport" was issued on September 21, 1950. The project was huge and influential. It consolidated the systems approach to Navy problems. There was a tendency for the Navy to compartmentalize functions in various Bureaus and to maintain communications among them on the basis of need to know. Rabi had broken down these barriers in so far as the activities of the Rad Lab were concerned. Zacharias insisted that that approach be maintained in the activities of Project Hartwell. S. S. Schweber, "The mutual embrace of science and the military: ONR and the growth of physics after World War II." in E. Mendelsohn, M. R. Smith, and P. Weingart (eds.), *Science, Technology and the Military.* (Dordrecht: Kluwer Academic Publishers, 1988).

Project Charles in 1951 was concerned with the Air Defence of the U.S. and led to the formation of Lincoln Labs and the Mitre Corporation. Project Vista, overseen by Cal Tech, dealt with problems connected with the defense of Europe during the summer of 1952.

[7] F. B. Llewellyn, "System Engineering with reference to its Military Applications." This is the text of speech Llewellyn, the executive secretary of the Science Advisory Committee, delivered on November 1, 1951. "Oppenheimer Papers," Library of Congress.

[8] Appendix A outlines the genealogy of some the institutions and the men involved.

[9] In general, see P. M. S. Blackett, *Studies of War. Nuclear and Conventional* (Edinburgh: Oliver and Boyd; New York: Hill and Wang, 1962) and the various papers of his reprinted therein; in particular "Tizard and the Science of War," Pt. I, 101–119; "Operational Research: 1948." Document 1, Part II, 169–198; "The Scope of Operational Research." Document 2, Part II, 199–204. See also P. M. S. Blackett, "Operational Research." *Advancement of Science British Ass.* 5 (April 1948), 28; P. M. S. Blackett, "Operational Research." *Operational Research Quarterly* 1/1 (March 1950) 3–6; P. M. S. Blackett, "Operational Research: Recollections of Problems Studied, 1940–45." *Brassey's Annual* (1953): 88–106; reprinted in *Studies of War*, 205–234. P. M. S. Blackett, "Evan James Williams. 1903–1945." *Obituary Notices of the Fellows of the Royal Society* 5 (1947) reprinted in *Studies of War*, 235-238. See also Sir Bernard Lovell, F.R.S. "Patrick Maynard Stuart Blackett," *Biographical Memoirs of Fellows of the Royal Society* 21 (1975): 1–115. Solly Zuckerman, "Scientific advice during and since World War II." *Proc. Royal Society of London A* 342 (1975): 465–47. Solly Zuckerman, *From Apes to Warlords: The Autobiography (1904–1946) of Solly Zuckerman.* (London: Hamish Hamilton, 1978). C. Kittel, "The Nature and Development of Operations Research." *Science* (ADD vol.) (1947): 150–153; M. Stone, "Science and Statecraft." *Science* (ADD vol.) (1947): 507–510; W. J. Horvath, "Operations Research – A scientific basis for Executive Decisions," *American Statistician* 2 (1947); Florence N. Trefethen, "A History of Operations Research," in *Operations Research for Management*, Joseph P. McCloskey and Florence N. Trefethen, eds. (Baltimore: Johns Hopkins Press, 1954).

[10] Blackett, *Studies of War*, 171. See also R. W. Clark, *The Rise of the Boffins* (London: Phoenix House, 1962).

[11] Perhaps the best introduction to OR and systems engineering is the volume entitled *Operations Research and Systems Engineering*, edited by Charles D. Fagle, William H. Huggins, and Robert H. Roy (Baltimore: The Johns Hopkins Press, 1960). In OR the mixed-team was usually small in number. Incidentally, mathematicians did not do well as operational analysts during WW II, presumably because they were not used to being part of "teams".

[12] O. Solandt, "Observation, Experiment and Measurement in Operations Research." *Journal OR Society of America* 3 (1955): 1–15.

[13] Characteristic names were Directorate of Naval Operational Research on the Naval Staff of the British Admiralty, the British Operational Research Section of the RAF Coastal Command, Combined Operations SE Asia Command. Neville Mott, Freeman Dyson, and H. R. Hulme are some of the physicists who were attached to some of these divisions. For an account of Desmond Bernal's work in Bomber Command and Combined Operations, his contributions to tactics and strategy while on the staff of Mountbatten in Libya and during the preparation and oversight of the Allied invasion of continental Europe, see Dorothy M. C. Hodgkin, "John Desmond Bernal," *Biographical Memoirs of Fellows of the Royal Society* 26 (1980): 17–84.

[14] See John Burchard, *Rockets, Guns and Targets: Science in World War II. History of NDRC* (Boston: Little Brown, 1948); P. M. Morse, "Operations Research", *Technology Review* February 1951, 191–217 and P. M. Morse, "Operations Research", *Technology Review* May 1953, 367–398; Jacinto Steinhardt, *Proc. U.S. Naval Institute* 72 (1946): 649–56.

[15] John von Neumann and Oskar Morgenstern, *Theory of Games and Economic Behavior* (Princeton: Princeton University Press. First edition, 1943; second edition, 1947). For a concise introduction to game theory see Martin Shubik, *Readings in Game Theory and Political Behavior* (Garden City: Doubleday, 1954).

[16] It is interesting to note that some operations research activities had already been carried out during WW I by Edison to analyze how to meet the threat posed by the German U-boats. See Lloyd N. Scott, *Naval Consulting Board of the United States* (Government Printing Office, Washington D.C., 1920); W. F. Whitmore, "Edison and Operations Research," *J. OR Society of America* 1 (1952/3): 83–85. It has been suggested that the lack of impact of Edison's useful and insightful recommendations was due to the fact they were made to a civilian official, the secretary of the Navy, who did not have operational responsibility for the deployment of ships and shipping.

[17] *The Rand Corporation. The first fifteen years.* Santa Monica, California: November 1963.

[18] *The Rand Corporation. The first fifteen years.* (Santa Monica, California: November 1963). For an overview of the approach see *Modern Systems Research for the Behavioral Scientist. A Sourcebook*, W. Buckley, ed. (Chicago: Aldine Publishing Company, 1961).

[19] T. P. Hughes, "The Seamless Web: Technology, Science, Etcetera, Etcetera," *Social Studies of Science* 16 (1986): 281–292. See also his *Network of Power* (Baltimore: Johns Hopkins University Press, 1983). During and after World War II the scope of engineering broadened to encompass the interrelation between the separate parts of a complex system, and "systems engineering" was the name given to that activity. It was the formal recognition of the importance of interactions between parts of a system. J. W. Forrester, one of the founders of M.I.T.'s Digital Computer Laboratory that produced the Whirlwind computer, and one of the guiding spirits in the development of the Sage system at the Lincoln Laboratory in the mid-fifties, accepted a position in MIT's Sloan School of Management to develop a unified "systems approach" for management education that would integrate the various activities of an industrial company into a coherent system. "This new framework, which will be developed in the next few years, will be based not on the functional divisions like manufacturing, sales, accounting and engineering, but on the underlying fundamental movements of materials, money and labor, all tied together into an information-flow and decision-making network," J. W. Forrester, "Systems Technology and Industrial

Dynamics." *Technology Review* 59 (June 1957): 417–428. See also Herbert A. Simon, *The New Science of Management Decision* (New York: Harper and Row, 196 ADD).

[20] This definition was slightly amended in Great Britain to read: "Operational Research is the use of scientific method in providing executive departments with a quantitative basis for decisions regarding the operations under their control." "Operational Research in the Research University," *Nature* 161 (1948).

[21] P. M. Morse and G. E. Kimball, *Methods of Operations Research* (New York: John Wiley and Sons, 1951), 1.

[22] Blackett, *OR Quarterly* 1/1 (1950): 3–4.

[23] In the USA: at MIT, Case, Carnegie Tech, Columbia, Cornell, Johns Hopkins, and the University of Pennsylvania; in the UK, at Birmingham and Imperial College. See J. F. McCloskey, "Training for Operations Research," *Journal OR Society of America* 3 (1953): 386–393.

[24] In 1953 George E. Kimball, a member of the NRC Committee on OR, reasserted this viewpoint: "Operations Research is not separated from such sciences as statistics and industrial engineering by any sharp boundary; there is now a large area recognized as its province. This area is that of the application of the scientific method to the operations of complex organizations. It is based on the conviction that the factors affecting such operations can be measured quantitatively and that there exist common laws obeyed by the basic variables... The main problems concerning operations research today are the discovery of such laws and the development of techniques, both techniques of measurement and techniques for rapid, simple application of existing laws." G. E. Kimball, "On a philosophy of Operations Research," *Journal OR Society of America* 2 (1954): 145.

[25] Maier, 26. *Scientific* management and Taylorism should also be seen as expressions of important distinctive characteristics of the Progressive Era: the adoption of novel ideas about the authority of science and the promulgation of science as a philosophy of social control. Peter D. Hall has noted in *The Organization of American Culture* (New York: New York University Press, 1982) that: "Thanks to the war [world war I] and their social reading of Darwin [the intellectual and social elite] could redefine their ideas about the locus of authority, shifting it from politics and religion to science, the world of matter in which the voice of God expressed itself far more authoritatively than through the voice of the people. In concrete social and political terms, this shift altered the public perception of the relation between wealth, ideas and power."

[26] Maier, *The Search for Stability*, 23. See also David Dickson, *The New Politics of Science* (New York: Pantheon Books, 1984), 319–321. There is a vast literature regarding the lure of "scientific management" as a means for increasing productivity and control over organizational activities. Max Weber indicated that the attraction of scientific organizational techniques was rooted in their potential for transforming an organization into "a precision instrument which can put itself at the disposal of quite varied – purely political as well as economic, or any sort of – interests in domination." Max Weber, *The Theory of Social and Economic Organization.* (New York: Free Press, 1947.)

[27] Le Chatelier, *Le Taylorisme* 73. See also the chapter entitled "Enseignement de L'Organisation," 142–160.

[28] See, for example, Kittel's article.

[29] See for example Thomas L. Saaty, *Mathematical Methods of Operation Research* (New York: McGraw-Hill Book Company, Inc., 1959).

[30] J. Bronowski, *Scientific American* 185 (1951): 75–77.
[31] J. Crowther and R. Whiddington, *Science at War*, 119–120.
[32] Committee on Elimination of Waste in Industry of the Federated American Engineering Societies, *Waste in Industry* (New York: McGraw Hill Book Company, 1921).
[33] Blackett, *Studies of War*, 201.
[34] Thorton L. Page, "A survey of Operations Research Tools and Techniques" in C. D. Flagle, W. H. Huggins, and R. H. Roy, *Operations Research and Systems Engineering* (Baltimore: The Johns Hopkins Press, 1960), p.120.
[35] B. L. B. Smith, *The Rand Corporation*, 18.
[36] T. Porter, HSS meeting Seattle, 28 October 1990.
[37] Peter Miller and Nikolas Rose, "Governing Economic Life" *Economy and Society* (1990); also Peter Miller and Ted O'Leary, "Accounting and the Construction of the Governable Person," *Accounting, Organizations and Society* 12 (1987): 235–265.
[38] Porter.
[39] Daniel Kevles, *The Physicists, The History of a Scientific Community in Modern America* (New York: Vintage Press, 1979).
[40] Thus Tizard became the chair of both the Advisory Council on Scientific Policy, which replaced the wartime Scientific Advisory Committee to the War Cabinet in 1947. The function of the Advisory Council on Scientific Policy was to advise the Lord President in the formulation and execution of government policy regarding non-military scientific matters; the Defense Research Policy Committee was to do the same for the Minister of Defense in military scientific policy. The panel on Technology and Operational Research (ACSP) was chaired by William Stanier, the Scientific Adviser to the Ministry of Supply. Cockcroft who replaced Tizard as the chair of the DRPC in 1952 was the head of the Atomic Energy Research Establishment at Harwell at the time. See Gummett, pp. 28–36.
[41] Thus for example Rudolf Peierls, who together with Frisch had initially proven the feasibility of an atomic bomb, and who had been a group leader at Los Alamos during the war never consulted for the British government on these matters after the war, because he was never asked. There are exceptions; the prominent ones are Blackett and Zuckerman. But their influence waned with the chilling of the Cold War (due to their identification with the left) and the advent of a conservative government in the early fifties.
[42] See H. L. Nieburg, *In the Name of Science* (Chicago: Quadrangle Books, 1966), esp. Chap. 13, "The New Braintrusters"; Ida R. Hoos, *Systems Analysis in Public Policy: A Critique* (Berkeley: University of California Press, 1972), esp. 45–67; Marcus G. Raskin, "The Megadeath Intellectuals," *New York Review of Books* (November 14, 1963); and Noam Chomsky, "The Mentality of the Backroom Boys," in *The Chomsky Reader* (New York: Pantheon, 1987), 269–288.
[43] *Report from Iron Mountain on the Possibility and Desirability of Peace*, with Introductory Material by Leonard C. Lewin (New York: Dial Press, 1967), 89–90; quoted in Hoos, *Systems Analysis in Public Policy*, 59. It is interesting to note that in 1971, Murray Gell-Mann, a member of PSAC and a Jasonite, gave a speech at the dedication of the new physics building at the University of California/Santa Barbara in which he indicated that: "the priority need is to develop a systems analysis with heart that society can rely on to choose between possible technologies." Murray Gell-Mann, "How scientists can really help." *Physics Today* May 1971, 23–25.

[44] See in this connection the Morse papers at MIT which detail his activities in bringing this about.
[45] P. M. Morse, "Must we always be gadgeteers?" *Physics Today* 3/12 (1950): 4–5.
[46] See the forthcoming biography of Jerrold Zacharias by Jack Goldstein (Cambridge: M.I.T. Press 1992).
[47] Thus among the participants in project Hartwell were Alvarez, Dicke, Getting, Lauritsen, Pierce, Purcell, Hill, Hubbard, Weisner, Piore, Berkner, Fisk.
[48] See the November 1, 1951 report by F. B. Llewellyn, the executive secretary of the Science Advisory Committee, entitled "Systems engineering with reference to military applications." Oppenheimer Papers. Archives Div. Library of Congress.
[49] There is another important issue to be addressed but which lies outside the scope of the present inquiry: "To what extent did the physicists represent pure science and its relevance and usefulness to applied science and technology?" It is worth noting that the physicists's influence waned in the early sixties. Among the many factors contributing to this, the conclusions of Project Hindsight undoubtedly played a role. "In 1963, the Pentagon had the technology of twenty essential and advanced weapons analyzed: Various nuclear warheads, rockets, radar equipment, a navigation satellite and an naval mine. As far as possible the contributions of separate scientific and technological advances made since 1945 to each weapon were traced. In this way 556 separate contributions were found. Of these 92 percent came under the heading of technology; the remaining 8 percent virtually all in the category of applied science, except for two which came from basic research." Daniel S. Greenberg. *The Politics of American Science* (New York: New American Library, 1967).

SKULI SIGURDSSON[1]

# UNIFICATION, GEOMETRY AND AMBIVALENCE: HILBERT, WEYL AND THE GÖTTINGEN COMMUNITY

In 1918 the mathematician Hermann Weyl (1885–1955) extended the general theory of relativity that Albert Einstein (1879–1955) had set forth in the years 1915–1916. At one level, Weyl's theory made it possible to unify the two field phenomena known at this time, namely those described by electromagnetic and gravitational fields. But more was at stake. At the beginning of the paper in which Weyl worked out the mathematical foundations of the theory, he observed that:

> According to this theory *everything real, that is in the world, is a manifestation of the world metric*; the physical concepts are no different from the geometrical ones. The only difference that exists between geometry and physics is, that geometry establishes in general what is contained in the nature of the metrical concepts, whereas it is the task for physics to determine the law and explore its consequences, according to which the real world is characterized among all the geometrically possible four-dimensional metric spaces.

In a footnote Weyl added triumphantly that he was sufficiently bold to believe that "the totality of physical phenomena could be derived from a single universal world law of the highest mathematical simplicity."[2]

As Thomas S. Kuhn has suggested, "the member of a mature scientific community is, like the typical character of Orwell's *1984*, the victim of a history rewritten by the powers that be .... There are losses as well as gains in scientific revolutions, and scientists tend to be peculiarly blind to the former."[3] The blind spots in old scientific texts indicate where an intellectual excavation might yield interesting results. The main aim of the present paper is to show that unification for Weyl and his contemporaries was understood not merely as a synthesis of the electromagnetic and gravitational fields but also as a unification of geometry and physics and as the quest for a universal world law accounting for both the structure of cosmos and matter. Its second purpose is to map the cognitive and social environment that nurtured and made feasible this quest for unification, namely the German university town of Göttingen. In the early twentieth century it was the location of an exceptionally vibrant community within which a belief in the mathematical comprehensibility of nature was widespread and facilitated very free exchanges between mathematicians,



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science,* 355-367.

applied mathematicians and physicists. I tackle these problems in reverse order. First I sketch the structure of the Göttingen community, and secondly the unification project itself. In the third section I consider Weyl's ambivalence towards the unification project in his later years.

## THE GÖTTINGEN COMMUNITY

The mathematical scene in Göttingen in the first decade of the twentieth century was dominated by Felix Klein (1849–1925), David Hilbert (1862–1943), and Hermann Minkowski (1864–1909). Klein was the great organizer who emphasized that cooperation and united efforts must now replace individual effort. Thus in 1893, in a lecture in Chicago Klein noted:

> Speaking, as I do, under the influence of our Göttingen traditions, and dominated somewhat, by the great name of Gauss, I may be pardoned if I characterise the tendency outlined in these remarks as a *return to the general Gaussian programme*. A distinction between the present and the earlier period lies evidently in this: that what was formerly begun by a single master-mind, we now must seek to accomplish by united efforts and cooperation.[4]

Klein realized that the dynamics of the scientific community had changed. In order to study nature it was necessary to secure financial resources, break down disciplinary boundaries, and establish strong links between science, technology and industry. Klein was a system builder and Göttingen was his system. In Göttingen he put his organizational talents to impressive use while Hilbert, the charismatic leader, built up a large school of mathematics.

Hilbert was a modern mastermind and supplier of ideas who attracted students on basis of his achievements and personality. Weyl was one of these students. He came to Göttingen in the spring of 1904. Most of the problems that Weyl considered at the beginning of his career were suggested in one form or another by Hilbert's work. Weyl clearly recognized this later and in a short obituary for Hilbert in 1944 he noted that: "Under the influence of his dominating power of suggestion one readily considered important whatever he did; his vision and experience inspired confidence in the fruitfulness of the hints he dropped." And Weyl added: "It is moreover decisive that he was not merely a scientist but a scientific personality, and therefore capable not only of teaching the technique of science but also of being a spiritual leader."[5]

In a second obituary Weyl wrote for Hilbert in 1944, he painted a striking picture of Hilbert and elaborated his views of Hilbert's scientific personality. When discussing Hilbert's impact on him and fellow students, he quoted a passage from Hilbert's memorial address for Minkowski in 1909.

Our science, which we loved above everything, had brought us together. It appeared to us like a flowering garden. In this garden there are beaten paths where one may look around at leisure and enjoy oneself without effort, especially at the side of a congenial companion. But we also liked to seek out hidden trails and discovered many a novel view, beautiful to behold, so we thought, and when we pointed them out to one another our joy was perfect.

Weyl himself continued:

I quote these words not only as testimony of a friendship of rare depth and fecundity that was based on common scientific interest, but also because I seem to hear in them from afar the sweet flute of the Pied Piper that Hilbert was, seducing so many rats to follow him into the deep river of mathematics.[6]

Weyl's description of Hilbert is strange and striking. Weyl's words simultaneously convey fascination and unease as he described Hilbert both as a spiritual leader and a Pied Piper. The unease was made palpable in an address Weyl gave in 1932 for Hilbert's seventieth birthday. During these difficult economic and political times at the end of the Weimar Republic Weyl said:

Hilbert was neither a youth leader, nor an extensive organizer, nor an all-round balanced personality, nor did he concoct world views. But we see in him, and he seems to me to be quite a strong example thereof, *how enormously positive are the effects of the naked scientific genius*, who remains faithful to his talents through industry and steadfastness in the creation of his works.[7]

Yet Weyl's remarks a decade later present the opposite picture and lead inevitably to the conclusion that Hilbert had indeed been a spiritual leader.

The main components in Hilbert's world view were his optimism, his passion and an unshakable and irresistible faith in the value of science. His faith in the value of science included a deep seated belief in the mathematical comprehensibility of the natural world. This belief was often signalled by referring to a pre-established harmony between mathematics and physics. Hilbert, like Klein and Minkowski, was very interested in physics and it would be misleading to regard him as a pure mathematician. As a matter of fact Hilbert devoted much of the second part of his career lecturing and working on physics, though he published very little on it. Apparently it was Minkowski who kindled Hilbert's enduring interest in physics. They had been close friends since their student days at the University of Königsberg in the early 1880s.

The easy movement between fields is evident in Richard Courant's description of the relation between physics and mathematics in Göttingen. When, early in the 1960s, Kuhn asked Courant whether there had been a separation between mathematicians and physicists in Göttingen, Courant answered:

> No, not at all. Hilbert of course set the tone by becoming more and more interested in physics, the quantum theory. Relativity was the beginning of his deep interest. Of course you knew, Hilbert had studied physics as a young man. He knew about old physics quite a bit. But then relativity came in, and Minkowski made all the difference.[8]

The Göttingen triumvirate found it reasonable to apply mathematics to a study of the physical world, and conversely to find mathematics cross-fertilized by physical concepts. The latter was stressed by Hilbert in his famous speech on twenty-three unsolved mathematical problems at the International Congress of Mathematicians in Paris in 1900: "But what a vital nerve important to mathematical science would be cut by the rooting up of geometry and mathematical physics!"[9] This legitimation of crossing the boundaries between mathematics and physics – evident also in the work of Emmy Noether and Ernst Zermelo – distinguishes the Göttingen tradition from, say, that in Berlin where pure mathematics in the tradition of Kronecker, Kummer and Weierstrass continued to reign supreme.

## THE UNIFICATION PROJECT AND RELATIVITY THEORY

Hilbert addressed unification at the International Congress of Mathematicians in Rome in 1908. There Hilbert discussed the difficulties inherent in trying to determine infinitely many unknowns from infinitely many equations, a concern following upon his work on the theory of integral equations. Much was uncertain in this realm but the rewards of success were great, as Hilbert stressed, for "if we are not allowed to be put off, by such considerations, then we will be like Sigfried, for whom the magic fire retreated by itself, and as a reward we will receive the beautiful prize of *a methodologically unified form of algebra and analysis.*"[10]

Unification took a more ontological turn with Hilbert's work on the general theory of relativity. Einstein and Hilbert had discovered the field equations of gravitation in the fall of 1915. Hilbert derived them with variational methods. According to Hilbert the emergence of the field concept gradually made possible a new conception of the physical world. In particular it was the physicist Gustav Mie who had showed how this new ideal of field theoretical unity succeeded the ideal of mechanical unity and "could be made amenable to a general mathematical treatment." Hilbert added that the magnificent problems and thought structures of Einstein's general theory of relativity found their "simplest and most natural expression" in the manner suggested by Mie and were furthermore "systematically extended and formally rounded off."[11] Hilbert erroneously believed the four electromagnetic potentials to be determined by the ten gravitational potentials. Max Born observed:

In this, Hilbert himself saw the height of his achievement, namely the discovery of the link between the two kinds of fields that until then had remained without any connection. ... he expected something greater from his field equations, i.e. reaching the goal that Mie had had in mind: the derivation of the electron and the atoms from the field equations.[12]

Born noted this goal had not yet been reached, and doubted that it could be reached in such a formal fashion.

Hilbert wanted to secure and unify the foundations of mathematics and physics. According to Hilbert, physics was but a four-dimensional pseudo-geometry, whose metric was connected via his theory to electromagnetic quantities, i.e. to matter. And with this knowledge an old geometric problem could now be solved: "Whether and in what sense the Euclidean geometry – about which we only know from mathematics that it is a logically consistent structure – is also valid in reality." After discussing Gauss's inability to verify empirically a non-Euclidean physics through angle measurement in a large triangle, Hilbert talked about how the physics of Einstein's general theory of relativity had a totally different relationship to geometry. The new physics started neither with Euclidean nor with any other fixed geometry in order to deduce the actual laws of physics. Instead general relativity yielded in one blow the laws of geometry and physics through one and the same Hamiltonian principle, i.e. through the fundamental equations of his theory. Hilbert's conclusion was:

Euclidean geometry is *an action-at-a-distance law alien to modern physics*: while the theory of relativity rejects Euclidean geometry as a general presupposition for physics, it teaches furthermore that geometry and physics are of a similar kind and rest as *one* science (*Wissenschaft*) on a common foundation.[13]

It was Klein's contribution that he immediately understood the implications of Hilbert's ideas and presented them in a clear and simple fashion in the first of three papers on the general theory of relativity, whereby Klein also rejuvenated himself. Hilbert's work on the foundations of physics and Klein's further elaboration introduced in the very broadest sense of the word Einstein's general theory of relativity to mathematicians. By their attention, Hilbert and Klein had legitimated Einstein's theory among Göttingen mathematicians. And in the background loomed Minkowski's four-dimensional representation of the special theory of relativity.

The Göttingen community appropriated Einstein's general theory of relativity in a framework that was an amalgam of earlier work in Göttingen on electron theory, the special theory of relativity, Mie's electromagnetic view of nature, Max Abraham's and Gunnar Nordström's research on relativity theory, and mathematical research on the calculus of variations, axiomatization and geometry. This thrust the development of the general theory of relativity in a direction different from Einstein's original intentions. The Göttingen appropriation implied that Einstein's theory was not only considered as a theory of gravitation as Einstein had conceived it, but also – and this is the crucial difference – as a model in which matter theory and mathematics came closely together.

Einstein immediately realized this. In a letter written to Weyl in the fall of 1916, at that time professor of mathematics at the Eidgenössische Technische Hochschule (ETH) in Zürich, he expressed delight with Weyl's warm interest and enthusiasm for his work. Einstein needed all the support he could get, and although the general theory of relativity still encountered considerable opposition, he found comfort in the fact "that as can be determined the average mental power of its supporters by far surpasses that of its opponents. This is a kind of objective evidence for the naturalness and sensibleness of the theory." He then spoke about his difficulties in understanding the Hamiltonian for matter, and admitted that that was an important task. But he doubted the value of the axiomatic method, which adopted all sorts of unphysical hypotheses. Thus he was doubly skeptical of Hilbert's approach: "The Hilbertian formulation for matter seems childish to me, in the sense of a child that recognizes no malice in the external world."[14]

The outbreak of World War I had a tremendous impact on Weyl. He had turned to the general theory of relativity after being released from military service in the spring of 1916. Upon his return to Zürich he sought a release from the horror of the War.

Einstein's "*Grundlage der allgemeinen Relativitätstheorie*", published in 1916, announced a truly epochal event, the reverberations of which extended far beyond the confines of mathematics. It also made epoch in my own scientific life. In 1916 I had been discharged from the German army and returned to my job in Switzerland. My mathematical mind was as blank as any veteran's and I did not know what to do. I began to study algebraic surfaces; but before I had gotten far, Einstein's memoir came into my hands and set me afire.[15]

Einstein's theory was mathematically challenging, intriguing philosophically and hinted at comprehending the orderliness in the universe. Thus it offered the reassuring possibility of a simultaneous spiritual and scientific regeneration.

Weyl's *Raum-Zeit-Materie*, the first high level textbook on the general theory of relativity, appeared in the spring of 1918. It was based on lectures he gave at the ETH during the summer semester of 1917, which Weyl had been encouraged to publish by Einstein's close friend Michele Besso. Weyl sent Einstein the page proofs of the book in the late winter of 1918, and said that he was already further extending the general theory of relativity and Mie's matter theory. "I have, I believe, managed to derive electricity and gravitation from a common source." He asked, if Einstein would communicate a paper entitled "Gravitation and Electricity" on these new developments to the Berlin Academy. Einstein did so in the spring of 1918.[16]

There Weyl presented an imaginative extension of the general theory of relativity and of Riemannian geometry which Einstein employed in his theory. Today this paper is remembered for having introduced the idea of gauge invariance which, suitably modified, has been of importance in modern physics in the theory of elementary particles and unified field theories. Before Weyl, one could not expect a vector, unless the geometry was Euclidean, to preserve its direction if parallely transported around a closed curve, although its length was preserved. The amount of tilting of the vector would be a function of the area enclosed by the curve and the curvature of space enclosed by that area. Thus the transition from Euclidean to Riemannian geometry was seen to consist in this: the change of the direction of a vector between two points depended on the path. Direction was thus not integrable in Riemannian geometry, but length was.

Weyl challenged the second assumption. The novelty in his approach was to take seriously the demands of local geometry (*Nahegeometrie*) and to insist that "*in a truly local geometry it is only possible to transfer length from one point to another infinitely close by.*"[17] Hence one could no more assume from the start that the length of vectors was integrable than their inclination. This plea sounds very similar to the declaration Hilbert had made in his research on the foundations of physics, namely that Euclidean geometry was an action-at-a-distance law alien to modern physics. At the beginning of the paper Weyl wrote:

In the Riemannian geometry characterized above, there has been preserved the last element of a global geometry – as far as I can see, without any objective justification; only the accidental origin of this geometry in the theory of surfaces seems to be to blame.[18]

A few days after Weyl had sent Einstein the page proofs of his article on gravitation and electricity, he sent Einstein a letter at the end of which he came back to the inconsistency of Riemannian geometry. Weyl claimed, "My geometry is the true local geometry" and the reason that Riemann did not treat it had "merely to do with historical reasons (origin in the theory of surfaces), but no objective ones. If you are right with regard to the real world, then I regret having to point out a mathematical inconsistency to the dear Lord (*lieber Gott*)."[19] When Weyl spoke about a mathematical inconsistency on behalf of the "dear Lord" he was referring to the devastating objections that Einstein had raised against the physical interpretation of Weyl's geometry which if successful would have enabled Weyl to unify gravitation and electricity.

The equivalence of inertial and gravitational mass had led Einstein to the geometric interpretation of the theory of gravitation. But in the theory of electricity no such fact was known, for there was no reason to believe in a universal influence of the electromagnetic field on so-called rigid rods or watches. As Einstein's objection to Weyl's theory demonstrated, the stability of spectral lines spoke decisively against the possibility of such universal electric influence. Fritz London commented on the unusually clear metaphysical conviction that enabled Weyl, in spite of Einstein's criticism, to believe that "nature would have to make use of this beautiful geometrical possibility that was offered to her."[20]

Cosmology determined some of the boundary conditions in Weyl's theory. In 1921 Wolfgang Pauli could not help expressing some wonderment at this turn of events: "We are thus led to conjecture that a connection exists between the size of the universe and that of the electron, which might seem somewhat fantastic."[21] Pauli's criticism highlights how technically the unified field theory program brought together the cosmos and matter. In that sense the "world" in "universal world law" was not empty rhetoric.

## WEYL'S AMBIVALENCE TOWARDS UNIFICATION

Weyl had been close to Hilbert as a student in Göttingen, but towards the end of World War I a noticeable estrangement had taken place. It became public when Weyl sided with the Dutch mathematician Luitzen Egbertus Brouwer in the controversy about intuitionism and the foundations of mathematics, the so-called "Grundlagenkrise" that erupted around 1920. Towards the end of the 1920s a rapprochement occurred. By 1925 Weyl's opposition to Hilbert on the foundations of mathematics had become less pronounced.[22] In 1927 Weyl wrote to a fellow Hilbert student, the mathematician Robert König, describing how happy he was that his relationship with Hilbert had become cordial again.

> He is awfully nice to me, and I am quite happy having such a harmonious relationship again with Hilbert, that again I can feel free towards him the way I did in our mathematical youth, where love and admiration flowed towards him; often I can hardly help being moved if I look at his small face which has become quite white under the mighty furrowed head. The spirit in which we do mathematics we have, after all, received from him.[23]

Weyl's two obituaries for Hilbert were reminiscences of the most intimate sort. Weyl downplayed the importance of Hilbert's work in physics – it could hardly be compared to his achievements in pure mathematics – while extolling Hilbert's achievements in pure mathematics. Why did Hilbert's plans in physics fail?

> The maze of experimental facts which the physicist has to take into account is too manifold, their expansion too fast, and their aspect and relative weight too changeable for the axiomatic method to find a firm enough foothold, except in the thoroughly consolidated parts of our physical knowledge. Men like Einstein and Niels Bohr grope their way in the dark toward their conceptions of general relativity or atomic structure by another type of experience and imagination than those of the mathematician, although no doubt mathematics is an essential ingredient.

Yet Weyl clearly recognized the value of Hilbert's application of integral equations to the kinetic theory of gases and the elementary theory of radiation, which Hilbert had made around 1910 before turning to the general theory of relativity. Weyl then discussed Hilbert's work on general relativity and unified field theories and noted that: "Hopes in the Hilbert circle ran high at that time: the dream of a universal law accounting both for the structure of the cosmos as a whole, and of all the atomic nuclei, seemed near fulfillment." But now in the 1940s Weyl felt that a unified field theory would have to include matter waves as well as gravitation and electromagnetism. It could no longer simply be an enlargement of "Einstein's now classical theory of gravitation."[24]

Weyl seems in part to have been asking himself how at the end of World War I he could have been so infatuated with a belief in a universal world law explaining both the structure of cosmos and matter. Now in the mid 1940s he felt that it was insufficient to unite electromagnetism and gravitation, but that quantum and nuclear phenomena had to be taken into account as well. By focusing on the unification of the physical fields known at that time, Weyl continued the re-definition of the aims of the unification project that he had begun in the late 1920s with the advent of the new quantum mechanics. As the above passage from the Hilbert obituary shows, in the mid 1940s Weyl would not even discuss the union between geometry and physics that seemed so attractive in the 1910s. The historian, following Kuhn, must thus regard critically recent narratives aimed at constructing a streamlined lineage from Weyl's early work on unification to gauge theories in modern physics: the past has to be brought faithfully to light and confronted in its bewildering richness.

One way of interpreting Weyl's obituaries for Hilbert is that they were Weyl's attempt to confront his own rich past and give it coherent structure. In other obituaries from the 1930s onwards and in correspondence with the biographers of Einstein and Arthur Stanley Eddington, Weyl stressed his primary identity as a mathematician.[25] This fitted the new self-image that emerged after Weyl's emigration to the United States in 1933. In his new country, Weyl concentrated on mathematics and devoted much less of his time to physics and philosophy, activities which had taken up so much of his energies while in Europe.

The effects of this intellectual migration are visible in Weyl's correspondence from the late 1940s with Fanny Minkowski, the sister of Hermann Minkowski. She intended to honor her brother's memory by writing about him, and she wanted to enlist Weyl's help in the effort, while Thomas Mann was to be her literary advisor. Weyl wanted to have none of this, claiming that he had never been close to her brother. He also emphasized that in his view Minkowski's deepest contribution lay in number theory, i.e. the geometry of numbers. It was in this field that young mathematicians carried on Minkowski's work, and that was "the most beautiful reward that a scientist can be granted." With regard to physics Weyl felt in spite of his work under Heinrich Hertz and his last work on the theory of relativity, and that although he had given the special theory of relativity its appropriate mathematical formulation, Minkowski stood noticeably further away from physics than he did to mathematics. Weyl told Fanny Minkowski that her brother had "aimed too high" in his famous lecture in Köln in 1908 by claiming that space and time by themselves were "doomed to fade into mere shadows, and only a kind of union of the two will preserve an independent reality."

Someone who contributes to a field foreign to himself is easily inclined, in the pride of also having mastered something foreign and lacking an overall view, to make an exaggerated assessment of his contribution. The lecture suffers also from the fact that he wanted to fix or immortalize a transitional phase in physics. ... Einstein's position was much freer from the beginning.[26]

Weyl was also telling Fanny Minkowski that at the end of World War I he himself had been too eager to contemplate a fusion of geometry and physics and prematurely believing that a universal world formula was within sight. Then he believed in having nearly reached a final stage of physics, in immortalization.

## CONCLUSION

In this paper I have discussed the unification project at the end of World War I. Relativity theory also made it possible to look at the "world" philosophically. Weyl wanted to give up an artificial separation of the world, the scene of reality, into space and time. "This objective world does not *happen*, but it *is* – simply; a four-dimensional continuum, but neither space nor time." After discussing how individuals experience their lives and history, he continued:

But the world itself has no history. Thus modern physics shows advantageously – long after physics has been liberated from the sense qualities – the great recognition of Kant that space and time are only the forms of our intuition without meaning for the objective. Admittedly different from the Kantian philosophy, physics has the courage to penetrate and represent in mathematical symbols the space- and timeless realm of "things in themselves" that is hidden behind the appearances.[27]

This remarkable statement demonstrates how taken Weyl was with Minkowski's view of relativity theory around 1920. It illuminates Weyl's denial in his letter to Fanny Minkowski and shows the seriousness attached to the well-known Kantian epistemological precepts in the German cultural context. How intellectually intoxicating it must have been to be on the verge of challenging them successfully and simultaneously demonstrating the power of mathematics. This world beyond was eternal, devoid of the contingency of personal lives and human history. Such avenues of emotional rescue had tremendous emotional appeal in the aftermath of World War I.

*University of Göttingen, Germany*

## NOTES

[1] I acknowledge the support of the Alexander von Humboldt-Stiftung and Verbund für Wissenschaftsgeschichte Berlin for writing this paper. It draws upon Sigurdsson, "Hermann Weyl, Mathematics and Physics, 1900–1927" (Ph. D. dissertation, Harvard University, 1991). All translations are mine unless otherwise indicated. I thank William Clark, Lorraine Daston, Joan L. Richards, Tilman Sauer and Silvan S. Schweber for reading and criticizing earlier versions of this paper.
[2] Hermann Weyl, "Reine Infinitesimalgeometrie," in Weyl, *Gesammelte Abhandlungen*, 4 vols. (Berlin: Springer, 1968), Vol. II, pp. 1–28, on p. 2; emph. in orig.
[3] Thomas S. Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1970 [1st ed. 1962]), on p. 167.
[4] Felix Klein, "The Present State of Mathematics," in Klein, *Gesammelte mathematische Abhandlungen*, 3 vols. (Berlin: Julius Springer, 1922), Vol. II, pp. 613–615, on p. 615; emph. in orig.
[5] Weyl, "Obituary: David Hilbert (1862–1943)," in Weyl, *Ges. Abh.* (ref.2), Vol. IV, pp. 121–129, on p. 128.
[6] Weyl, "David Hilbert and His Mathematical Work," in Weyl, *Ges. Abh.* (ref. 2), Vol. IV, pp. 130–172, on p. 132.
[7] Weyl, "Zu David Hilberts siebzigstem Geburtstag," in Weyl, *Ges. Abh.* (ref. 2), Vol. III, pp. 346–347, on p. 347; emph. in orig.
[8] Richard Courant, interviewed by Thomas S. Kuhn, 9 May 1962, Archive for History of Quantum Physics.
[9] David Hilbert, "Mathematische Probleme," in Hilbert, *Gesammelte Abhandlungen*, 3 vols. (Berlin: Julius Springer, 1935), Vol. III, pp. 290–329, on p. 295.
[10] Hilbert, "Wesen und Ziele einer Analysis der unendlichvielen unabhängigen Variablen," in Hilbert, *Ges. Abh.* (ref. 9), Vol. III, pp. 56–72, on p. 57; emph. in orig.
[11] Hilbert, "Die Grundlagen der Physik (1924)," in Hilbert, *Ges. Abh.* (ref. 9), Vol. III, pp. 258–289, on p. 258. This paper is a condensed version of Hilbert's two notes on the foundations of physics of 20 November 1915 and 23 December 1916.
[12] Max Born, "Hilbert und die Physik (1922)," in Born, *Ausgewählte Abhandlungen*, 2 vols. (Göttingen: Vandenhoeck & Ruprecht, 1963), Vol. II, pp. 584–598, on pp. 595–596.
[13] Hilbert, "Die Grundlagen der Physik (1924)," in Hilbert, *Ges. Abh.* (ref. 9), Vol. III, pp. 258–289, on p. 278.
[14] Albert Einstein to Hermann Weyl, Berlin, 23 November 1916, Nachlass Hermann Weyl HS 91: 536, ETH Library Archives Zürich. Hereafter abbreviated *NWeyl*.
[15] Weyl, Lecture at the Princeton Bicentennial Conference, December 1946, *NWeyl*, HS 91a: 18.
[16] Weyl to Einstein, Zürich, 1 March 1918, *NWeyl* HS 91: 538a.
[17] Weyl, "Gravitation und Elektrizität," in Weyl, *Ges. Abh.* (ref. 2), Vol II, pp. 29–42, on p. 30.
[18] *Ibid.*, on p. 30.
[19] Weyl to Einstein, Zürich, 19 May 1918, *NWeyl* HS 91: 545a.
[20] Fritz London, "Quantenmechanische Deutung der Theorie von Weyl," *Zeitschrift für Physik* **42** (1927), pp. 375–389, on p. 377.
[21] Wolfgang Pauli, *Theory of Relativity* (New York: Dover, 1981 [English transl. 1958]), on p. 202.

[22] Paul Bernays to David Hilbert, Charlottenburg, 25 November 1925, Nachlass David Hilbert 21, Niedersächsische Staats- und Universitätsbibliothek, Göttingen.
[23] Weyl to Robert König, 21 February 1927, quoted in König, "Hermann Weyl," *Bayerische Akademie der Wissenschaften: Jahrbuch* (1956), pp. 236–248, on p. 243.
[24] Weyl, "David Hilbert and His Mathematical Work," in Weyl, *Ges. Abh.* (ref. 2), Vol. IV, pp. 130–172, on p. 171.
[25] Weyl, "Emmy Noether (1935)" in Weyl, *Ges. Abh.* (ref. 2), Vol. III, pp. 425–444; Weyl to Carl Seelig, Zürich, 19 May and 26 June 1952, Seelig papers, HS 304: 1062 and 1063, ETH Library Archives Zürich; and Weyl to A. Vibert Douglas, Zürich, 31 October 1953, *NWeyl*, HS 91: 173.
[26] Weyl to Fanny Minkowski, Princeton [?], 24 March 1947, *NWeyl*, HS 91: 378. Minkowski's lecture "Space and Time" is reprinted in H. A. Lorentz *et al.*, *The Principle of Relativity* (New York: Dover, 1952 [English transl. 1923]), pp. 75–91, on p. 75.
[27] Weyl, "Die Einsteinsche Relativitätstheorie," in Weyl, *Ges. Abh.* (ref. 2), Vol. II, pp. 123–140, on p. 131; emph. in orig.

ALEXANDER A. PECHENKIN

# THE TWO-DIMENSIONAL VIEW OF THE HISTORY OF CHEMISTRY

## SOME PRELIMINARY REMARKS

As far as this article is concerned with the metaphor of the dimensions of the history of science, it needs some comments on G. Holton's discussion of the two-dimensional view of science.[1] G. Holton calls the standard philosophical view of science which has its roots in empirism or positivism a two-dimensional view. To explain that view he uses a mnemonic device of two orthogonal axes representing the two dimensions of a plane. These dimensions are phenomenal and analytic. A scientific statement, in the "standard" view, is analogous to an element of area in the plane, and the projection of it onto axes are the aspects of the statement that can be rendered, respectively, as the phenomenal aspect (protocol of observation) and the analytic one (protocol of calculation). In other words, any scientific statement has "meaning" only so far as it can be shown to have phenomenal and/or analytic components in the plane.

G. Holton emphasizes that the two-dimensional view of science overlooks the existence of active mechanisms at work in the day-to-day experience of those who are actually engaged in the pursuit of science. For example, it is of little help in handling questions a historian of science has to face. Therefore, G. Holton introduces the third dimension in his view of science, the thematic dimension.

In contrast to Holton's writings, this article is about the dimensions of the history of science rather than science itself. Although the same mnemonic device of two orthogonal axes on a plane is used in this article, that device has another meaning. For Holton, an element of "the mathematical space" is a scientific statement, in this article it is an historical event, that is, the development or appearance of a scientific paper, or an influential scientific address, or a specific discovery, etc.

G. Holton has developed his mnemonic device criticizing the standard philosophic view of science close to positivism and empiricism. I am developing that metaphor in contrast to the one-dimensional approach to the history of science typical of many historical and philosophical writings, especially writings in the history and philosophy of chemistry.



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 369-377.

Holton shows the third dimension of scientific discource, the thematic dimension. I will try here to specify the two-dimensional scheme of the history of chemistry. These dimensions are the "physicalization" of chemistry (a copy of the mathematization of physics) and the system approach to chemical facts. In terms of this scheme, an historical event in chemistry is analogous to an element of area in the plane, and projected is the aspect of the event that can be rendered as physical aspects and as aspects of the system research. It is possible to describe the third, forth, and further dimensions of the history of chemistry. For example, one may treat the third dimension as a dimension of mathematization of chemistry that is developing both within the framework of the physicalization of chemistry and out of the framework (for instance the mathematization of phenomenological chemical kinetics). However, in this article I am concerned with the two-dimensional scheme of the history of chemistry, arguing with the one-dimensional approach to the history of the science.

## THE ONE-DIMENSIONAL APPROACH TO THE HISTORY OF CHEMISTRY

The one-dimensional approach to the history of chemistry was widespread in literature in the USSR in 1950–1980. Allow me to discuss the approach of B. M. Kedrov, the famous Soviet philosopher and historian of chemistry (1903–1985)[2] It is not difficult to recognize that Kedrov had been inspired by F. Engels's writings on chemistry and, generally, by Hegel's logic. Kedrov described the following "internal logic" of the development of chemistry: the study of quality (the theory of phlogiston), the study of quantity (the chemistry of gases launched by C. W. Scheele and J. Priestley under the dominance of the phlogiston theory), the study of measure (the discovery and formulation of stoichiometric laws, namely, the law of constant proportion, the law of equivalents, the law of multiple proportions), the study of essence (the atomic theory launched in chemistry, in Kedrov's opinion, by J. Dalton).

Kedrov proposed the following triangular scheme to describe the development of atomism in chemistry.

composition – property

structure[3]

The horizontal line "composition – property" in this scheme symbolizes the problem of explanation of properties on the basis of chemical composition. This traditional chemical problem became more strict and specific as Dalton formulated the atomic theory. However, when the phenomena of polymerism and, especially, isomerism were discovered, chemists started recognizing the incompleteness of the explanation of properties on the basis of composition. They started moving step by step toward the conception of structure which was first formulated in organic chemistry (at the beginning of 1860) and later in non-organic chemistry.

The above triangle scheme of the history of chemistry from 1800 to 1860 was also proposed and argued in more detail by G. V. Bykov, a prominent Soviet historian of organic chemistry.[4]

Kedrov's rational reconstruction can be criticized from different positions. I would like to stress in this article that Kedrov overlooked the mechanic models of explanation of chemical facts or rejected them as "dead-end metaphysics". Nevertheless, those models played a great part in the development of 18th and 19th century chemistry, providing the ideological base for the discovery and formulation of some of chemistry's laws. The picture of historical events in 18th and 19th century chemistry fails to be complete if the development of mechanical models of chemical phenomena is not taken under examination.

According to Kedrov, modern scientific chemistry began with the theory of phlogiston which expressed however direct *Show* (Hegel's *Shein*) which was set up for *Reality*. But the Lavoisier oxygen theory, which had replaced the theory of phlogiston by the end of the 18th century, revealed, in his opinion, the way to the cognition of real *Essence* of chemical phenomena. Kedrov overlooked the Newtonian and atomic grounds of the theory of phlogiston that had been described by historians of chemistry.[5] However, the theory of phlogiston arose within the Newtonian paradigm treating phenomena on the basis of the force interaction of gravitating bodies. This paradigm dominated until the end of the 1800, and many theories of chemical affinity traced its roots to the theory of phlogiston. We are therefore forced to take this theory very seriously as the first scientific theory of affinity. Moreover, the Guldberg–Waage Mass Action Law, the fundamental law of chemical kinetics, was been discovered and formulated on the basis of the model of gravitating bodies. But following the principle of observability, Guldberg and Waage replaced the force of affinity by the velocity of reaction, and the masses of gravitating bodies by the concentrations of reactive substances.

Kedrov, who refered very often to C. Schorlemmer's book *The Rise and*

*Development of Organic Chemistry* had overlooked an important remark of Schorlemmer. The genesis of modern scientific chemistry, according to Schorlemmer, had been provided by the physicalization of chemical knowledge.[6]

Proclaiming the above triangle scheme of the cognition of chemical structure, Kedrov and Bykov did not appreciate the dynamic (mechanic) models of a molecule that had been put forward by Dumas, Laurent and other chemists. Those models had helped to phrase the problem of the interconnection between the structure of a molecule and its reaction ability, resulting in the ideological presumptions of electronic, and later quantum theories of chemical structure.[7]

While rejecting Kedrov's one-dimensional approach to the history of chemistry and proclaiming a many-dimensional approach to it, I preserve the idea of historicism implied by the one-dimensional approach. By that I mean the idea of a general trend, a "line" in a number of historical events. For me, this idea has gained in importance to meet the attacks of extreme relativism typical of the contemporary philosophy of science. However, in contrast to the one-dimensional ideology, I insist on a plurality of general trends and lines in the history of science.

## THE PHYSICALIZATION OF CHEMISTRY

Thus, I propose the two-dimensional scheme of the history of chemistry. This division is about the dimension called "the physicalization of chemistry".

The history of chemistry exhibits three stages of physicalization. At the first stage, physical ideas started to penetrate chemistry. This process began in the 18th century. Chemistry was greatly influenced by the Newtonian idea of force, practically all the concepts of chemical affinity (from those that were formulated within the framework of the above phlogiston theory to the Mendeleev concept) were cut according to the pattern of the Newtonian concept of gravitational interaction. Chemical affinity was interpreted as gravitation between particles of matter akin to universal gravitation. It should be added that laws of chemistry were not derived from mechanics. They were formulated by analogy the laws of chemical affinity were based on chemical experiments but carried a physical idea in them: It was supposed that they show the degree of gravitation between substances.

At the second stage of physicalization, chemical theory was constructed on the basis of physical theory. This was also associated with the presumption that a rigorous interpretation of chemical phenomena can be found in physical theories and that, being too complicated to admit a direct application of physical methods, they are described by approximate methods. In other words, chemical systems were regarded as disturbances, compositions, and modifications of simple systems governed by physical laws.

The second stage was ushered in by work in chemical thermodynamics late in the last century. Thermodynamics was constructed as a theory of ideal systems. The principles of this theory show how to formulate and tackle problems for ideal systems (such problems played the role of paradigms in thermodynamics). Chemical thermodynamics dealt with real (non-ideal) substances (solutions, gases, etc.) and was erected by adjusting the methods of rigorous (physical) thermodynamics formulated for ideal systems to suit the tasks dealing with more complicated real systems. The rigorous thermodynamic formulae were approximated by special empirical parameters.

At the end of the Twenties and into the Thirties, quantum chemistry was formulated and further developed as a specific addition to quantum mechanics. It provided an interesting example of the second stage of the physicalization of chemistry as well. In certain respects the development of quantum chemistry repeated the development of chemical thermodynamics. In principle, quantum mechanics can explain the phenomenon of chemical bonds and the structure of molecules. At the same time, being a quantitative theory, it is able to describe the simplest molecular system (the ionized molecule of hydrogen $H_2^+$). Quantum chemistry is a total sum of the approximate methods of quantum mechanics. This field rests not on the direct solution of the differential equation of quantum mechanics, but on the discussion of chemical bonds and molecular structures, according to the pattern of those problems of quantum mechanics that are formulated for simple molecular systems and which allow for a rigorous or sufficiently rigorous solution. Thus, the method of molecular orbitals, one of the basic methods of quantum chemistry, interprets the complex molecule as a composition of one-electron systems, akin to the ionized molecule of hydrogen. Electronic correlations, that is, the disturbance introduced by interelectronic relations and interactions, are taken into different degrees of account in the variants of this method.

The third stage of the physicalization of chemistry can be called

Prigogine's stage as it was first presented by I. Prigogine and his coworkers' writings. To some extent, this stage is close to the second stage. Prigogine and his coworkers provided the rigorous mathematical and physical description of a simple kinetic system, the three-molecular chemical reaction, $A \rightarrow X$, $2X + Y \rightarrow 3X$, $B + X \rightarrow Y + D$, $X \rightarrow E$. They called their basic model "The Brusselator" and considered it to be a paradigm in the physical examination of more complicated kinetic systems, including systems of biochemistry.

One can observe, however, a considerable paradigm change in physics by reading Prigogine's writings. In contrast to quantum mechanics which first arose in the course of physical research, Prigogine's nonlinear, nonequilibrium thermodynamics resulted from the sophisticated elaboration of chemical kinetics and chemical thermodynamics. This theory was cut according to the pattern of chemical thought.

A comparison of the following two quotations can confirms this claim of a paradigm change in physics. The Feynman lectures on physics state that "physicists always have a habit of taking the simplest example of any phenomenon and calling it Physics, leaving the more complicated examples to become the concern of other fields say, of applied mathematics, electrical engineering, chemistry, or crystallography".[8]

Nicolis and Prigogine write that "A nonlinear thermodynamics is a thermodynamics of chemical reactions".[9] They considered chemical phenomena not only as a field of application of physical theory, but as its original subject matter.

## THE DEVELOPMENT OF SYSTEM RESEARCHES IN CHEMISTRY

Let us turn to the other dimension of the history of chemistry which has been called "the development of system research in chemistry." This dimension can be presented by the following scheme:

composition – property

structure – function

organization – behaviour

self-organization

This scheme begins with the above Kedrov–Bykov triangle "composition

–property–structure", symbolizing the development of the conception of structure in chemistry. The second triangle "structure–function–organization" symbolizes the development of chemical kinetics which began with attempts to explain the reactivity (chemical function) of a molecule on the basis of its structure, and resulted in the recognition of the principal incompleteness of any explanation of reactivity on that basis. Chemical kinetics became a fundamental theory when it left the frontiers of the conception of structure and turned to a new set of concepts which is marked in the above scheme as "organization".

An important contribution to the development of chemical kinetics was the conception of the mechanism of a reaction, elaborated in the 1870s and 1880s on the grounds of the above Mass Action Law. However, this conception could not grasp the very nature or a chemical process. The mechanism of a reaction is its temporal structure. It does not have a great deal to do with the system of kinetic factors of a reaction: the nature of a solvent, the texture of vessel walls, the velocity of a mixer, etc. These factors are considered to be disturbances of the mechanism. The golden age of chemical kinetics came in the 1920s when the theory of chain reactions, the first theory of complex reactions, was highly developed and the interconnections between the mechanism of a reaction and its kinetic factors were taken under theoretical investigation and expressed in laws and formulae.

With the idea in mind of the domination of kinetics in chemistry, N. N. Semenov claimed that the kinetic reinterpretation of the conceptual techniques of chemistry was the principal goal of theorization in this science.[10] Similar claims were made by other chemists in the 1930s and 1940s.[11]

Although the conception of organization has not taken such an important place in chemistry as the conception of structure, chemists have come close to this conception, describing the unity of the mechanism of a reaction and its kinetic factors. Thus, S. Z. Roginsky, a Soviet specialist in catalysis, has put forward the conception of the internal cybernetics of a chemical system.[12]

The third triangle, "organization – behaviour – self-organization" symbolizes the development of modern theories of self-organization in chemistry. Following this triangular scheme, formation of the theories of self-organization was associated with the problem of explanation of the behaviour of chemical systems, that is, of the evolution of their functions in time.

Prigogine's nonlinear, nonequilibrium thermodynamics provides an interesting example of the theory of self-organization. This theory was formed in the course of attacks on the explanation of complex behaviour which was discovered for some chemical systems. As Prigogine writes, the Belousov–Zhabotinsky reaction, showing the concentrational oscillations, has provided the main experimental basis for his thermodynamics[13]. To be sure, an important step in the explanation of this reaction was achieved when the mechanism was described. However, the general conditions of the Belousov–Zhabotinsky reaction were formulated in nonlinear, nonequilibrium thermodynamics by introducing the concepts "stability", "self-oscillations", "dissipative structures", etc.

It should be emphasized that there is a theory of self-organization in chemistry which, in contrast to Prigogine's theory, can be called the chemical one. By this I mean the theory of evolution of catalytic systems proposed by A. P. Rudenko, a Soviet chemist, attempting to explain the changes in the catalytic activity of the systems of heterogenous catalysis.[14] If Prigogine's theory has been constructed according to the standards of mathematical physics, then Rudenko's theory is a chemical theory based on physical ideas (in accord with my classification that theory is associated with the first stage of the physicalization of chemistry).

Following the above scheme of the three triangles, one can observe how theoretical chemistry ran away from the physicalization of it. When quantum mechanics was formulated and the basis for the physical explanation of the structure of molecules was created, a paradigm change in chemistry took place: Chemists concentrated mainly on the research of the chemical process. Although physics succeeded in explaining the elementary steps of chemical reactions, it was a less successful in attacking the problem of the interconnections of elementary steps in the structure of a reaction.

Now, the development of nonlinear, nonequilibrium thermodynamics and the nonlinear theory of dynamic systems seems to provide the physical foundations for an explanation of the complex behaviour of chemical systems. Nevertheless, the special chemical theories of self-organization not derived from physics have been formulated for the explanation of some catalytic phenomena.

More arguments in favour of my two-dimensional scheme of the history of chemistry are given in my book.[15] Nevertheless, my scheme should be considered as a hypothesis which helps in understanding historical facts.

*Institute of History of Science and Technology*
*Russian Academy of Sciences*

## NOTES

[1] G. Holton. *The Advancement of Science, and its Burdens.* (Cambridge: Cambridge University Press, 1986), pp. 6–27.
[2] B. M. Kedrov. *On the Marxist History of Natural Sciences.* (Moscow, 1968) (in Russian).
[3] B. M. Kedrov. *The Three Aspects of Atomism.* Vol. 2 (Moscow, 1969), p. 294 (in Russian).
[4] G. V. Bykov. "The Prestructural Theories of Organic Chemistry in Russia", Proceedings of the Institute of History of Science and Technology (Moscow, 1958) (in Russian).
[5] H. Metzger. *Newton, Stahl, Boerhaave et la doctrine chemique* (Paris, 1930); V. P. Zubov. *The Development of Atomism till the Beginning of XIX* (Moscow, 1965) (in Russian).
[6] C. Schorlemmer. *The Rise and Development of Organic Chemistry* (London, New-York: Macmillan, 1894), Russian translation: Moscow, 1937, p. 46.
[7] B. N. Menshutkin. *Chemistry and Ways of its Development.* (Moscow, 1937) (in Russian).
[8] R. Feynman, R. Leyton, and M. Sands. *Faynman Lectures on Physics.* Vol. 2 (Reading Mass., 1964), p. 31–I.
[9] G. Nicolis and I. Prigogine. *Selforganization in Non-Equilibrium Systems* (New York, 1977), p. 49.
[10] N. N. Semenov. *The Chain Reactions* (Moscow, 1934), p. 19 (in Russian).
[11] L. P. Hammett. *Physical Organic Chemistry* (New York, London: McGraw-Hill, 1940), CH. I.
[12] S. Z. Roginsky. "An Introduction", *The Problems of Kinetics and Catalysis.* Vol. 14 (Moscow, 1970), p. 8 (in Russian).
[13] I. Prigogine. *From Being to Becoming: Time and Complexity in the Physical Sciences* (W. H. Freeman and Co, 1980), p. 130.
[14] A. P. Rudenko. *The Theory of the Self-development of Open Catalytic Systems* (Moscow, 1969) (in Russian).
[15] A. A. Pechenkin. *The Interconnection between Physics and Chemistry: Philosophical Analysis* (Moscow, 1986) (in Russian).

ANNA KOSTOULA

# THE PROBLEM OF METHOD IN THE STUDY OF THE INFLUENCE A PHILOSOPHY HAS ON SCIENTIFIC PRACTICE. THE CASE OF THERMOELECTRICITY

Developments in the historiography of science force the historian, interested in the influence a philosophy has on science, to make clear in what scientific aspect he expects to find this effect. The treatment of the issue up to now implied that the decision examine the specific philosophy – an organized way of thinking – as a sensitizing agent of the scientist's mind which enabled him to perceive and recognize a phenomenon. The analysis implied that the mechanism producing the phenomenon, the laboratory habits (theoretical and technological), remained unchanged.

In the case of Naturphilosophie, we have to do with a philosophy that explicitly states demands for "another" scientific practice, thus forcing the historian to examine the effects, if any, on the whole of the scientific behavior. However, Naturphilosophie does not offer a well-organized corpus of rules for science. Some of its representatives busied themselves with the theoretic question of scientific practice (Hegel, Fichte, Schelling, etc.), answering the question of how it should be done, while others undertook translating this into everyday practice. Oersted and Goethe did both. In Naturphilosophie we have to do with a movement of a wide spectrum and not with a certain group of people that abide by a specific philosophical structure.

Naturally, the first task posed by the circumstances is to determine which part of the spectrum the scientist in question represented. The customary route is to search the writings of the scientist for expressions of his philosophical preferences. If no general statements concerning his philosophy can be found but only biographical information about personal connections with the circle or, for that matter, vaguely expressed sympathies, the historian then looks at the scientific work for signs of philosophical preferences. In the case of Naturphilosophie, this would mean references to bipolar situations or the unity of phenomena.

One, at first, thinks that it is easier and more straightforward to study a scientist who has verbalized his philosophical preferences, because the historian then has the personal *Vorbild* of the scientist before him. But this is the problem. It is a *Vorbild*, an aspiring image.



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science,* 379-384.

The person that theorizes his preferences might build a theoretical system expressing at length, and wit exactness and style his wishes about the way things should or should not be done. If he dares to use them in practice (or if the conditions are such that allow him to do so), is another issue. And when he explicitly uses them as norms for his scientific behavior we are faced with the question as to what degree his method succeeds.

Therefore, in both cases – a philosophically expressive scientist and the "practioner" of the philosophy – the historian has the "actual" scientific work to look at, which very often means the presentation of the scientist's laboratory work.

In forming the spectrum of the philosophical current one keeps in mind (explicitly or inexplicitly) its ideal, the *Vorbild*: the general ideal and the one each scientist has formed for himself.

The "private" *Vorbild* is important because it guides the researcher in examining the reasons that forced the scientist in question to deviate from the more general *Vorbild* presented by the philosophy. It (the "private" *Vorbild*) facilitates the historian's effort to specify the deviations from the norm. Thus, in assessing the influence a philosophy has on scientific practice, the historian needs the general *Vorbild* of the philosophy as well as the individual scientist's version of it.

For Naturphilosophy, the reconstruction of the general ideal, the *Vorbild*, is difficult. However, widely accepted characteristics can be gathered. I used Goethe as a source because Goethe is a good receiver and transmitter of the hidden and unexpressed qualities of his time and class (*Bildungsburgertum*) in both theoretical and working examples. Goethe was the mouthpiece of his time. He possessed the gift to put into words even the moments of silence of his class. He organized around him, using the Theory of Colors as a center, many people concerned with the creation of a Naturphilosophical model in science. It was not the only group, but one whose activities were well recorded. The nature of other acknowledged cooperations are not as well known. (Oersted-Ritter or Ritter-Schelling, with the help of Novalis.)

It is not clear if Goethe's group showed more stability or if Goethe's tendency to record extensively makes it appear as such. Maybe his connection with the Weimar Court (he was *Geheimrat*) made him an important point of connection for the literary and scientific circles. His position in the court was viewed by the *Sturm-und-Drang* followers – the precedents of the Romantics – as an umbrella for the movement.

To determine the aspiring image of each scientist in question, an examination of his general, philosophically minded statements suffices. To determine the form it took in practice, I turned to the scientific writings. I applied a scholastic type of reading to the scientific publications.

Many readings were applied to the text under consideration. It was first read through the author's eyes, i.e., through the promises he makes. In the second reading, an attempt was made to determine the guiding questions of each set of experiments. During the final reading, the experimental results were examined, with the question in mind if they actually responded to the aspiring principles and to what degree.

My choice of persons was guided by two scientific events: the discovery of Thermoelectricity and of Electromagnetism. Thermoelectricity's discovery and the first laboratory studies of the phenomenon were the focal point of my interest. Both discoveries, especially Electromagnetism, presented not only a new line of research but also an example – a case – in which the main principle of Naturphilosophie – the unity of natural phenomena – found its laboratory expression. Unlike Goethe's color experiments, the discovery of Electromagnetism brought about an experimental setting well-accepted by the scientific community.

Seebeck, the man who set up the laboratory arrangement that made the thermoelectric effect obvious, and who recognized it as such, took a central position in my study.

The discovery of Thermoelectricity was made and reported by Seebeck in December, 1821, as part of an extensive study concerning the nature of Electromagnetism discovered half a year earlier by Oersted. Thus, an understanding of the conditions of the discovery meant an examination of the author's way of treating Oersted's newly discovered experimental setting and Oersted's theories.[1]

The phenomenon of thermoelectricity presented a laboratory connection for the major categories of phenomena then known: electricity, magnetism, heat, and the possibility of producing chemical phenomena. Therefore, it was of importance for the Naturphilosopher.

The understanding of Oersted's treatment of Electromagnetism was also dictated by the following. Oersted, a Naturphilosopher par excellence for the historians of the Romantic Movement, was the first person to respond to Seebeck's discovery. He carried the news to France and there made a joint study with Fourier, which was published in 1823. Thus, an understanding of this treatment of Thermoelectricity meant an understanding of the wider program in which the discovery of thermoelectricity was placed.

The experimental setting in which the Thermoelectric effect was demonstrated was very important for the Romantics. It simultaneously brought to light under the scrutiny of the laboratory, the main categories of phenomena (heat, electricity, magnetism, and chemical phenomena). Thermoelectricity symbolized the unity, an important aspect of all versions of naturphilosophical scientific programs.

The determination of the program each scientist had in mind was inferred from his first responses to Electromagnetism – the first loud proof of the utility of Naturphilosophical principles in scientific practice.

The discovery of Electromagnetism gave the Naturphilosophically inclined scientist the right to openly work his method into mainstream science. Goethe, on the other hand, had built another branch outside the conventional optical laboratory tradition with his Theory of Colors and, thus, his work was excluded altogether from the mainstream scientific practice of the time. It should be remembered that these events took place in a climate in which "right", i.e., accepted scientific practice, used experimentation as the only source of ideas. The public image of scientific activity of the time was composed of a dogmatic application of the Newtonian credo 'Hypothesis non fugo', with some Baconian elements.

The use of ideas – hypothesis – was a major issue for the Romantics, although it was not their only concern, and it was not exclusively Romantic. Whewell, clearly not a Naturphilosopher, also expresses himself on this issue.

As indicators of the climate of the time, we should remind ourselves of the strong reactions to Goethe's theory of Color and the general contempt provoked by Ritter's failure to produce the laboratory proof of the connection between electricity and magnetism.[2]

Seebeck presented the discovery of Thermoelectricity as a second part of a study on Electromagnetism. The results and conclusions were read in the Akademie der Wissenschaften in Berlin and subsequently printed in the *Abhandlungen der Akademie*.

Seebeck's presentation is unique in that it structures a working example of a scientific study based on the principles of Naturphilosophie without once referring overtly to the principles of that philosophy. At first, the articles can be read as typical, i.e., factual reports. A more careful reading, a reading that constantly looks for the hidden reasons of each experiment and carefully looks at their order of presentation, uncovers the philosophical intent. The only theoretical proposition Seebeck ventures is when he discusses the various explanations of Oersted's discovery. He presents his own opinion not as the result of philosophical principles, though it clearly is, but as the pure, deeper voice of the experimental work on the subject up to that time, which he is able to hear. Seebeck proposes that the effect of the current on the magnetic needle and the related phenomena are the result of something very fundamental from which everything starts: The Imbalance of Action – *Ungleichkeit der Action.* This can be caused by different conditions of matter in the plates, such as temperature.

After this, the experimental study itself and the order of presentation is such that it proves the functioning of this general cause in a specific case. Seebeck leads the reader to see differences in temperature as one kind of imbalance.

Another example of the way Seebeck uses his philosophical principles is his insistence on referring to Thermoelectricity as "Thermomagnetismus", which brings to mind Hegel's phenomenology. With the term "Thermomagnetismus" Seebeck emphasizes the phenomenology of scientific events.

It is interesting to see how Oersted, a Naturphilosopher and a theoretician to the degree that he was accused by his contemporaries for being too much of a philosopher, presented his discovery, the significance of which – for his philosophy and for the larger scientific community – he is very aware.[3]

There are differences in style between Oersted's and Seebeck's presentations. Reading Oersted's article, the reader has the impression that the author humbly pulls the curtain of the public stage in which he has rolled his laboratory setting which produces the effects he claims he has observed. Only after he has carefully described the characteristics of the magnetic force (all are experimental results) does he propose an explanation – a theoretical structure – in which one can recognize his philosophical preferences. But, even then, the proposal is of a very tentative nature, offered to the wider community for testing.

In comparison, Seebeck offers his proposals with arrogance, as the truth he has read (and was able to read) in nature. It is the truth extracted by him from the experiments. He uses the experiments like a priest. He is not only the producer of facts but also the revelator, the person that makes the facts talk. Seebeck's behavior is aristocratic and comes closer to the Naturphilosophical ideal (*Vorbild*).

My curiosity for the rate of this marvellously composed structure, (i.e. the effect it had on other researchers) led me to similar readings examinations of the works on Thermoelectricity that followed Seebeck's. The researchers accept Seebeck's experimental work, but not in the light of his philosophy, although some do show sympathy for the use of ideas in science.

A close reading of these publications looks carefully for the question that guides the experiments. The question which the results actually answer, which does not necessarily coincide with the one claimed by the scientist. This type of reading of the presented experimental procedures proved that the experiments obeyed strictly quantitative questions – questions that aim at the control of the phenomenon. They facilitate its manipulation, but do not lead to the causes or to the one cause, looked for by the Romantics.

Seebeck's experimental findings are treated as an integral part of this structure, without any respect for the philosophical implications or sense of difficulty from the process of adaptation. To have a sense of the meaning of this, I want to again remind you of the vehemence with which Goethe's color experiments were rejected.

*Aristotle University of Thessaloniki, Greece*

NOTES

[1] T. J. Seebeck, Ueber den Magnetismus der galvanischen Kette. In: Ostwalds Klassiker, hrg. von A. J. v. Oettingen, Leipzig 1895.
[2] Keld Nielsen. Another kind of light: The work of T. J. Seebeck and his collaboration with Goethe. Part I. In: Hist. Studies in the Physical and Biological Sciences, 1989, vol. 20, part. I.
[3] H. Kristine Meyer and H. C. Oersted, Naturvidenskabelige Skrifter...3 vols (Kopenhagen, 1920), v. 1.

STUART WALKER STRICKLAND

# REOPENING THE TEXTS OF ROMANTIC SCIENCE: THE LANGUAGE OF EXPERIENCE IN J. W. RITTER'S *BEWEIS*

I grew up nursing an illusion: that science was a humanity. Maybe this deliberate naïveté explains how breathlessly I fell in with romantic science. But that is a story with many roots. In any case, it is not a story of retreat, nor an antiquarian withdrawal from the problems of the present world. My turn to romanticism has always born the mark of its original motivation: to goad my contrary commitments to literature and to science – the rivalries I have allowed to grow, even fostered, within me – into a dialogue, to force them to confront each other and to speak.

In my recent work I have been following the career of the romantic physicist Johann Wilhelm Ritter (1776–1810).[1] Ritter himself lived in two worlds. To Friedrich Schlegel and the literary community of Jena, he was the embodiment of romantic science. To Alexander von Humboldt and other leading figures of German science, he was a careful and promising young experimenter, worthy of the highest praise. But, as the intellectual circles around Jena began to disintegrate, as the Napoleonic Wars disrupted German academic life, and particularly as empirical researchers sought to define their work in opposition to speculative idealism and romantic *Naturphilosophe*, these two worlds diverged and Ritter tumbled into the crevasse which opened up between them.

Frustrated in his attempts to secure a position at the University of Jena, Ritter accepted an offer to join the Bavarian Academy of Sciences and moved to Munich in the summer of 1805. One of his first scientific projects upon arriving in Munich was to seek the Academy's support for a pilgrimage to the Italian Tyrol. There he would investigate reports of a young villager named Francesco Campetti. Campetti, it had been claimed, possessed a special sensitivity enabling him to discover buried pieces of metal and underground sources of water. Ritter saw here the chance to continue his earlier research into the affinities between the apparently dead and living domains of nature. Following his enormously successful work with animal electricity, he expected his investigations of Campetti to open up a fertile new field along the borders between physiology and physics. Furthermore, Ritter presented this research as a continuation of



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 385-396.

the meticulous experimental methods upon which he had built his scientific reputation.

Bursting with enthusiasm over his most recent discovery, Ritter rushed Campetti back to Munich, asked the Academy to appoint a commission before which he could demonstrate his results, and began publishing a new journal called *Siderismus*. Although his experiments attracted widespread public curiosity, many of his new colleagues in Munich and the editors of the most influential scientific journals were suspicious. In their view, Ritter had overstepped the bounds of legitimate science. The controversy which ensued stigmatized Ritter, left him personally and intellectually isolated, and came to signify the worst excesses of romantic *Naturphilosophie*.

The farther I have pursued Ritter's fate, the more I have come to appreciate how intimately his life, his career, his misfortunes are bound up with the stories which separate the Enlightenment from modernist conceptions of science. These are the stories of the collapse of a tradition of grand, totalizing systems of nature; of the expulsion of figurative language from scientific writing; of a strange alliance between a philosophically sophisticated elite and popular folk knowledge; and of a burgeoning class of professional researchers who rose up to break this alliance by deploying a powerful new technology: the division of intellectual labor. These are the stories around which I have come to understand romantic *Naturphilosophie*.

The work I will be presenting here is part of a larger study of the rhetoric of self-experimentation during the romantic period. My plan will be to focus on the way in which appeals to experience as an anonymous authority have worked to constrain language choices in the representation of subjective experience. Beginning with the galvanic self-experiments of Ritter, Alexander von Humboldt and Christoph Pfaff, I will examine how experiences whose persuasive force was rooted locally in the particular integrity of the subject/author were reconciled with a prevailing premium placed on disinterested appeals to a kind of universal experience which hid the author from view. Around these representations of galvanic self-experiments I will construct a larger story about authority and experience, a story which begins with Lavater's diaries of self-observation and continues through Helmholtz's attempts to tame subjectivity in order to create a philosophically self-conscious science of perception. Ultimately, I hope to work out how the tensions within these texts have reshaped subsequent appeals to experience and how they continue to condition our anxiety over the rhetorical construction of the particular, privileged yet limited perspective of the subject.

For the moment, however, I will limit myself to a discussion of ambiguities in Ritter's work and to the tensions which emerge in his rhetoric as he turns from experiments with dissected frogs to those suffered by his own body. To appreciate the significance of Ritter's work for studying the language of subjective experience it may be helpful to develop a sense of the deep ambivalence in contemporary estimations of the value of experience and the dangers of speculation.

Writing to Alessandro Volta in 1807, Ludwig Gilbert, editor of the *Annalen der Physik*, a journal to which Ritter had hitherto been among the most prolific contributors, assessed the state of physics in Germany. Gilbert described the havoc wrecked by military defeat and by a wave of speculation which threatened to erode the scientific community:

> As for the condition of physics in Germany, it is regrettable that the unfortunate war appears to be crushing its blossom precisely in those regions where the natural sciences were pursued with the most zeal and in the true spirit of research. That is to say, in the Prussian states, Saxony and Hannover. In this there is a great difference between northern and southern Germany. In the latter a spirit of distasteful speculation and mysticism has spread through physics, taking hold even of men who would otherwise honestly and skilfully experiment. Munich in particular (since the University of Würzburg is as good as gone) now appears to be the seat of this fanaticism. Ritter, who always leaned somewhat in that direction, now seems to have surrendered himself completely.[2]

For Gilbert, the unjustly disparate effects of the war reflect a spiritual chasm separating northern from southern Germany. This, in turn, provides a geographical correlate to support the dichotomy he advances between the honest experimenter and the mystical fanatic, between the spirit of research and the spirit of speculation. Ritter is cast here both as an honest victim swept away by historical circumstances and as a kind of traitor, an intellectual turncoat whose move from Jena to Munich confirms his spiritual treason.

No motif is more inveterate in the historiography of *Naturphilosophie* than the binary opposition of speculation and empiricism. Along the path of empiricism we find the security of the inductive method, careful observation and the lessons of experience. But should we stray into the domain of speculation, we risk entanglement with philosophical deducers, hypothesis builders and the coquettish imagination. If Gilbert's language encourages us to weave our historiography upon this dichotomous loom, it also exposes a loose thread. By acknowledging that Ritter had always been inclined toward speculation, Gilbert inadvertently undercuts the mutually exclusive relation upon which his portrayal of Ritter's fall depends. He invites us to look for continuity between Ritter's earliest investigations of galvanic circuits and his later sidereal research. He tempts us to reexamine the Ritter whom he had described just a few years before as "a highly scrupulous and precise observer,"[3] and to reopen Ritter's first book, his *Beweis, daß ein beständiger Galvanismus den Lebensproceß in dem Thierreich begleite.*

Published in 1798, the *Beweis* is concerned with the conditions under which an effective galvanic circuit may be constructed.[4] Luigi Galvani's announcement, seven years earlier, that he had discovered a new kind of electricity – his claim that when the exposed nerve and muscle of a frog's leg were bridged by a metal arc the muscle would contract – aroused a flurry of attempts to reproduce and explain this apparently new phenomenon.[5] Ritter's *Beweis* describes a series of experiments in which assorted metals and liquids are brought into contact with nerves and muscles in various arrangements. Circuits are classified as either effective or ineffective, depending on whether or mot they elicit a muscle contraction.

A simple example, and the one with which Ritter begins the *Beweis*, involves a piece of silver, a piece of zinc, a muscle removed from the lower extremity of a freshly killed frog, and the nerve which is organically bound to this muscle. When the silver is placed in contact with the nerve and the zinc is placed in contact with the muscle and the two pieces of metal are then joined together, a violent contraction is supposed to result. As long as the circuit is kept closed the muscle should remain at rest. But when the metals are separated and the circuit is broken a second, weaker contraction ought to occur. If the order of the metals is reversed – if the zinc is connected to the nerve and the silver to the muscle – then the weaker contraction should accompany the joining of the two metals and the stronger one is to be expected upon their separation.

The *Beweis* promotes a cluster of related claims, but its central concern, the proposition around which almost all of the experiments are assembled, involves the period during which the galvanic circuit seems to be at rest. Ritter argues that even though no motion can be perceived and the muscle appears to "sink into a deep sleep,"[6] the phenomenon which he calls galvanism nevertheless continues as long as the circuit remains closed. In a lecture delivered the previous autumn to the Jena *Naturforschende Gesellschaft*, Ritter distinguished between the motion (*Bewegung*) and the sensitivity (*Empfindung*) of the circuit:

Motion is produced only at the instants of closure and separation, whereas sensitivity is induced as long as the circuit remains closed.[7]

Motion is easy to see, easy to describe; but sensitivity is elusive. To support his position Ritter resorts to a host of self-experiments in which his sense organs and mucous membranes stand in for the frog's leg in the galvanic circuit. His descriptions of the sensations he endured after the initial shock of closing himself in the circuit provide the ground for his argument that there is more to galvanism than meets the eye.

In his overt insistence that he is describing only the laws governing galvanic phenomena, in his recurrent appeals to experience as the only legitimate arbiter of conflicting theoretical claims, and in his reluctance to name the underlying cause of the experiences he relates, Ritter secures the high ground of Newtonian rhetoric above the shifting currents of contention between those who would commit galvanism to the Charybdis of bodily fluids and those who would toss it against the Scylla of metallic contact. Indeed, the critique of hypotheses with which Ritter prefaces the *Beweis* even appears to anticipate the language with which his critics would later chide him:

If, then, it does no good to pile hypotheses on top of hypotheses, if a genuine thirst for knowledge cannot be satisfied by a somehow, a perhaps, an it-might-be-possible, an it-could-be, we may only reach this goal by adhering to experience. Only in her hand will we walk happily; but if we forsake her, if we abandon ourselves to the wings of our imagination, we may indeed have pleasant dreams, but awaken all the more unpleasantly.[8]

It would be rash, however, to read Ritter's disavowal of hypotheses as an unambiguous claim to relinquish speculation; nor may we infer from his embrace of experience a submission to naïve empiricism. To appreciate the complexity of the semantic web in which experience, speculation, hypotheses and empiricism implicate one another, we must first recognize that the *Beweis* is a text divided against its own convictions. If Ritter disparages hypotheses, he nevertheless aspires to a single true and complete theory of nature. And if he withdraws behind the authority of experience, he remains unable to escape his own authorial responsibility for rendering and conducting the individual experiences which permeate the text.

The concept of experience was the subject of vigorous debate among Ritter's contemporaries – not only among philosophers, but, as Brigitte Lohff has shown, among physiologists, chemists, and physicists as well. An important distinction drawn within this discourse was that between individual experiences and the process through which these were brought together to form a science of experience.[9] In a letter to his patron, Ernst II., Herzog von Sachsen-Gotha and Altenburg, Ritter seems to acknowledge such a distinction as he explains his own role in arranging the experiences he has gathered:

> Since my departure from Gotha I have indeed made various new investigations. But my main occupation remains to work up the unusual quantity of experiences which I was fortunate to be able to gather in Gotha into general results and to apply these results to the propitious explanation of an almost equally large number of other phenomena. I have yet to completely finish this business, though only its richness has been to blame for its delay.[10]

Experiences must be digested by the investigator of nature; they must be worked up into explanations of phenomena. In the text of the *Beweis*, however, experience takes on a more active role. The investigator exits and experience steps in as the voice of authority. Experience teaches,[11] experience offers examples,[12] experience confirms suppositions,[13] experience speaks for or against theories,[14] and experience substantiates claims.[15] It is not Ritter, but experience who levels "beautiful hypothetical edifices" with a single blow. It is she who dictates that, "As soon as the least true contradiction divulges itself, a theory cannot be the true one and must be abandoned."[16] As experience assumes a more prominent role in the *Beweis*, Ritter's own movements become obscure.

Some clues may emanate, however, from Ritter's introduction to a German edition of Spallanzani's work on phosphorescence. Here he discriminates between the business of the *Philosoph* and that of the *Naturforscher*.[17] The philosopher follows the path of deduction; the investigator of nature follows the path of reduction. The philosopher resolves the whole into its constituent parts. The investigator collects the parts and assembles them into the whole. Though travelling in different directions, the philosopher and the investigator are nevertheless bound to meet so long as each remains faithful to his chosen way.

But ambiguity lurks in this deceptively conspicuous division of labor. Are these two distinct activities or are they complementary aspects of one and the same endeavor? According to Ritter's prescription, the investigator works neither arbitrarily nor according to his own intentions; rather, he follows nature and composes the parts "as the parts themselves submit and bind one another together."[18] This would seem to imply that the investigator works blindly and without anticipating the whole toward which his efforts are leading. Yet Ritter measures a science of experience by its ability to provide explanations, explanations which emerge at the juncture of deduction and reduction.[19]

Rather than directly confronting the difficulties involved in justifying this alliance, rather than resolving the conflict between an assembly of parts which anticipates the whole and one which follows their own natural inclinations, Ritter appeals to the everyday experiences of his readers. The experiences from which explanations should be built, he advises, are "...not rare or to be attained only through wide detours, but, on the contrary, the most everyday..."[20] By sticking to everyday experiences, Ritter facilitates his obligation "to bring [experience] before the eye of the reader as much as possible."[21] In the process, his responsibility for ordering the experiences he has collected is palmed off on convention, on habitual patterns presumably shared by the investigator and the reader. The appeal to everyday experience invokes a common frame of reference which masks the associations constructed in the text itself and thus postpones the need to justify the author's own agency in this process of construction.

But if we return to the *Beweis*, we see how problematic this appeal to everyday experience must be. Experience is preferable to hypothesis building just because it promises to tell Ritter something he does not already know, to confound conventionally-based expectations. Even though galvanism and electricity appear to be two different phenomena, experience reveals their identity. Even though a frog's leg appears to be at rest while the galvanic circuit is closed, the *Beweis* purports to prove that galvanic action is continuous.

In the self-experiments through which Ritter seeks to establish the latter claim the tension between experience's role as an authority and the subjectivity of the author's personal experience is laid bare. Ritter's shift from accounts of frog experiments to descriptions of experiments performed on his own body involves a significant shift in narrative perspective. These experiments are recounted in the first person. Ritter assures his readers of results which he himself has experienced. The experimental conditions of these self-experiments are also described in far greater detail than any others in the text, ostensibly to facilitate repetition, but perhaps equally to create the illusion of having oneself witnessed the experiments described.[22] The *Beweis* provides what James Paradis, in his study of Boyle's experimental essay, has called the "curious literary manufacture of phenomenological experience".[23] Such descriptive detail is necessary because these self-experiments are in fact far removed from the ordinary experiences of everyday life. They are hidden in the dark recesses of the experimentor's sensations of pain and require special artifices to bring them "before the eye of the reader". They require the text to shift its weight from the impersonal authority of experience to the personal assurance that the scenes recounted are those which Ritter himself has experienced. In the end, Ritter is forced to acknowledge that the phenomena he describes are "only subjectively perceivable".[24] But his defence, that all experience is ultimately subjective,[25] is also a concession.

Although it announces itself as a proof and is punctuated by laws, the *Beweis* repeatedly reminds us that it is but a fragmentary artifact – disjointed and incomplete. Ritter introduces it with the admission that he has fallen short in his "striving for perfection" and apologizes for offering such an unfinished work:

> With much cause do I ask that the whole be regarded more as a very imperfect *attempt* to arrange my subject systematically, that everywhere more consideration be given to its contents than to its form, though the former also loses something in the latter.[26]

There is more to this remark than a prudent modesty which forestalls criticism. If we disregard Ritter's attempt to divert our attention from the structure of the text, we may become aware of a deep tension between closure and openness.

One of the most striking structural characteristics of the *Beweis* is its apparent inability to draw itself to a close. Although the proof seems to reach its conclusion with a neat circular repetition of the book's title, the sense of closure which this repetition suggests is weakened by a subsequent section which declares that the results cannot be allowed to stand alone without consideration of their practical implications. What then begins as a modest suggestion for the application of galvanism in medicine, quickly overflows into a discussion of animal magnetism, the relations between the body and the soul and, in the next two sections. an elaboration of previously implicit conjectures about the galvanic basis of life and the identity of galvanism, electricity and chemistry. The final section, the *Beschluß* or "resolution," breaks this discussion off abruptly and concludes by pronouncing the open-endedness of Ritter's objective of finding, through the further investigation of galvanic phenomena, a "general natural law" which would govern the "physiology of the universe":

> The imperfection of this writing itself can prove to you that it is not the work of an individual; only the united force of many would be able to complete it.[27]

It is hardly surprising that, in the face of such vast ambition, Ritter's work should appear incomplete, nor that he should seek to enlist potential colleagues in an common project. Paradoxically, however, far from being speculative appendages to an otherwise purely inductive investigation, Ritter's aspirations to articulate an all-encompassing system and his concomitant belief in the fundamental nature of galvanic phenomena are manifest throughout the *Beweis* and lay at the very heart of his critique of hypotheses.

Ritter's distaste for hypotheses is rooted in a particular conception of history, a conception which also informs his efforts to move from individual experiences to a comprehensive theory of nature. The confusing myriad of competing theories of galvanism led to numerous attempts during the first decade of the nineteenth century to construct coherent histories of various scientific disciplines. From these histories one disturbing fact emerged: previous generations of skilled and clever researchers had been wrong! The lesson of history was caution. "If it is an evil not to know a truth," Ritter quotes Fontana, "it is a still greater evil to believe an error."[28] The project of writing history would become, as in the case of Hans Christian Orsted's 1807 "Considerations of the History of Chemistry," an effort to recover the sense in which apparent errors had nevertheless contributed to progressive development. But the consolation that even false theories "must have something in common with the true one,"[29] was insufficient either to assuage the anxiety evoked by the historical perspective or to prevent hypotheses from becoming the scapegoat of error. As Orsted wrote, describing the false path taken by the phlogiston theorists, "The hypothesis made them blind to that which nature showed them."[30] Or, to use Ritter's favorite metaphor, hypotheses interrupt the interrogation of nature. "Indeed she often tells us more than we first believed we would hear;" he informs his readers, "and for such cases the investigator of nature must make the greatest possible effort to lend her an obliging ear."[31]

For Ritter, at least, this historical lesson is learned not just from a chronicle of events, but from a history told as a cyclical narrative of separation and return. As is evident from his lecture on *Physics as Art. An Attempt to Divine the Tendency of Physics from Its History*, this is a story in which science will reunite humanity with all of the apparently disparate forces of nature.[32] Hypotheses are suspect because they may obstruct the historical destiny of science. But the plea to keep the conversation open in order to learn something unexpected certainly seems disingenuous if

Ritter already knows the ultimate conclusion towards which the history of natural science is heading.

Consider for example, the proposition that galvanism, electricity and chemical reactions are all manifestations of the same fundamental phenomenon. A substantial portion of part one of the *Beweis* is devoted to demonstrating that all bodies which serve as conductors in galvanic experiments also conduct electricity and that, with a few exceptions, non-conductors of galvanism do not conduct electricity. Significantly, however, these exceptions do not lead Ritter to doubt his position. Instead of stating his case before the court of experience, Ritter refrains from formulating this position *as* a hypothesis until near the end of part two when he cautiously leaves the issue open to further investigations; but it has nevertheless already served as an implicit focal point around which many of the experiments in part one were arranged. The strategy of keeping the conversation open is thus applied as much to protect tacit hypotheses as to preclude competing ones.

The desire to learn from nature something we do not already know, to be led by experience rather than to lead experience ourselves, drives the language of the *Beweis*. But in order to show that things are more than they seem to be, that something which lies beyond everyday experience holds the world together, Ritter embeds the fragments of individual experience in a Mosaic narrative of departure and return – a structure within which the destiny of physics and the history of humanity converge.

I should say in closing that I agree completely with Professor Baltas' view that there is no way of simply "stepping out" of our conceptual framework.[33] But I do believe that the very process of doing history can alter that framework. I prefer to see it not so much as a skin in which we are stuck as an old sweater which we have inherited and which has shrunk to fit us. This sweater is bound to stretch and even in some places to rip if we try to pull it down over an oddly shaped body like romantic *Naturphilosophie*. The conceptual framework we have before we try to understand is something altogether different from the one we will have after we have tried to understand. I believe, perhaps naïvely, that it is possible to learn something from history.

My aim is decidedly *not* to present a closed and complete reconstruction – rational, sociological or otherwise – of the romantic philosophy of science. But I do believe that the story of *Naturphilosophie* has rested for too long on the easy distinction between speculation and empiricism. We can no longer afford to turn a deaf ear on the competing commitments hidden beneath the still surface of Ritter's critique of hypotheses and his espousal of experience. Only by pursuing these contradictions, only by holding them together and making them speak, may keep our historical conversation open; only thus may we have some chance of hearing more than the echoes of our own questioning voices.

*Northwestern University*

NOTES

[1] "Circumscribing Science: Johann Wilhelm Ritter and the Physics of Sidereal Man" (Ph. D. diss., Harvard University, 1992).
[2] *Epistolario di Alessandro Volta. Edizione Nazionale. Volume Quinto, 1805–1827* (Nicola Zanichelli Editore: Bologna, 1955), 110. All translations are my own.
[3] Gilbert to Volta, 18 February 1803. *Epistolario di Alessandro Volta. Edizione Nazionale. Volume Quarto, 1800–1805* (Nicola Zanichelli Editore: Bologna, 1953), 258.
[4] Johann Wilhelm Ritter, *Beweis, daß ein beständiger Galvanismus den Lebensproceß in dem Thierreich begleite. Nebst neuen Versuchen und Bemerkungen über den Galvanismus* (Weimar: Industrie-Comptoir, 1798). For alternative readings of Ritter's *Beweis*, see Hinrich Knittermeyer, *Schelling und die romantische Schule* (Munich: Ernst Reinhardt, 1929), 142–151; Barry Gower, "Speculation in Physics: The History and Practice of *Naturphilosophie*," *Studies in History and Philosophy of Science* 3 (1973): 327–339; and Walter D. Wetzels, "Johann Wilhelm Ritter: Romantic Physics in Germany," in *Romanticism and the Sciences*, ed. Andrew Cunningham and Nicholas Jardine (Cambridge: Cambridge University Press, 1990), 201–203.
[5] Luigi Galvani, *De viribus electricitatis in motu musculari commentarius* (Bologna, 1791). For a detailed examination of the reception of Galvani's discovery among German-speaking researchers, see Maria Trumpler, "Questioning Nature: Experimental Investigations of Animal Electricity in Germany, 1791–1810" (Ph. D. diss., Yale University, 1992).
[6] *Beweis*, 121.
[7] Ritter, "Über den Galvanismus; einige Resultate aus den bisherigen Untersuchungen darüber, und als endliches: die Entdeckung eines in der ganzen lebenden und todten Natur sehr thätigen Princips; – vorgelesen in der Naturforschenden Gesellschaft zu Jena, am 29. October 1797," in *Physisch-Chemische Abhandlungen in chronologischer Folge* (Leipzig: C. H. Reclam, 1806).
[8] *Beweis*, x.
[9] See Brigitte Lohff, *Die Suche nach der Wissenschaftlichkeit der Physiologie in der Zeit der Romantik* (Stuttgart and New York: Gustav Fischer, 1990).
[10] Ritter to Ernst II., Herzog von Sachsen-Gotha, 14 June 1802, in Kurt Poppe, "Johann Wilhelm Ritter und Ernst II., Herzog von Sachsen-Gotha Etc. Zwei unbekannte Briefe aus den Jahren 1802–1803," *Jahrbuch des Freien Deutschen Hochstifts* (1972): 183.
[11] *Beweis*, 28 and 36.
[12] *Ibid.*, 25 ff.
[13] *Ibid.*, 61, 80, and 161.

[14] *Ibid.*, 34.
[15] *Ibid.*, 44.
[16] *Ibid.*, ix.
[17] *Darstellung der neuern Untersuchungen über das Leuchten des Phosphors im Stickstoffgas u.s.w. und der endlichen Resultate daraus für die chemischen Theorie* (Jena: Frommann, 1800).
[18] *Ibid.*, iii.
[19] *Ibid.*, vi–vii.
[20] *Ibid.*, x.
[21] *Ibid.*, 13.
[22] *Beweis*, 85–86.
[23] James Paradis. "Montaigne, Boyle, and the Essay of Experience," in *One Culture. Essays in Science and Literature*, ed. George levine (Madison, WI: The University of Wisconsin Press, 1987), 76.
[24] *Beweis*, 94.
[25] *Ibid.*, 105.
[26] *Ibid.*, xii.
[27] *Ibid.*, 174.
[28] *Ibid.*, viii.
[29] *Ibid.*
[30] Hans Christian Orsted, "Betrachtungen über die Geschichte der Chemie," *Journal für die Chemie und Physik* 3 (1807): 209.
[31] *Beweis*, vii.
[32] Ritter, *Die Physik als Kunst. Ein Versuch, die Tendenz der Physik aus ihrer Geschichte zu deuten* (Munich: Joseph Lindauer, 1806).
[33] A. Baltas, "On the Harmful Effects of Excessive Anti-Whiggism," *Contemporary Trends in the Historiography of Science*, Corfu, 27 May 1991.

YIORGOS N. VLAHAKIS*

# PROBLEMS AND METHODOLOGY OF EXPLORING THE SCIENTIFIC THOUGHT DURING THE GREEK ENLIGHTENMENT (1750–1821)

The history of science like the history of all human ideas, is a history of irresponsible dreams...
Karl Popper, *Conjectures and Refutations*, 1963, p. 216.

## INTRODUCTION

Besides the serious, methodological problems a historian of science usually faces, there are some which are by-products of the special case under examination. For those who have decided to explore the "virgin" world of the structure and development of scientific thought in Greece during the so-called "Greek Enlightenment" (1750–1821), these many and sometimes peculiar problems present a case for study themselves. The main problem is the very subject under investigation, because the researcher must first convince the scientific community that this period is worthy that science in this area of Greek culture actually existed during the occupation of the Ottoman Empire. The researcher must also be careful to avoid repeating the mistake made concerning science in the Middle Ages. Some of the "ordinary problems which arise in the exploration of this magnificent but still unknown world" will be presented here.

## THE PROBLEMS

A primary problem for those willing to investigate the pre-revolutionary period of science is the problem of approaching the sources of the documents which are of some importance. The cultural centres of the Hellenic civilization during the period of interest were spread all over the Balkan peninsula, which was under the uniform government of the Ottoman empire. Many of the precious documents that could help in the

* This paper is dedicated to Dr. I. Karas, my teacher.



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 397-404.

formation of an objective history of science have vanished or were destroyed, including archives of valuable manuscripts.

The manuscript-problem was intensified by the fact that published scientific books remained at that time very few compared to the total number of works published from 1700–1821. From a total of 2,741 titles (3,079 volumes) only 6.92% were scientific books while 77.5% were religious books.

Furthermore, just a few copies of most were written and they were more sensitive to damage caused over time.

A classic example is that until now, not a single copy of the only Greek translation of Newton's "Principia", done by Nikolaos Zerzoulis in Jasi, has been found. It is a work, the importance of which can easily be understood by everyone.[2]

It must be mentioned that the difficult unpleasant work of locating the existing documents has, for the most part (almost 100%) been completed. That is, thanks to literary historians, who have ploughed the field until recently, concentrating mainly on textual criticism, following a quite anachronical way of presenting their conclusions.

Having broken this conventional "history", it is time for a diachronical history of science to be undertaken, keeping in mind the avoidance of certain whiggish approaches to the subject.

Another important question arising and awaiting an answer from the historian of science who has already located and tried to study the works of the most significant Greek scholars, is the determination of the writers' objectives.

The frequent references, mainly in the books' prefaces,[3] indicate that the main task of the publication of scientific books was not the presentation and support of certain scientific opinions, but a contribution to the vanquishing of ignorance and superstition. This goal was directly connected to the liberal ideas coming mainly from France and the general trend for a revolution of independence against Ottoman Empire. What Greek scholars really tried to achieve was bringing into the limelight of everyday life the applications of sciences in general and particularly physics.

As most Greek scholars had no scientific specialization, we have not benefited much from their scientific ideas. But a handful of them expressed original theories or their works may be considered important for the construction of a special scientific terminology in Greek, which was absent until that time.[4]

However, we are able to clearly understand the climate of these years and Greek scholars' efforts to organize their work on an educational basis.

In this way, Greek scholars had as a first priority the conveyance of the appropriate knowledge from Europe, which on this occasion forms a parallel coexistence to Aristotle's and Newton's scientific schemes. This was owed to powerful props. Even in the last years of its decline, Aristotelianism prevailed over the thought of Greek (and other) scholars.

It is an unquestionable fact that in Greek areas during the last few decades of the 18th century, two ideologically different groups of scholars formed. Groups that have influenced on almost the same degree for a certain time the climate of the era.

This coexistence can be compared to the turbulent flow of two different fluids that forms for a few moments an illusion of a mixture.

Eventually, there were supporters of extreme positions among the members of these groups. The extremists who belonged to the so-called "conservative" group expressed almost ridiculus ideas, from a scientific point of view, in their yearning to support the Aristotelian scientific concept, while the members of the other group rejected Aristotelianism in its entirety and came to severe contradictions with the official Orthodox Church.[5]

As a conclusion on this subject, one could support the idea that it is not a matter of pure luck that strong supporters of the newtonian concept refer many times in their works to Aristotelian ideas, although in many cases unconscientiously. An analogous behavior can be distinguished in the West during the first years of the scientific revolution.

Most of the works of Greek scientists excluding some exceptions during the last years of the movement, were translations compiled, usually, from more than one European scientific book. In these compilations only the parts describing experiments, basic laws and theorems were kept; usage of difficult mathematics and algebra was avoided. The supported theories were chosen on a philosophical basis, taking in account that they had to be suitable and anti-superstitious. Greek scholars were rationalists in the way they tried to develop it.

Experiments were generally a tool to illustrate the benefit of the new "magic" spirit of physics against the classical demonological interpretation of nature. As it happened during the early 18th century in central Europe, Greek scholars tried to perform their experiments in public. But these experiments – where and when the appropriate instruments were available, which was rarely the case – were mainly experiments of demonstration; the verification of a certain law or theorem remained a secondary task.

Using a very wisely chosen criterion serving the tasks of the movement, Greek scholars translated books not of the highest level, like the original essays describing the new theories and mathematical developments (there are naturally a few exceptions), but books of educational character, written by Western scientists who wished to spread and popularize the Newtonian concept of scientific construction. This choice was very significant as it could ensure a considerably high degree of students participating in the lectures given by the scholars. The very low mathematical and physical background of the majority of Greek students and the lack of high-level institutes and universities should be mentioned.

Among the European scientists whose books were used, of the highest importance were the French physicist, Abbé Nollet, the German, F. A. C. Grenn, the Dutch physicist and experimentalist, Peter van Musschenbroek, and the English physicist Benjamin Martin, better known as an instrument maker. Martin's *Grammar of the Natural Sciences* has been translated by Anthimos Gazis, one of the most distinguished Greek teachers of that period.

Perhaps it should be emphasized that many of the Greek scholars did not consider it convenient to refer to the authors of the originals, or the exact editions they used. One of these "unknown" writers was Peter van Musschenbroek. Abbé Nollet and Christian Wolff were two other scientists who had suffered from the same behavior. It was mostly the Italian revised editions of their books that were used widely by Greek scholars.

Among the most famous Greek scholars were Nikephoros Theotokis and Eugenios Boulgaris, who extensively used Musschenbroek's *Elementa Physicae* in their own works *Stichia Physikis* (Elements of Physics) and *Areskonta tois philosophois* (What philosophers really like) accordingly.[6]

Discovering the European scientists whose works were used by Greek scholars is another important problem that the historian of Science has to overcome, especially if he wants to point out the so-called "common sources". In this effort a researcher has to trace:

(a) the European countries a certain scholar had visited as a student or later;
(b) the foreign languages he knew;
(c) the references to European scientists in the Greek text (probably more detailed references indicate a source, but this is not always the case); and

(d) the title of the Greek book which sometimes reveals influences.

Up to this point, speaking of science, we meant physics. During the period under examination not only had physics come to a certain point of interest and development, but geography as well. This was because, under the social conditions prevailing in the region, people had a strong desire to be informed about the situation in other countries and the morphology of the world.

It should be mentioned that 'Cosmography at that time was part of geography. Therefore some serious scientific and philosophical questions (like the shape of the earth and the "system of the universe") were answered through the pages of such texts as Iossipos Moisiodaxe's (a radical supporter of Newton's theories) *Theory of Geography*.[7] Generally, there was no objection among the scholars that the shape of earth was almost spherical and not a plane. In the book *Geographia Neoteriki*, whose title could be translated as "Geography Based on New Opinions" written in 1791 by Fidipides and Konstantas, the following text appears in the first three lines:[8]

> The Earth, where we are living is not an infinitely extended plain valley as it seemed to be, but almost a sphere surrounded by air.

As for the "system of the universe", had known the opinions of Copernicus since the beginning of the 18th century, the hypothesis that the Earth was not the centre of the cosmos was not accepted until the end of the century, after a rough procedure. Most Greek scholars were supporters of Tycho Brahe's theory, which was the theory accepted by the official Church.

Another point that has to be mentioned is that even in the first published book of physics (*Elements of Physics* written by Nikephoros Theotokis, Leipzig 1766, 2 Vols.), there are several and important references to chemistry (or "Chymistry" in the earliest versions) as a distinguished scientific field. The first book completely dedicated to the new science was the *Philosophie Chimique* of A. Fourcroa, translated and published by Anthimos Gazis in 1802. The conservation of the term "philosophy" in this pioneering effort to construct the new science of Chemistry, is of special interest. It is a conservation that is absolutely connected with the same term in the first books of Physics or "Natural Philosophy".

Another question begging an answer is the formation of a scientific community by scholars in the Balkan region. Facts collected until now tend to prove that a scientific community, with peculiar characteristics, composed of members of a certain relationship did exist. One of the unusual characteristics of this scientific community was that none of its members could be considered the intellectual leader. Furthermore, it was a community structured on a basis quite different from what we call today a typical scientific community. Physicists, mathematicians, chemists, philosophers, even politicians participated in this community, which, after 1811, had its own voice, the fortnightly journal, *Ermis o Logios*.[9]

The problems mentioned above have been manipulated in the traditional, qualitative way. But, aside from the qualitative approach, we can study the influence of the West quantitatively, using the "data" which are now available. An example illustrating this quantitative approach is the following. From the paraphrased names of European scientists in the most important Greek scientific books, from their distribution in years of birth, nationalities, and field of specialization, we can conclude that almost 50% of European scientists acted during the first period of the scientific revolution while the other 50% acted when the Newtonian scientific concept was accepted as the principal one. As for the nationalities of these scientists, most of them were from France, England, Germany and Italy. The interpretation of this situation is not difficult. French and English scientists have contributed a lot to the development of classical physics; Germans have been mostly known as mathematicians and astronomers. The few Swedish scientists were mainly chemists who reflected the high level of that science in Scandinavia. Physicists and chemists comprised 50% of the total number of scientists referred to in Greek scientific books.

Another historiographical problem that must be faced is the examination of the true beliefs of Greek scholars in some critical questions, which, unfortunately, also have ideological dimensions. It is a common practice of researchers to use only the parts of texts which support their opinions, but this practice undoubtedly, misleads us. One example is *the case of* Nikephoros Theotokis who, in the *Elements of Physics*, seems to accept, at least, the heliocentric system, while in his *Geography*, published twenty-five years later, he seems to be a supporter of the geocentric system. This second opinion usually is forgotten by the researchers, sometimes on purpose, because they want to present Theotokis as one of the leaders of the supporters of modern theories and they believe that this fact could be against that purpose. This issue is also related to the problem of internal coherence of texts written by a certain scientist. In our opinion, such a coherence is not always necessary, although in the case examined, the incoherent step was a step backwards, if one is thinking in a whiggish way.

## THE RECONSTRUCTION

Completing this short inventory of historiographical problems concerning the study of the Greek Enlightenment from the epistemological point of view, there is a last question (which could also be the first).

*Is the reconstruction of this historical fact possible?* The answer cannot be absolutely negative or affirmative.

In fact reconstruction is possible, but with a certain degree of deformation. It is like the instant picture of a fastmoving car or the flight of a seagull under the Greek sun. Mainly, the only thing we can do is to connect arbitrarily the pieces of the puzzle we have in our hand, leaving blank the blanks, having the obligation to inform people how subjective our creation is.

We have always to remember that, as the meaning of words change through time the same happens with scientific theories. So although we have in mind a continuity in this evolution, it may not be true. The only real thing is that the history of science is a history of irresponsible dreams.

*H. N. Hydrographic Service, Athens, Greece*

## NOTES

[1] Ianis Karas, *Oi Thetikes Epistimes ston Elliniko 18. Aiona* (Natural Sciences in the Greek 18th Century), Gutenberg Publ., Athens, 1977, p. 42.

[2] Linos G. Benakis, Apo tin istoria toy metabysantinou aristotelismou ston Elliniko choro. Amfisvitisi ke iperaspisi tou philosophou ston 18. Aiona. Nikolaos Zerzoulis-Dorotheos of Lesvios, *Filosophia* (Philosophy) 7, (1977) pp. 420–421.

[3] Dimitrios N. Darvaris, *Epitomi Physikis* (Synopsis of Physics), Vienna 1812, preface: ‘Physics is not only the most enjoyable but the most useful science as well... It sets us free from the injurious and miserable superstition under the yoke of which so many people without any knowledge of Physics are terribly tortured.”

[4] This is the case of Veniamin Lesvios and Theofilos Kairis who introduced two theories based on the existence of some kind aetherial matter which was named pantachikiniton by the first and phlogiston by the latter.

[5] Iossipos Moisiodax, *Apologia* (*Apology*), Vienna 1780, p. 17: “Thus Aristotelism, pure or mingled with Platonism is an enemy of the healthy philosophy for the eternity...”

[6] George Vlahakis, *I “Fisiki” tou Nikephorou Theotoki, stathmos stin epistemoniki skepsi tou*

*18ou Aiona (Physics of Nikiphoros Theotokis, a turning point in the scientific thought of the 18th century)*, Ph. D. Thesis, Athens Technical University, Athens 1990.

[7] A procedure of presentation of the three systems is proposed by Iossipos Moisiodax in his *Apologia*, p. 36: "The question for example is how the shining sun sets and rises every day. The philosopher gives firstly a general idea of the three systems of the planets, secondly he explains each one according to the background of the listeners, thirdly he emphasizes the weakness of Ptolemaic and Tychonic systems and at the end after the examination of the case under study with a historical method he leaves the final choice of the correct system to his systems.

[8] Fillipides-Konstantas, *Geographia Neoteriki (Geography based on new opinions)*, Vienna 1791, p. 1.

[9] *Ermis o Logios* was a periodical publication in Vienna supporting the progressive opinions of the period under consideration. Its edition stopped abruptly in 1821, when the Austrian authorities obliged the editors to publish the "excommunication" of the Greek War of Independence by the Patriarch of the Orthodox Church Grigorios the fifth.

MICHEL BLAY

# HISTORY OF SCIENCE AND HISTORY OF MATHEMATIZATION: THE EXAMPLE THE SCIENCE OF MOTION AT THE TURN OF THE 17TH AND 18TH CENTURIES*

For the sake of clarity I shall enounce, from the outset, the meaning I give to the expression "history of mathematization":
- It is a historical approach whose object is to examine the precise role of mathematics and its development, as a dynamic and creative factor, in the realm of mathematical physics. Or to put it differently:
- It is an examination, based on historical examples, of the specifically mathematical impact of mathematical physics in conceptual coordination, coherence and genesis.

This paper aims to show the fruitfulness of such an approach in understanding, for example, the specific development of the science of motion at the turn of the 17th and 18th centuries.

During the first decade of the 18th century, using the new methods of differential and integral calculus, Varignon wrote many papers devoted to the science of motion.[1] These papers, inserted in the volumes published by the Paris Académie Royale des Sciences,[2] deal primarily with questions that Newton already tackled, some years earlier, in his famous *Philosophiae Naturalis Principia Mathematica*.[3] However, whether he is concerned with questions that pertain to determining the expression of central forces (*Principia*, Book I), or with the study of the motions of projectiles in resisting mediums (*Principia*, Book II), Varignon clearly emphasizes, in the first Memoirs on each of his investigations, the specificity of his procedure and the novelty of his approach – a specificity and a novelty characterized both by the generality of the results obtained and the ease with which Newton's examples are merely deduced from these results as so many particular cases.

Thus in his 31 March 1700 Memoir Varignon writes:

> Il en résulte aussi une formule très simple des *forces centrales*, tant centrifuges que centripètes, lesquelles sont le principal fondement de l'excellent ouvrage de M. Newton *De Phil. Natur. Princ. Math.* ... me contentant de faire voir ici avec quelle facilité elle expédie les exemples que voici, dont la plupart sont de M. Newton sçavoir ceux des prop. 7, 8, 9, 10, de son premier livre.[4]



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 405-420.

A similar view is taken up again in the opening lines of his first Memoir of 1707 devoted to the motion of projectiles in resisting mediums:

Monsieur Newton dans le Livre qu'il nous a donné *De Principii Math Philos. natur.* Livre 2. Set. 1.2 et 3. M. Leibnitz dans les Actes de Leipsik de 1689, pag. 39 etc., M. Hugens dans son discours de la cause de la pesanteur pag. 168 etc. Et M. Wallis dans ses Oeuvres Mathématiques Tom. 2 chap. 101, pag. 438. etc., ont traité fort doctement de la résistance du milieu au mouvement des corps. Voici ce qui m'est aussi venu en pensée sur cette matière, le tout compris en une Proposition générale, d'où résulte en plusieurs manières, non seulement tout ce que ces quatre grands Géomètres ont conclu de leurs hypothèses; mais encore ce qui suit de plusieurs autres faites à volonté: tout cela paroîtra dans les Problèmes suivant, et dans leurs Corollaires.[5]

While keeping in mind that the style adopted here by Varignon harbors in all likelihood a polemical strain (these Memoirs were composed in the heat the controversy over Leibnizian calculus that disturbed the activities of the Académie Royale des Sciences),[6] one must carefully examine his remarks, since these Memoirs represent by no means a mere introduction or initiation to the methods and the results of the *Principia*, but rather offer a complete reworking. Such is the ambiguity of Varignon's endeavor with respect to the *Principia*.

How will Varignon manage, by putting to use the methods of differential and integral calculus, to "expédier" with "facilité" the difficult problems of central forces solved earlier by Newton? Or, to be more precise, how did it become possible to introduce the differential formalism? And how did introducing it alter the problem of central forces?

To answer these questions, we must first return to two of Varignon's papers written shortly prior to 1700.

## 1. THE TWO 1698 MEMOIRS AND THE CONSTRUCTION OF THE CONCEPT OF VELOCITY IN EACH INSTANT

It is in two Memoirs, read at the sessions of the Académie Royale des Sciences respectively on Saturday 5 July and Saturday 6 September 1698, that Varignon, by working out the concept of "vitesse dans chaque instant", gave the science of motion a new start.[7]

The object of the first Memoir is to yield a general expression of the velocity making it possible to treat every motion in the case of rectilinear trajectories, whatever the velocity's mode of variation: "Règle générale pour toutes sortes de mouvements de vitesse quelconques variées à discrétion.'[8]

On the whole Varignon's procedure consists in considering as uniform the velocity of a body in each instant of its motion or, to take up an expression employed by Fontenelle in the *Histoire de l'Académie Royale des Sciences* for the year 1700:

... M. Varignon n'a pas laissé de traiter les mouvemens variés comme les uniformes, & de tirer des uns les mêmes conséquences que des autres.[9]

In his Memoir Varignon begins by defining several variables[10] by means of which the conceptual process can be carried out (Figure 1).

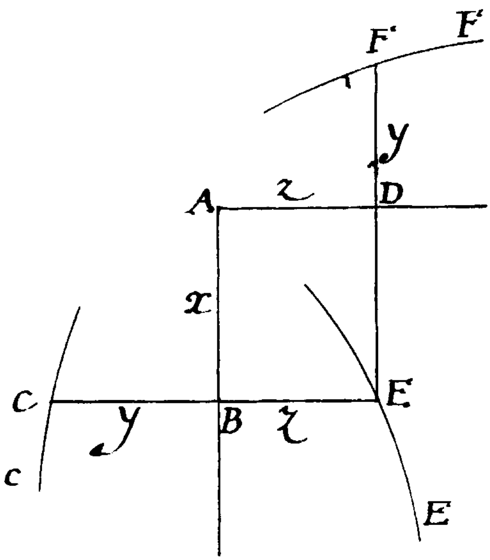



Fig. 1. *A.Ac.Sc. Registres*, t. 17 bis, fol. 298 r°.

... (tous les angles rectilignes de la figure qu'on voit icy, etant droits) soient AB = x les espaces parcourus en quelque sens qu'on voudra, BE = z les temps employez à les parcourir, et BC = y = DF les vitesses à chaque point B de ces espaces.[11]

The variables, space, time, and velocity, now being represented by the "axes" or line segments AB, BE, and BC (or DF), it then becomes possible, by associating these variables two by two, to generate what we would call today the space diagram EE, and the two velocity diagrams CC and FF:[12]

... de maniere que la courbe EE (quelle qu'elle soit) formée par tous les points E, exprime les espaces par ses abscisses AB, et les temps employez à les parcourir, par ses ordonnées correspondantes BE; la courbe CC (quelle qu'elle soit aussi) formée par tous les points C, exprime de même les espaces par les abscisses AB, et les vitesses à chaque point B de ces espaces, par ses ordonnées correspondantes BC; enfin la courbe FF (quelle qu'elle soit encore) exprime aussi par ses ordonnées DF ces mêmes vitesses comparées aux temps BE ou AD qui leur servent d'abscisses.[13]

Let us now take into account what happens in a particular instant:

Cela posé, les instants seront = $dz$;[14] l'espace parcouru dans chaque instant, sera = $dx$; et la vitesse avec laquelle dx aura été parcourue sera = $y$.[15]

Thus in an instant equal to $dz$ the space travelled by the mobile will be equal to $dx$, whereas the velocity of the mobile will be equal to $y$. Varignon goes on to assert that the velocity in each instant may be "regardée comme uniforme":

De sorte que cette vitesse ($y$), dans chaque instant pouvant être regardée comme uniforme, a cause que $y \pm dy = y$, la notion seule des vitesses uniformes donnera $y = dx/dz$ pour la regle de tous les mouvemens variés comme on voudra, c'est a dire quelque rapport d'espace, de temps, ou de vitesse, qu'on suppose; la vitesse de chaque instant etant toujours et par tout egale au quotient de l'espace parcouru dans chaque instant divisé par cette même differentielle de temps.[16]

It is essential to emphasize, as Varignon will do explicitly only in a 6 July 1707 Memoir: "Des mouvements variés à volonté, comparés entr'eux et avec les uniformes",[17] that

l'espace et le temps etant des grandeurs heterogenes, ce n'est point proprement elles qu'on compare ensemble dans le rapport qu'on appelle vitesse, mais seulement les grandeurs homogenes qui les expriment, lesquelles sont ici, et seront toujours dans la suite, ou deux lignes, ou deux nombres, ou deux telles autres grandeurs homogenes qu'on voudra.[18]

Such being the case, it is obvious that the line of reasoning that leads Varignon to the concept of velocity at each instant relies directly on the procedures of the calculus of differences,[19] as in the *Analyse des Infiniment petits, pour l'intelligence des lignes courbes*, published in Paris in 1696 by L'Hospital. In this book we read, for example:

I. Demande ou Supposition.

2. On demande qu'on puisse prendre indifféremment l'une pour l'autre deux quantités qui ne diffèrent entre elles que d'une quantité infiniment petite: ou (ce qui est le même chose) qu'une quantité qui n'est augmentée ou diminuée que d'une autre quantité infiniment moindre qu'elle, puisse être considérée comme demeurant la même.[20]

Varignon is thus able to identify $y \pm dy$ with $y$. The velocity $y$ may then be "considérée comme demeurant la même" during the interval of time $dt$ during which the length $dx$ was traveled. In consequence "la notion seule des vitesses uniformes" yields immediately, indeed, the expression of the velocity $y$ in each instant $y = \mathrm{d}x/\mathrm{d}t$.

Varignon will clarify this concept of velocity in each instant or, as he also calls it, instantaneous velocity[21] to which he returns, in particular, in his posthumous work published in 1725 in Paris under the title: *Traité du mouvement et de la mesure des eaux coulantes et jaillissantes. Avec un Traité préliminaire du mouvement en général. Tiré des ouvrages manuscrits de feu Monsieur Varignon, par l'Abbé Pujol:*

Remarque ... cependant chaque vitesse instantanée est égale et uniforme en elle-même, parce qu'en supposant qu'elle répond à un instant infiniment petit, elle ne peut avoir pendant cet instant infiniment petit, qu'une variation infiniment petite, et par conséquent nulle par rapport à la variation qui arrive dans un tems fini.[22]

His characterization of velocity in each instant enables Varignon to formulate, in his 5 July 1698 Memoir a "Règle générale". Here he is merely taking up again his characterization under three different aspects, emphasizing successively, by means of algebraic manipulations, velocities, times and spaces:

Regle generale

Des vitesses  Des temps  Des espaces

$$y = \frac{\mathrm{d}x}{\mathrm{d}z} \quad \text{ou} \quad \mathrm{d}z = \frac{\mathrm{d}x}{y} \quad \text{ou} \quad \mathrm{d}x = y\mathrm{d}z.$$ [23]

From these three formulas it obviously follows that:

Quelles que soient presentement *la vitesse d'un corps* (accelerée, retardée, en un mot variée comme on voudra), *l'espace parcouru, et le temps employé a le parcourir*; deux de ces trois choses etant données a discretion, il sera toujours facile de trouver la troisiême par le moyen de cette regle, même dans les variations de vitesse les plus bisarres qui se puissent imaginer.[23]

Varignon's construction of the velocity in each instant was therefore ultimately the result of a twofold conceptual process:

- the first making it possible, by transcending the traditional debate over homogeneous magnitudes, to express the concept of velocity as a quotient.
- the second making it possible, by introducing the idea according to which in an instant of time the velocity may be taken to be constant, to bring in readily the concepts of Leibnizian calculus

Such being the case, Varignon's formulation of the "Règle générale" concludes the first part of the 5 July 1698 Memoir. The second part aims at putting this "Règle générale" to use by means of five examples.

Two months later, on September 6, 1698, Varignon completes this first study by a second Memoir devoted to motion in curvilinear trajectories: "Application de la Règle générale des vitesses variées, comme on voudra, aux mouvements par toutes sortes de courbes, tant mécaniques que géométriques. D'où l'on déduit encore une nouvelle manière de démontrer les chutes isochrones dans la cycloïde renversée".[24]

First, Varignon summarizes the results of the 5 July 1698 Memoir:

Le 5 juillet dernier (1698) je prouvay en general à l'Academie que de quelque variation de vitesse qu'un corps se meuve, ce qu'il en a à chaque instant, est toujours égal au quotient de ce qu'il parcourt alors d'espace divisé par cet instant.

D'où je conclus qu'en prenant AB = $x$ pour tout l'espace parcouru, les ordonnées BC = $y$ d'une courbe quelconque CC pour les vitesses a chaque point B de cet espace, et les ordonnées BE = $z$ d'une autre courbe quelconque EE pour les temps employez a venir de A en B; l'on aura toujours $dz = \frac{dx}{y}$.[25]

What Varignon is now concerned with is to generalize this result to the case of motion in curvilinear and no longer rectilinear trajectories.

Varignon then takes the uniform velocity in each instant to be no longer the ratio of an infinitely small line segment d$x$ to an infinitely small time d$z$, but that of the arc of an infinitely small curve (assimilated with a line segment capable of being determined by means of the Pythagorean theorem applied to infinitesimal increments)[26] to an infinitely small time (Figure 2):

... si au lieu de AB le corps se meut de même vitesse correspondante de AK (parallele aux ordonnées) en G le long d'une courbe quelconque GG,[27] dont les ordonnées sont BG = $v$, pendant des temps exprimés par les ordonnées BH = $s$ d'une autre courbe quelconque encore HH; il suit de même que l'on aura toujours $ds = \sqrt{(dx^2 + dv^2)}/y$, ou (en prenant $a = 1$ pour observer la loy des homogènes $ds = \frac{a}{y}\sqrt{dx^2 + dv^2}$, d$z$ se changeant icy en d$s$, et d$x$ en $\sqrt{dx^2 + dv^2}$.[28]

Varignon gives this preference to the expression of the instant of time d$s$ over that of the velocity at each instant $y$. For, as the remainder of the Memoir will make clear, what interests Varignon here is not to find the velocity, but rather to find the time taken to travel along all sorts of curves, either mechanical or geometrical, the velocities and the trajectories being given; the ultimate goal, as the title announces, is to prove, by means of this new method, the isochronism of fall in reversed cycloids.[29]

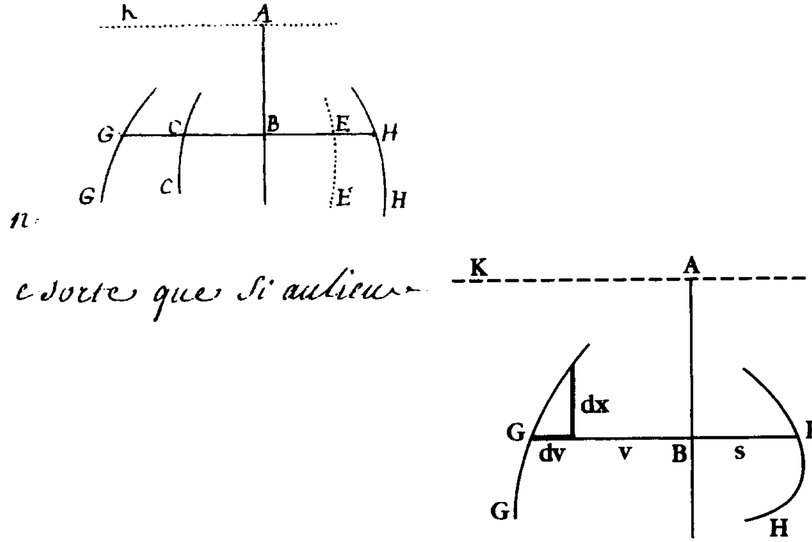



*Fig. 2. A.Ac.Sc. Registres*, t. 17 bis, fol. 387 v°.

.. Si l'on substitue la valeur de $y$ en $x$ par l'équation donnée de la courbe CC,[30] et $v$ encore en $x$ par l'équation aussi donnée de la courbe GG, l'on aura en $x$ et en $s$ l'equation cherchée de la courbe HH dont les ordonnées BH ($s$) expriment les temps employez a venir de AK en G le long de la courbe GG.[31]

Thus ends the first part of the Memoir, the second part, as in the previous Memoir, being taken up by examples, which enable Varignon to bring out the relevance of his new concepts for treating problems of motion. He has only to sum up the general import of the 5 July and 6 September 1698 Memoirs as well as the fruitfulness of the "Règles" that were set up:

Ces deux exemples suffisent pour faire voir comment on se doit servir de la formule $ds = \frac{a}{y}\sqrt{dx^2 + dv^2}$, pour trouver les temps employez a parcourir toutes sortes dye courbes tant mecaniques que geometriques, quelque variée que soit la vitesse des corps qui les parcourent, n'y ayant qu'a substituer telle autre courbe qu'on voudra, au lieu de la parabole sur les ordonnées de laquelle on vient de regler ces vitesses pour s'acommoder a l'hypothese de Galilée touchant la chutte des corps.

Non seulement la courbe des temps se peut ainsy trouver, celle des chemins et des vitesses etant données; mais encore chacune de celles-cy se trouvera, les deux autres etant données en operant comme dans les exemples de la Regle generale que je demontray le 5 juillet dernier a l'Academie: c'est pourquoy je ne m'y arreteray pas davantage.[32]

Varignon, who uses the resources of Leibnizian calculus, is thus now in possession of the concept of velocity in each instant, valid both for rectilinear and curvilinear motions.

How will Varignon apprehend questions belonging to the science of motion and, in particular, those raised by Newton in the first book of the *Principia*?

## 2. THE DIRECT PROBLEM OF CENTRAL FORCES IN THE MEMOIRS OF 1700

### 2.1. *The Three Memoirs of 1700*

As early as 1700, in three successive Memoirs, Varignon seeks to give the problem of central forces all the clarity and generality that the new methods of Leibnizian calculus along with his research on the concept of velocity in each instant can bring.[33]

These three Memoirs dated 30 January, 31 March and 13 November 1700, are entitled respectively:

- Manière générale de déterminer les forces, les vitesses, les espaces, et les temps, une seule de ces quatre choses étant donnée dans toutes sortes de mouvemens rectilignes variés à discrétion.[34]
- Du mouvement en général par toutes sortes de courbes; et des forces centrales, tant centrifuges que centripètes, nécessaires aux corps qui les décrivent.[35]
- Des forces centrales, ou des pesanteurs nécessaires aux planètes pour leur faire décrire les orbes qu'on leur a supposés jusqu'icy.[36]

As the titles of the first two Memoirs imply, one encounters again here the conceptual order of the two Memoirs of 1698, whereby, in particular, the concept of velocity in each instant is introduced: the first one will be devoted to "forces centrales", or the force vers (un centre) $c$ of a body or again the "tendance au point $c$ comme centre",[37] in the case of bodies describing rectilinear trajectories. The second Memoir deals with the case of "forces centrales" where the bodies describe curvilinear paths around a given fixed center. In the third Memoir Varignon abandons examples of a purely mathematical nature in favor of the study of planetary orbits.

We will confine our analysis here to the case of central forces causing bodies to follow rectilinear trajectories.

### 2.2. *Central forces of bodies describing rectilinear trajectories*

#### 2.2.1. *The general expression of central force*

a. *The velocity in each instant and the increment of the velocity.* The first Memoir dated 30 January 1700 follows as a development of that of 5 July 1698 on rectilinear motion. Varignon points this out in an opening paragraph missing from the final version published in the *Mémoires de l'Académie Royale des Sciences*, but that is found in the Minute book of the sessions of the *Académie Royale des Sciences*:

Le clnq juillet de 1698 je demontray a l'Academie une Regle generale pour toutes sortes de mouvemens varies a discretion, et telle qu'en prenant sur la figure presente (dont tous les angles rectilignes sont droits) AH pour tout l'espace parcouru, les ordonnées VH = VG d'une courbe quelconque, BV ou VK pour les vitesses à chaque point H de cet espace, et les ordonnées HT d'une autre courbe aussi quelconque TD pour les temps employés a venir de A en H; une de ces trois courbes etant donnée, l'on en deduira toujours les deux autres.[38]

Varignon then goes on to make it clear, in the same unpublished text, that in this new Memoir his view is enhanced: he will now raise the question of the intensity of the central force in association with the variation of motion:

Mais comme en cela je ne comprenois point alors la force vers C qu'a le corps en chaque point H, indépendamment de sa vitesse et que j'appelleray dorenavant *force centrale* a cause de sa tendance au point C comme centre".[39]

To solve this problem, Varignon assumes the different variables required and materializes their relations[40] by means of a graph whose construction, in its principle, is identical with that presented in the 5 July 1698 Memoir (see Figure 3). In the *Mémoires de l'Académie Royale des Sciences* Varignon's new study thus opens with these words:

Tous les angles rectilignes étant droits dans la figure que voicy, soient six courbes quelconques TD, VB, FM, VK, FN, FO, dont les trois premières expriment par leur abscisse commune AH, l'espace parcouru par un corps quelconque mû comme l'on voudra le long de AC. Soit de même le temps employé à le parcourir, exprimé par l'ordonné correspondante HT de la courbe TD; la vîtesse de ce corps en chaque point H, par les ordonnées aussi correspondantes VH, VH, des courbes VB, VK; ce qu'il de force vers C, à chaque point H, indépendamment de sa vîtesse (je l'appelleray dorénavant *Force centrale* à cause de sa tendance au point C comme centre), s'exprimera de même par les ordonnées correspondantes encore FH, FG, FE, des courbes FM, FN, FO.[41]

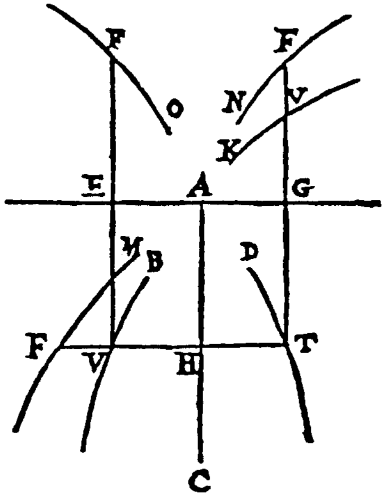



*Fig. 3.*

Henceforth Varignon names these six fondamental curves TD, VB, FM, VK, FN and FO respectively (let us recall that the body describes any rectilinear motion AC):

... la courbe TD, à laquelle les ordonnées HT se terminent en T, s'appellera la *courbe des temps*.[41] (The abscissa is AH).

Les deux courbes VB, VK, auxquelles les ordonnées correspondantes et égales VH, VG, se terminent en V, s'appelleront les *courbes des vitesses*.[41] (The abscissa are respectively AH and AG or HT).

The curves VB and VI correspond to what I called, in my presentation of the 5 July 1698 Memoir, the first and second diagrams of velocity.

Enfin les trois courbes FM, FN, FO, auxquelles les ordonnées correspondantes encore et égales FH, FG, FE, se terminent en F, s'appelleront les *courbes des forces*.[41] (The abcissa are respectively AH, HT, or AG, and VH or EA).

In short, one may say, in a slightly modernized language, that the curve TD expresses the variations of the space AH in function of the time HT; the curve VB that of the space AH in function of the velocity VH; the curve VK that of the time AG in function of the velocity VG; the curve FM that of the space AH in function of the central force FH; the curve FN that of the time AG in function of the central force FG; and, finally, the curve FO that of the velocity EA in function of the central force FE.

Varignon then associates with each of the four variables, space, time, velocity and force, a definite algebraic symbol:

... soient les espaces parcourus AH = $x$, les temps employés à les parcourir HT = AG = $t$, les vîtesses en H (que j'appelleray *finales*) HV = AE = GV = $v$, les forces centrales correspondantes HF = EF = GF = $y$.[42]

As a result of this procedure of algebraization, Varignon is able to set up, in the first place, the kinematic concepts already considered in the 5 July 1698 Memoir and necessary for his subsequent investigation of central forces in the case of rectilinear motions:

De là on aura $\mathrm{d}x$ pour l'espace parcouru comme d'une vitesse uniforme $v$, à chaque instant; $\mathrm{d}v$ pour l'accroissement de vitesse qui s'y fait; $\mathrm{dd}x$ pour ce qui se parcourt d'espace en vertu de cet accroissement de vîtesse; et $\mathrm{d}t$ pour cet instant.

A ce compte, la vîtesse ne consistant que dans un rapport d'espace parcouru d'un mouvement uniforme, au temps employé à le parcourir; l'on aura déjà $v = \frac{\mathrm{d}x}{\mathrm{d}t}$ pour une première Régle, laquelle donnera $\mathrm{d}v = \frac{\mathrm{dd}x}{\mathrm{d}t}$ en faisant $\mathrm{d}t$ constante. [42]

Varignon is thus now in possession of two differential expressions pertaining to the concept of velocity:

- on the one hand, that of the velocity in each instant : $v = \frac{\mathrm{d}x}{\mathrm{d}t}$
- on the other, that of the increment of this same velocity during the same instant $\mathrm{d}t$ : $\mathrm{d}v = \frac{\mathrm{dd}x}{\mathrm{d}t}$ .[43]

Actually, varied motion can be interpreted as a result at each instant of a uniform motion of velocity $v$, equal to the total velocity acquired during the previous interval of time, and a motion equally uniform of velocity $\mathrm{d}v$, the velocity acquired at the very beginning of the time interval $\mathrm{d}t$.[44] This second velocity can be neglected with respect to the first within the time interval $\mathrm{d}t$.

b. *From the increment of velocity to the central force.* Setting up the differential expressions pertaining to central force will hing on the Galilean model for falling bodies:

De plus les espaces parcourus par un corps mû d'une force constante et continuellement appliquée, telle qu'on conçoit d'ordinaire la pesanteur, étant en raison composée de cette force et des quarrés des temps employés à les parcourir; l'on aura aussi $\mathrm{dd}x = y\mathrm{d}t^2$, ou $y = \frac{\mathrm{dd}x}{\mathrm{d}t^2} = \frac{\mathrm{d}v}{\mathrm{d}t}$ . Ce qui fait encore une Règle $y = \frac{\mathrm{d}v}{\mathrm{d}t}$, qui avec la précédente $v = \frac{\mathrm{d}v}{\mathrm{d}t}$, satisfait à tout ce qu'on se propose icy de résoudre.[45]

Here Varignon adopts the conceptualization of the accelerating force as proposed by Newton in Lemma X of Section one of the first book of the *Principia*. To compare, in its light, the first sentence of Varignon's text with Newton's formulation of Lemma X is highly revealing:

The spaces which a body describes by any finite force urging it, whether that force is determined and immutable, or is continually augmented or continually dimished, are in the very beginning of the motion to each other as the squares of the times.[46]

Corollary 3 of this Lemma also stipulates:

Cor. III. The same thing is to be understood of any spaces whatsoever described by bodies urged with different forces; all which, in the very beginning of the motion, are as the product of the forces and the squares of the times.[46]

In consequence, as during the time interval dt the force, supposed "constante et continuellement appliquée", produces an increment of space equal to ddx, one may write, while strictly following Varignon:

$$\mathrm{dd}x = y\mathrm{d}t^2 \text{ or } y = \frac{\mathrm{dd}x}{\mathrm{d}t^2}$$

$$\text{given } y = \frac{\mathrm{dd}x}{\mathrm{d}t.\mathrm{d}t}\text{, therefore } \mathrm{d}v = \frac{\mathrm{dd}x}{\mathrm{d}t} \text{ hence } y = \frac{\mathrm{d}v}{\mathrm{d}t}.$$

However, as I noted above that the variation of the velocity dv, because of the form of its expression, must appear at the very beginning of the time interval dt, it is essential in this case to bear in mind that a force acts instantaneously at the beginning of this time interval dt then ceases to act until the beginning of the next time interval. In this view the expression "force constante et continuellement appliquée" betrays certain underlying problems pertaining to the production and nature of motion. The difficulties can be brought out in the following way by returning to the expression of the increment of velocity: in the case where the force actually acts in a continuous and constant manner during the time interval d$t$, so that the increment of velocity acquired at the end of this time interval is still d$v$, the space traveled will nog longer be dd$x$ but ½ dd$x$. Yet Varignon reaches the expected result; for, when passing from a proportionality to an equality in the expression of the force, he neglects to take into account the coefficient of proportionality.

By expressing himself slightly differently and adopting the standpoint of the determination of the expression of the force, one may say that the difficulties of Varignon's conceptualization lie mainly in the fact that the expression of the increment of velocity d$v$ and that of the force $y$ imply, as it were, an ambiguous modelisation of the force's mode of action. In the sense that the latter is supposed to act, according to the expression that Varignon is trying to determine, either at the very first instant (such is the implicit consequence of the solely mathematical manipulations that lead to the expression for d$v$), or in a constant and continuous manner during the entire time interval d$t$ (such is the explicit hypothesis involved in calculating $y$).

In the last analysis, one may take Varignon to ask, from the standpoint of the coherence of his conceptualization, for not having analysed in a rigourous manner the construction of the concept of the increment of velocity and for having undoubtedly once again gotten carried away, by only considering the mathematical operation "faire dt constante", because of the great pleasure he derived from setting the machinery of the new calculus going.

It then appears here that, although he reaches a satisfactory expression for $y$, Varignon unfortunately leaves unanswered several questions both technical and conceptual.

### 2.2.2. *Statement and use of the General Rules*

The general expression for central force in the case of rectilinear motions now being, as it were, acquires, Varignon is able to state what he calls the "Règles générales des mouvements en lignes droites":

$$1°.\ v = \frac{\mathrm{d}x}{\mathrm{d}t} \qquad 2°.\ y = \frac{\mathrm{d}v}{\mathrm{d}t}\left(\frac{\mathrm{dd}x}{\mathrm{d}t^2}\right).$$ [47]

What do these "Régles générales" show us? Precisely that the concepts of velocity in each instant and of accelerating force in each instant, which have just been successively constructed by Varignon, can actually be deduced from one another by means of a simple calculation using Leibnizian algorithms, and, in consequence, as Auguste Comte writes, "d'après ces formules, toutes les questions relatives à cette théorie préliminaire du mouvement varié se réduiront immédiatement à de simples recherches analytiques, qui consis teront ou dans des différentiations, ou, le plus souvent, dans des intégrations".[48] After stating his General rules, Varignon determines their use in the very next lines:

Usage. Je dis présentement qu'une des six courbes cy-dessus, étant donnée à discrétion, on pourra toujours en déduire les cinq autre par le moyen de ces deux Règles, supposé les résolutions et les intégrations nécessaires des égalités en question.[49]

Then, having analysed two quite simple examples, Varignon sums up the generality of his method:

Les mêmes choses se trouveront de la même manière dans toute autre hypothèse; il n'y aura de différence que la difficulté du calcul, laquelle n'auroit fait qu'embarasser icy. Ainsi ces deux exemples suffisent.[50]

In the light of the analysis I have developed here, and which represents only one aspect of Pierre Varignon's diverse research, it is his work as a whole that ultimately takes on a new significance. Varignon's work is all to often either completely forgotten or, at best, presented by historians of classical mechanics in a few sentences dealing exclusively with his contributions to statics ("polygone de Varignon" or vector polygone).

The reason for this lack of interest for the greater part of Varignon's work is quite easy to understand given our approach: the historians of classical mechanics, who for the most part only see mathematics, after the fashion of the positivists, as an instrument or tool at the disposal of science, emphasize in their studies the historical analysis of the concepts or the principles of mechanics and neglect to think of mechanics, or, more generally, physics, as mathematical physics, that is, they most often fail to fully recognize the mathematical activity proper of mathematical physics in the conceptual coordination, coherence and genesis. Yet here lies precisely Varignon's endeavor.

As a result, the history of the science of motion appears in a new light.

*Centre National de la Recherche Scientifique, Paris, France*

## NOTES

*Abbreviations*

*A.Ac.Sc. Registres*: Archives de l'Académie des Sciences. Registres manuscrits des Procès-verbaux des séances de l'Académie des Sciences.

*AH; AM*: *Histoire de l'Académie Royale des Sciences avec les Mémoires de Mathématique et de Physique pour la même année. Tires des Registres de cette Académie* (*AH*: *Histoire*; *AM*: *Mémoires*).

*Analyse:* G. de l'Hospital, *Analyse des infiniment petits pour l'intelligence des lignes courbes* (Paris, 1696).

*Briefwechsel*: *Der Briefwechsel von Johann I Bernoulli* (Birkhäuser Verlag, 1955 ...). Vol. I: Correspondence with L'Hospital, edited by O. Spiess (1955). Vol II-1: Correspondence with Varignon, edited by P. Costabel and J. Peiffer (1988).

* This paper has been translated into english by Anastasios Brenner.

[1] "Liste des Oeuvres de Pierre Varignon", *Briefwechsel*, II-1, 387–408.

[2] *AM*, since 1699 (1702).

[3] Explicit references to Newton's *Principia* appear quite early under Varignon's pen: *Briefwechsel*, II-1, 96 and 229 (Letters to Johann I Bernoulli dated 24 May and 12 July 1699).

[4] *AM*, 1700 (1703), 84.

[5] *AM*, 1707 (1708), 382.

[6] On this issue see, in particular, Michel Blay "Deux moments de la critique du calcul infinitésimal: Michel Rolle et George Berkeley", *Revue d'Histoire des Sciences* (1986), 223-253.

[7] I have edited these Memoirs under the title "Quatre Mémoires inédits de Pierre Varignon consacrés à la science du mouvement", *Archives Internationales d'Histoire des Sciences* (1989), 218–248. For an indepth study of these matters, see my book *La naissance de la mécanique analytique* (Paris, PUF, 1992).

[8] *A.Ac.Sc. Registres*, t. 17, fol. 298 v° – 305 r°.

[9] *AH*, 1700 (1703), 85.

[10] The term "variable" belongs to the language of the period. See *Analyse*, "Définition I".

[11] *A.Ac.Sc. Registres*, t. 17, fol. 298 v°.

[12] I recall that in modern terms using the concept of function ($x$ representing space, $t$ time, and $v$ velocity) motion is defined by its temporal equation $x = f(t)$. Analysis of this function will give the behavior of the point on the trajectory. The graph of the function $a = f''(t)$ is called the diagram of accelerations. Varignon also introduces, to use current terminology, the function $v = g(x)$, whose graph we name the first diagram of velocities (the second being that of $v = f'(t)$)

[13] *A.Ac.Sc. Registres*, t. 17, fol. 298 v°.

[14] Varignon does not define explicitely here the concept of instant. On the other hand, in his *Traité du mouvement et de la mesure des eaux coulantes et jaillissantes. Avec un Traité préliminaire du Mouvement en général. Tiré des ouvrages manuscrits de feu Monsieur Varignon, par l'Abbé Pujol*, published posthumously in Paris in 1725, one reads: "Définition V. Un instant est pris ici pour la plus petite portion de temps possible, et par conséquent doit être regardé comme un point indivisible d'une durée quelconque", 3. In a Memoir dated July 6, 1607 under the title "Des mouvemens variés à volonté, comparés entr'eux et avec les uniformes", the definition appears in a slightly different mathematical setting: "Définition I. Par le mot d'instant, nous entendons ici une particule de tems infiniment petites, c'est-à-dire, moindre que quelques grandeur assignable de tems infiniment petite, c'est-à-dire, moindre que quelque grandeur assignable de tems que ce soit: c'est ce qu'en langage des Anciens l'on appelleroit *minor quavis quantitate data...*", *AM*, 1707 (1708), 222.

[15] *A.Ac.Sc. Registres*, t. 17, fol. 298 v°.

[16] *Ibid.*, fol. 298 v° – 299 r°.

[17] *AM*, 1707 (1708), 222–274.

[18] *Ibid.*, 223.

[19] I will use either the expression "Leibnizian calculus" or "calculus of differences" rather than "differential calculus", which could imply too modern an approach. In the *Analyse des infiniment petits pour l'intelligence des lignes courbes* (Paris, 1696) of L'Hospital, one reads: "Définition II. La portion infiniment petite dont une quantité variable augmente ou diminue continuellement, en est appelée la *différence*", 1.

[20] *Analyse*, 2.
[21] Varignon will also call this concept "vitesse instantanée", *AM*, 1701 (1708), 224. I nevertheless prefer to speak of the velocity at each instant; for Varignon interprets the expression d$x$/d$t$ not as a derivative but rather as a quotient. See especially *A.AC.Sc. Registres*, t. 17, fol. 386 r°.
[22] *Traité du mouvement...*, 22.
[23] *A.Ac.Sc. Registres*, t. 17, fol. 299 r°.
[24] *Ibid.*, t. 17, fol. 386 r° - 391 v°.
[25] *Ibid.*, fol. 387 r°.
[26] See *Analyse*, 3.
[27] KG represents the distance covered along its trajectory GG. This point K, the origin of the motion along the trajectory, will be very useful in the rest of the Memoir.
[28] *A.Ac.Sc. Registres*, t. 17, fol. 387 r°.
[29] In 1697 and 1698 Varignon already worked on the problem of isochronous fall in the case of the reversed cycloid, but this procedure remained geometrical and did not belong to an approach using the concept of velocity at each instant.
[30] "La courbe cc" corresponds to the first diagram of velocities (see above note 12).
[31] *A.Ac.Sc. Registres*, t. 17, fol. 387 v°.
[32] *Ibid.*, fol. 391 r° - 391 v°.
[33] *Supra Note* 7.
[34] *AM*, 1700 (1703), 22–27; *A.Ac.Sc. Registres*, t. 19, fol. 31 r° - 37 r°.
[35] *AM*, 1700 (1703), 893–101; *A.Ac.Sc. Registres*, t. 19, fol. 133 v° - 141 v°.
[36] *AM*, 1700 (1703), 218 - 237; *A.Ac.Sc. Registres*, t. 19, fol. 360 v° - 364 v°.
[37] *AM*, 1700 (1703), 22. The expression "force centrale" seems somewhat ambiguous; for Varignon, at no point, explicitly introduces any consideration of the mass.
[38] *A.Ac.Sc. Registres*, t. 19, fol. 31 r°.
[39] *Ibid.*, fol. 31 r°.
[40] *Ibid.*, fol. 31 r°.
[41] *AM*, 1700 (1703), 22.
[42] *Ibid.*, 23.
[43] Varignon by "faisant dt constante" merely makes use of procedures current at the time of Leibnizian calculus. On this issue, see H. J. M. Bos, "Differentials, Higher-Order Differentials and the Derivative in the Leibnizian Calculus", *Archive for History of Exact Sciences* (1974–1975), 3–90.
[44] See also E. J. Aiton, "The Inverse Problem of Central Forces", *Annals of Science* (1964), 86, note 17.
[45] *AM*, 1700 (1703), 23.
[46] *Principia*, translated into English by Andrew Motte in 1729 (University of California Press, 1934, 1962), 34–35.
[47] *AM*, 1700 (1703), 23.
[48] Auguste Comte, *Philosophie première, cours de philosophie positive, leçons 1 à 45* (Paris, Hermann, 1975), dix-septième leçon, 268.
[49] *AM*, 1700 (1703), 23.
[50] *Ibid.*, 26.

PIAMA GAIDENKO

# THE ARTISTIC CULTURE OF THE RENAISSANCE AND THE GENESIS OF MODERN EUROPEAN SCIENCE

Secularisation of every area of life was one of the most important features of Renaissance culture. Each sphere experienced a growing autonomy. Human personality itself became autonomous and came to be regarded as the centre of the universe and of great value in its own right.

This anthropocentric view distinguishes the Renaissance both from the theocentrism of the middle ages and the cosmocentrism of antiquity. Mediaeval people, unlike those of classical times, did not relate themselves to the cosmic element but to the transcendent creator of the world. As a result human personality, Augustine's 'inner man', acquired a previously unheard of value. Yet in the mediaeval understanding of the world this value entirely and fully rested on the union between God and man. It was a reflection of man's divine origin and of God's presence within him. In other words, the personality did not have autonomous value, it was not valuable in itself.

During the Renaissance people tried to cut themselves free of this transcendent root. They did not so much seek support in the universe (though the natural philosophy of the Renaissance testifies to their partial appeal to this source), as in themselves. They saw their own now more profound and complex souls as the centre of the world – their souls and their bodies which were now displayed in a new light that enabled them to see materiality as a whole in a new way. We also find this frame of mind in a great number of the period's scientists, Galileo in particular. Galileo was convinced that human reason was the equal of divine reason. This was not, we should add, in the breadth of its grasp of various objects, the number of which was infinite. No, Galileo had in mind the profundity with which human reason might penetrate its subject:

> ... taking man's understanding *intensively*, in so far as this term denotes understanding some proposition perfectly, I say that the human intellect does understand some of them perfectly, and thus in these it has as much absolute certainty as Nature itself has. Of such are the mathematical sciences alone; that is, geometry and arithmetic, in which the Divine intellect indeed knows infinitely more propositions, since it knows all. But with regard to those few which the human intellect does understand, I believe that its knowledge equals the Divine in objective certainty, for here it succeeds in understanding necessity, beyond which there can



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science*, 421-432.

be no greater sureness.[1]

How were philosophical principles and views of the world refracted in the scientific theories of this period? To find out we must thoroughly analyse the scientist's way of thinking and then compare it to the stylistic features of the period's art, for it is there that this worldview found its most direct reflection. Such a comparative analysis of Renaissance painting and the works of a scientist (in this case Galileo) uncovers some of the distinguishing features of each. And these allow us to speak of a certain unity of culture that was variously refracted in both the science and the art of this epoch.

One of the reasons why perspective painting came into being was that the artist of the Renaissance thought it his task to reproduce the effect of what he saw. What was represented in his picture should give the illusion of corresponding to the subject he was drawing or painting. In his treatise *On perspective painting*, Piero della Francesca, the famous 15th-century Italian artist, put it in the following way:

> Painting is nothing other than a showing of surfaces and bodies shortened or enlarged in a finite plane in such a way that the real things that the eye sees at various angles appear to be real in the said limits. Just as each magnitude has one part nearer the eye than another, and the nearer always appears to the eye in these defined boundaries as a larger degree than that further away; and just as the intellect cannot by itself judge their dimensions, i.e. which is nearer and which further away; so I assert that perspective is essential. It distinguishes proportionately all magnitudes and, as a genuine science, demonstrates the shortening and lengthening of any magnitudes using lines.[2]

Like many other artists of the Renaissance, della Francesca was at the same time a mathematician. His treatise is devoted equally to painting and to geometry or, to be more precise, it was devoted to practical geometry for that was how della Francesca looked at art.

It is insufficient, however, to say that painting according to della Francesca was applied geometry. Classical and mediaeval scholars both interpreted geometry, after all, as the science of spatial relationships existing independent of the human eye. Moreover, the essence of these proportions had its roots not in space as such but in numbers, as Plato had written. As far as geometry in della Francesca's understanding is concerned, its main task was to establish strict correspondence between visible objects and the observing eye. He therefore reinterpreted the Euclidean definitions of a point, line etc., in terms of his own task.

If a point for Euclid is that which has no parts then it is clear that the eye has no role to play here. And if a line is length without breadth then any judgement of it can only be made by reason alone. Euclid's geometry thus does not deal with visible things, i.e. with the empirical world, fragments of which then became the subject of reproduction in Renaissance paintings. Visible things were only illusions, in the Platonist view; yet it is their reproduction on canvas that now became the task of the artist. He had to transform the axioms of Euclidean geometry in such a way that they could be 'grasped by the eye'. A point changed from something "ideal" into something "imaginary": it became something like the "material point" of classical mechanics – a very small body, on the borders of sensible perception. It is this same "hybrid" of the purely mathematical and, at the same time, the corporeal that constitutes Galileo's "material point". The same happens with the line. It acquires breadth but only a very small, barely visible amount.

The eye thus acquires the prerogatives that were formerly only possessed by reason. Hence it becomes possible to hymn the human organ of sight, as in Leonardo da Vinci's treatise "On painting". During the Renaissance man became the centre of the universe, and the human eye became the centre of that centre. This is why the world as it appears to our eye little by little takes the place of the world as it appears to our intelligence.

The Renaissance reliance on optical illusion differs from the mistrust of the senses exhibited by classical thinkers. By just how much we can see in the symbolic striving of the philosophers of antiquity to free themselves from optical perception so as to acquire greater mental perception. Hence the legend about Democritus: for the Greeks a wise man had to be blind because physical sight concealed the truth from him. There is no single classical thinker who would have called the eye, in Leonardo's words, the universal judge of the fate of all bodies. Such a judgement only became possible when human beings ascended above all nature and were deified, not only as spiritual but also as sensual beings.

To get a clearer idea of how the principles of painted representation had changed in the Renaissance compared to the Middle Ages we shall quote Averlino's interesting discussion about the mirror. The terms he used are much more crucial to an understanding of the principles underlying perspective painting than might at first appear. "I think you have understood from the aforesaid how a plane is depicted," says Averlino to his imaginary companion. "I have understood," answers the latter, "but I would like to see how it is done. Tell me, for instance, why these squares do not appear as squares." The reason, answers Averlino, is that you are seeing them lying on a plane. If you saw them directly in front of you they would seem like squares to you.

In order to convince yourself that this is right, look at the floor on which a quadrilateral board has been placed, or look up at the ceiling from below. All the beams are at an equal distance one from another, but when you look at them they seem to be either further apart or nearer together. The further they recede from you, the closer together they appear to be .... If you want to gain a better view of them take the mirror and look into it. You will clearly see that this is the case.

"And painting is indeed a science and the lawful daughter of nature," wrote Leonardo,

for she was engendered by nature. To put it more correctly, however, we should say that she is the granddaughter of nature, since all visible things were engendered by nature and painting was born of these things. Therefore we justly shall name her the granddaughter of nature and a relative of God.[3]

Why is what Averlino wanted to show his companion more clearly seen in the mirror? The reason is that when the eye looks directly at the ceiling beams it is unconsciously guided by the mind. As a result the eye does not immediately notice that the beams grow closer together the further away they are: a person knows that they are everywhere an equal distance apart, and the mind constantly makes adjustments to what is seen by the eye. Therefore some means must be found to dispense with this fusion of mind and eye. The solution is to use a mirror, for when a person looks at a reflection he or she do not so easily associate the visible image there with reality. In this way the object is presented in a purely sensual way, and freed from any type of "intrusion" by the human cognitive faculty.

It is instructive to compare this with what we find, for instance, in Plato. He divided knowledge into two kinds, belief and intelligence, and he linked the first to the visible and the second to the intelligible. He further divided each and arrived at four types of knowledge: "Then it will suffice ... if we call the first division knowledge, the second understanding, the third faith, and the fourth conjecture ... the last two together, belief and the first two, intelligence and say that belief deals with becoming and intelligence with being".[4] He explained that the relation between knowledge and faith was just like that between intelligence and belief, or understanding and conjecture. The latter of these pairs is a lower form of knowledge. In the following passage Plato showed what the latter represented in his elucidation of the distinction between the visible and the intelligible. For the sake of comparison, he suggested,

...take a line divided into two unequal segments, and divide each segment in the same proportion, and suppose the first segment to belong to the visible and the second to the intelligible. The different segments will represent in their comparative clearness and obscurity in [the first segment of] the visible sphere, images. By images I mean, firstly, shadows, secondly, reflections in water, and in things that are close-grained, smooth, and bright, and all similar things ... Make the second segment that whereof this first is an image: ordinary living creatures, everything that grows and everything that is made.[5]

As we can see, Plato placed the reflections of sensual objects precisely in the lowest category of knowledge, conjecture: in other words, just what we see in the mirror and, accordingly, should depict in a picture when using perspective. This is understandable. Plato considered, after all, that the eye gave us an illusory idea of reality and must obligatorily be corrected with the help of reason:

... in these studies some instrument in each man's soul, which was being destroyed and blinded by other pursuits, is here purged and rekindled a salvation more precious than that of thousands of eyes. For with this instrument alone is truth perceived.[6]

The revolution brought about in painting through the introduction of perspective is of world-historical significance. Usually it is emphasised that at this time there was a shift in philosophy as a discipline away from Aristotelianism towards Plato and neo-Platonism. This is quite true. However, the Platonism of the Renaissance differed substantially from that of antiquity and this difference must be taken into account in studying the culture and learning of the former period. The analysis of Renaissance painting, as it happens, can throw light on this difference.

In a certain sense, mediaeval painting was closer to Platonism, and to classical philosophy and learning in general. It tried to depict things as they are perceived by our mind's eye (our inner vision) and therefore was incompatible with perspective. Pavel Florensky, in his study of the mediaeval Orthodox icon, showed that the relations of perspective in icons "are in glaring contradiction with the rules of linear perspective, while the views of the latter could not help being seen [then] as crude illiteracy in drawing".[7]

We can already see that perspective carried with it a new understanding of the world, and expressed a new worldview: the serious and passionate support given by Renaissance artists to the law of perspective and the very nature of their supporting arguments are proof of this.

If we now look at the creation of this new art in terms of the evolution of scientific thought it is important not just to record this turning point in the Weltanschauung. We must also thoroughly analyse the elements that made up the theory of perspective, not least because the latter did not find an entirely identical expression among different artists. There is a characteristic difference in this respect, for example, between Piero della Francesca and Leonardo da Vinci. Both insisted that the artist must observe all the rules of perspective. At the same time, however, Piero saw perspective first and foremost as the means for conveying the proportional relations between objects depicted. Leonardo, meanwhile, regarded it as the means for attaining the sensual verisimilitude of the objects depicted.

This difference between the two artists found expression both in the stylistic features of their works and in their theoretical reasoning. Piero also pointed, it is true, to illusionism as an essential element in perspective painting. For him, however, it did not occupy the prominent position allotted to it by Leonardo. The explanation is that Piero was very much a painter of the 15th century whereas Leonardo already belonged in spirit to the High Renaissance. The former was thus not so far removed from the mediaeval sensibility for illusion not to provoke his inner resistance.

Piero della Francesca was a monumental artist. For him the most important task was the organisation of a wall surface. Painting had not yet become a fully autonomous art for him, and had still not lost its deep inner links with architecture. The transition from painting walls to work on canvas etc. from mural to easel painting, had profound prerequisites in a world view, as it happened, and no less profound consequences. "In the transition to an art depicting the experience of the sense," notes the art historian Hans Sedlmayr,

> pictorial (abbildende) architecture was excluded and its role passed on to the new type of painting. The traits of such a sensual painting are the depiction of the 'non-objective', and of shadows, lighting and perspective.[8]

It is perhaps closer the truth to describe Piero's perspective painting as phenomenalism rather than illusionism. In his works perspective serves more as the means of organising the depicted objects within the area of the picture: it is less capable of creating the illusion of the sensual presence of that which is depicted on a plane surface as in Leonardo. Both phenomenalism and illusionism alike presupposed that all that is absolute is related to the subject, who by the time of the Renaissance was no longer God but man. The difference between them lies in the following. The observing subject in illusionism acts empirically and therefore equates the depiction with the viewer's sensual perception. The artist here relies mainly on the latter. This is less true of phenomenalism. Both artist and viewer are less oriented towards sensual perception and for the most part pay attention to the mathematical laws of spatial proportions. This is of paramount importance for our subject since an analogous process also occurs in the science of the Renaissance. Just as perspective became the method whereby nature was depicted, so geometry became the method whereby nature was understood. At the same time, as Plato explained, geometry as a scientific discipline has, in a sense, a "dual subordination", relying on reason on the one hand, and on the imagination on the other (i.e. a sensual and emotional rather than an intellectual and moral capability). To the extent that geometry relies on reason it is the science of proportional relations and related, in Plato's words, to arithmetic, the science of numbers and the highest of all mathematical disciplines. To the extent that it relies on imagination, however, geometry studies proportion in its spatial refraction. Therefore, it touches the sphere of "belief" where a subjective and illusory principle is already in control and there is an end to the realm of authentic and objective knowledge.[9]

This "dual subordination" of geometry made itself felt as soon as the artists tried to rely on it, using the method of perspective.

Just as the painting of the 15th and 16th centuries appealed to perspective, so the science of the period appealed to geometry. And here the central figure was Galileo. He tried to replace Aristotle's physics, based on the principle of reason, by mechanics which, in Galileo's concept, would be something like the geometry of the physical world. The attempt led Galileo to exactly the same contradiction the artists had come up against: he wanted to devise a science that explained natural phenomena while, in fact, science in his hands was transformed into a description of how changes in those phenomena take place.

The mechanics Galileo constructed on the basis of geometry demanded to remain in the world of appearances. Its actual subject proved to be the establishing of functional dependence between phenomena, i.e. of the law of nature. Leonardo organised the depiction in his paintings with the aid of space; in analagous fashion, Galileo's mechanics established the functional dependence between different phenomena with the aid of time. In neither, however, was an appeal to intelligible entities supposed. Like the artists of the quattrocento, Galileo wished to rely on geometry but avoid the illusionism it entailed.

Let us examine in specific terms, how Galileo restructured the principles of mediaeval physics.

There is a discussion in one of Galileo's works that is highly relevant for an understanding of how he approached the study of motion in the free fall of bodies. In the *Conversations and Mathematical Proofs* Salviati describes how a body thrown into the air moves. His fellow discussant Sagredo then comments:

It seems to me that this reasoning gives sufficient grounds for answering the question the philosophers have raised about the reasons why the natural motion of heavy bodies accelerates. Examining the body thrown upwards I find that the force that the thrower transfers to it gradually decreases; the body continues to rise until it exceeds the contrary force of gravity. As soon as it reaches equilibrium, however, the body ceases to rise and passes through the state of rest during which the initially transferred impulse is not in any way destroyed but only its initial excess over the weight of the body, which caused it to move upwards, is cancelled. Since the increase of this external impulse continues, of which the preponderance of gravity is a consequence, the reverse motion or fall of the body begins. At first this happens slowly as a result of the counter-action of the force transferred to the body, a large part of which it still retains. Since, however, the latter gradually decreases and in ever greater degree is overcome by gravity the gradual acceleration in motion arises.[10]

Sagredo was here presenting Galileo's own conceptions, using which he had originally hoped to give a causal explanation of projectile motion entirely in the spirit of the physics of impetus. However, later Galileo showed why he was forced to abandon this means of explanation. The Aristotelian Simplicio raises objections to Sagredo's words. He points out that only the violent motion upwards can be explained in this way; it is impossible to explain the acceleration of the body which is not thrown upwards, however, but falls from a definite height thereby leaving the state of rest. Sagredo refutes Simplicio's arguments. Nevertheless, although Galileo was striving to overcome the fundamental distinction made in the physics of impetus between natural and violent motions, he resolved the dispute between his imaginary companions in the following way:

I think that now is not a suitable time to take up the question of the causes of acceleration in natural motion: so many different opinions have been stated on this subject by various philosophers. Some attributed it to its drawing closer to the centre; others thought it due to the gradual decrease in the resistant medium; yet others considered a certain influence of the environment that closed in behind the falling body and exerted pressure on it, constantly nudging it as it were. All of these suppositions and many more besides deserve examination but that, however, would be of little help. Now for our Author it will be sufficient if we examine how he studies and presents the capacity for accelerated motion (whatever might be the cause of acceleration). He accepts that the moments of speed, beginning with the transition from a state of rest to motion, proceed and grow in the same simple ratio as time itself ... If it proves that the capacities which are to be proven below are true for both the motion of naturally and accelerated falling, bodies then we shall be able to say that the definition we have given also embraces the said motion of heavy bodies. We shall then be able to say that our proposition concerning the growth in acceleration corresponding to the growth in time, i.e. the continuity of motion, is quite true.[11]

It is, in other words, not essential to find the cause of acceleration in falling bodies, more important is to find the law that describes the phenomenon of acceleration. This thesis is to a certain extent analogous to what we observe in painting. As the result of many years of investigation, Galileo reached the conclusion that it was more feasible for mechanics to establish the law describing how a phenomenon behaves than to determine its intelligible essence, as the physics of impetus had striven to do. Just as the 16th-century artist depicted sensually given phenomena by trying to find a way, with the aid of the rules of perspective, to order them on the canvas so Galileo no longer tried to see phenomena as mere external shells, referring to another intelligible reality. The artist found the means of ordering the sensual fact by using space and its geometrical laws or, to be more precise, by using the rules of measuring spatial relations between objects as they were dependent on their disposition in relation to the artist's eye (and correspondingly to that of the viewer). The scientist, in this case Galileo, found the means of giving an ordered description of a natural process using time. And he found a law, i.e. a means for ordering this phenomenon, without resorting to an intelligible cause – the law of the free fall of bodies: "The even or unified accelerated motion is that whereby in equal intervals of time equal intervals of speed are added after exit from the state of rest".[12]

Here, however, there may arise a legitimate question. For many centuries a trend in investigations was pursued in classical and mediaeval science that did not claim to uncover the essence of motion. Could we not then say, therefore, that Galileo's striving to establish a law of motion instead of disclosing its intelligible cause was just a continuation of this tradition? This idea is all the more tempting since this trend was close to Platonism, and thereby might seem to confirm the thesis of Galileo's Platonism. Astronomy from the time of Eudoxus, and from that of Ptolemy down to Copernicus, was guided by the "principle" of 'saving the phenomenon'. Astronomical theories were thereby regarded as convenient mathematical fictions among which it was better to prefer those which agreed best with the observed facts ('saving the phenomenon'). This principle was based on a characteristic distinction made in classical (and also in the closely-related mediaeval) science between the mathematical and the physical approaches. The mathematician could construct a model and use it to describe the motion of the celestial bodies. His construction, however, did not claim to uncover the real causes of this motion. Only physics, supposed the classical and mediaeval astronomers, could give that answer, not mathematics.

The separation of physics as the science explaining causes from mathematics as the science constructing hypotheses for 'saving the thing' was based on yet another presupposition. Astronomy, it was believed, in which mathematical fictions were applied always dealt with instruments and its conclusions were therefore only approximate.

However, Galileo disputed just these arguments. The approximation of celestial mechanics and of mechanics as a whole was central to Galileo's thinking. In his works he also refuted the other argument linked to the division between physics and mathematics. The two arguments are internally linked. As soon as the same accuracy attainable in mathematical proof could be achieved in experiment there would no longer be any necessity to seek for another means of understanding the physical world than that given by mathematics.

Thus by bringing the subject-matter of mathematics and physics closer together, and by insisting on the necessity of dealing with idealised subjects and not with the subjects of the empirical world, Galileo immediately resolved an entire range of problems.

First, he eliminated the distinction between physics as a science explaining the cause of motion and mathematics as the science enabling that motion to be described, i.e. to formulate its law. Second, he removed the basic difference separating the sciences of mathematics and physics from the art of mechanics. Third, he abolished the traditional idea that mathematics was the science of unchanging entities. He thereby lay the foundations for a new type of mathematics that could describe, as it happens, motion and change and establish the law of change. Fourth, he raised the question whether it was not more important for physics to establish a law describing the process of change in a phenomenon than to seek for the intelligible causes of the latter phenomena.

The condition that made possible the solution of all these problems for Galileo was the experiment. This was not an empirical but an idealised experiment or, putting it another way, it was like a materialised mathematical construction. This entire revolution in principles rests on the assumption that the physical world is mathematical in essence, and therefore the mathematical treatment of all natural realities is quite legitimate. Thus for Galileo there was no longer talk of 'saving the phenomenon' as there had been with Ptolemy. He already did not recognise a "gap" between physical experiment and mathematical theory: for Galileo the mathematical construction did not only 'save phenomena', it expressed its very essence. However, since the empirical picture of the motion of bodies was extremely unlike the mathematical construction, the scientist had to create special, ideal or "mathematical" bodies or a system of bodies. The experiment was precisely that ideal construction where physics and mathematics, according to the concept, should coincide.

However, we have still not resolved one question. If we compare Galileo's mechanics to perspective painting, then where in Galileo is that "reference to the subject" that we saw in Piero della Francesca and Leonardo da Vinci?

It is to be found in the very heart of Galilean mechanics, in the assumption that a real physical process may coincide with an intellectual construction. This is one of the reasons why Galileo spent so much effort trying to prove that his experiments were absolutely accurate; this was why he insisted that once all obstacles had been removed and the experiment cleanly conducted then that would be sufficient to convince everyone of the complete trueness of the law he had thereby established. Certain contemporary historians of science, for example P. Feyerabend, reproach Galileo with resorting to various ruses and tricks in his experiments. The great Italian scientist is thus accused of scientific unscrupulousness.

Yet we believe that the true content of what Feyerabend considered as revealing a lack of scruples in Galileo was understood much more profoundly by Immanuel Kant.

> When Galileo experimented with balls of a definite weight on the inclined plane, when Torricelli caused the air to sustain a weight which he had calculated beforehand to be equal to that of a definite column of water, or when Stahl, at a later period, converted metals into lime, and reconverted lime into metal, by the additon and subtraction of certain elements; a light broke upon all natural philosophers.[13]

To use Kant's terms, Galileo's phenomenalism suggested that we "must not attribute to the object any other properties than those which necessarily followed from that which he had himself, in accordance with his conception, placed in the object."[14]

We may take issue with Kant when he also interprets the mathematics of antiquity in the same spirit. As an interpretation of Galileo's method, however, Kant is right in pointing to the constructivist principle of the latter's physics. It lay precisely in this striving to show that the essence of reality is a mathematical construction, and that certain entities explaining worldly phenomena need not be sought beyond the limits of the empirical world. The functional link should be sought between the elements themselves. It is precisely this approach that Galileo's mechanics shares in common with the principles of Renaissance painting.

*Institute of Philosophy, Russian Academy of Sciences*

NOTES

[1] Galileo Galilei, *Dialogue concerning the two chief world systems, Ptolemaic and Copernican*, Berkely and Los Angeles, 1953, p. 103.
[2] Piero della Francesca, *De prospettiva pingendi*, 1899; tr. in Mastera iskusstva ob iskusstve, Moscow, 1966, Vol. 2, p. 79.
[3] *Ibid.*, Vol. 2, p. 117. *Ashburnham manuscript 1* (Bibliotheque Nationale 2038), reverse of page 20.
[4] Plato, *Republic*, tr. A. D. Lindsay, London 1976 VII, 534a.
[5] *Ibid.*, VI 509–10.
[6] *Ibid.*, VII, 527.
[7] P. A. Florensky, "Reverse perspective" in *Works on Symbolic Systems*, Tartu, 1967, p. 281 (in Russian).
[8] H. Sedlmayr, *Architektur als abbildende Kunst*, Vienna, 1948, p. 21.
[9] See O. Becker, *Grundlagen der Mathematik in geschichtlicher Entwicklung*, Freiburg and Munich, 1964, pp. 100–2 and 104.
[10] Galileo Galilei, Conversations and mathematical proofs relevant to two new branches of science, 1638. Tr. from Russian in *Galileo, Selected works*, Moscow, 1964, vol. 2, pp. 242–3.
[11] *Ibid.*, pp. 243–4.
[12] *Ibid.*, p. 240.
[13] I. Kant, *The Critique of Pure Reason*, New York, 1900, p. 24.
[14] *Ibid.*, p. 23.

MARIA K. PAPATHANASSIOU

# ARCHAEOASTRONOMY IN GREECE: DATA, PROBLEMS AND PERSPECTIVES

Archaeoastronomy, as its name suggests, is the meeting point of astronomy and archaeology through the introduction of astronomical methods into the study of architectural remains. The history of archaeoastronomy in Greece begins in the last two decades of last century with the works of Heinrich Nissen,[1] Sir Norman Lockyer[2] and Sir Francis C. Penrose,[3] who measured and studied the orientation of a number of Greek temples on the Greek mainland, the islands, the Ionian coast and South Italy. This theory dates from 1890 when Lockyer visited Greece and observed the difference in orientation between the old and the new Parthenon as well as the change in direction of the axes of other temples. He then went to Egypt and in the months ending March 1891 he measured the orientation of Egyptian temples. Knowing that some churches were orientated towards sunrise on the feast day of their patron saint, he thought the same might be true for ancient temples. According to his measurements and calculations there are six solsticial Egyptian temples, i.e. temples whose long axis was directed towards the point of the rising Sun at the winter solstice (four temples) or at the summer solstice (two temples).[4]

Taking into account the relation between Egyptian deities and stars according to Egyptian mythology and the available epigraphical and literary evidence, Lockyer also investigated the possibility that other temples were orientated to the heliacal rising or setting of certain prominent stars.[5] He concluded that 7 temples were orientated to Sirius, 12 to $\alpha$ Columbae, 9 to $\alpha$ Centauri, 3 to $\alpha$ Ursae majoris, 7 to $\gamma$ Draconis, 4 to Canopus, 5 to Capella, and 2 to Spica.[6] Meanwhile Penrose measured and studied the orientation of Greek temples and concluded that they also were orientated to the heliacal rising or setting of certain bright stars and Lockyer included this information in the last chapter of his book on Egypt.[7] The most striking outcome of their investigation was their claim to have calculated the exact date (i.e. year and day) of the foundation of these temples. I give here an outline of their method in order to comment on it later.

By using a theodolite they measured very accurately (up to first and



*Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science,* 433-443.

even second minutes of arc) the orientation of the main axis of the temple as well as the profile of the local horizon in the direction of this axis. Then, in the case of a *solar* temple (i.e. a temple whose long axis is orientated to sunrise), given the geographical latitude of the site, the present obliquity of the ecliptic (23°30 ′), and the correction for refraction, Lockyer found that the true horizon sunrise amplitude[8] had a value "x". Then, given this "x", he found a value "y" for the declination, which should also be the value of the obliquity of the ecliptic in the year of the foundation of the temple,[9] the day of the month being that on which the amplitude of the long axis of the temple was the same as that of the sunrise.

The case of *stellar* temples is more complicated. One reason is the effect of the precession of the equinoxes (50.2 ′ ′ of arc per year) on the declination of the stars, which implies that in the course of about 13000 years the amplitude of the rising (or setting) places of a star may in an extreme case vary by the order of 47° (twice the amount of the obliquity of the ecliptic) along the (local) horizon north or south. This change of the amplitude of the rising and setting places of the stars is considered to offer the means of determining the date of foundation of the temples.[10] According to Penrose the years of the dates of the temples can be only considered as approximate but relatively trusted, while the days of the month are less uncertain, as they depend upon the Sun's place, which results immediately from the orientation. The angle of orientation depended primarily on the day of the principal festival, but it would be liable to a "slight modification" for the sake of combining a heliacal star with the sunrise, so that both the rising Sun and the star were observed through the two central columns of the front collonade in the adytum of the temple.[11]

Another reason is that the angular distance between the Sun and the star when the star first becomes visible heliacally depends on the brightness of the star. Lockyer used the value of 11° for the depression of the Sun[12] in his calculations for Egyptian temples. But Penrose based on his own observations came to the conclusion that a first magnitude star in fair average weather in Greece or Italy could be seen, when rising heliacally, at an altitude of 3°, the Sun being 10° below the horizon; that a second magnitude star should require an altitude of 3°30 ′ with the Sun 11° depressed; but that of a third magnitude star the Sun's depression should not be less than 13°.[13]

In his search for a star to be associated with a given temple Lockyer used a precessional globe as a guide in order to obtain a first approximation within 1° of declination.[14] Penrose used an astrolabe he had made himself, the fixed part of which was a stereographic projection of the celestial sphere (taken on the pole of the ecliptic, using a mean obliquity and showing the pole of the equator, right ascension, hour lines and parallels of declination, and the principal available stars, taking 1850 as the standard year). If by this astrolabe he found a bright star as heliacally rising or setting, then he calculated the "exact" foundation date.[15]

There follows a composite list of the results deduced by Penrose and adopted Lockyer, in which 34 Greek temples are grouped according to their orientation towards the heliacal rising (R) or setting (S) of some prominent stars; the last column gives their supposed foundation date.[16]

| STAR | TEMPLE | AREA | PHAEN. | DATE |
|---|---|---|---|---|
| Pleiades | arch. temple of Athena | Athens | R | 2020 B.C. |
| | Asclepieium | Epidaurus | R | 1370 |
| | Hekatompedon | Athens | R | 1495 |
| | Temple of Bacchus | Athens | R | 1180 |
| | Temple of Athena | Sounion | S | 845 |
| Antares | old Heraeum | Argos | R | 1830, Octo. 24 |
| | Dionysos "en limnais" (upper temple) | Athens | S | 1700, June, 20 |
| | old Erechtheum | Athens | S | 1070 |
| | Temple in... | Korinth | S | 770 |
| | Temple on the mountain | Aegina | S | 630 |
| | Zeus Panhellenios | Aegina | S | 630 |
| Arcturus | Dionysos "en limnais" (lower temple) | Athens | S | 850, July, 19. |
| Spica | Heraeum | Olympia | R | 1445 |
| | Nike Apteros | Athens | S | 1130 |
| | Themis | Rhamnus | R | 1092 |
| | Temple in ... | Athens | R | 780, Sept. 23. |
| | Nemesis | Rhamnus | R | 747 |
| | Apollo (east. doorway) | Bassae | R | 728 |
| | Artemis | Ephesos | R | 715 |
| | Theseum | Athens | R | 470, Octo. 5. |
| | Zeus Olympios (new) | Athens | S | 174, March 27. |
| $\alpha$ Arietis | Athena (old temple) | Tegea | R | 1660 |
| | Athena (new temple) | Tegea | R | 1140 |
| | Zeus Olympios (old) | Athens | R | 790 |
| | Asclepieium | Athens | R | 720, April 5. |
| | Hera? | Plataea | S | 650 |

(continued)

| STAR | TEMPLE | AREA | PHAEN. | DATE |
|---|---|---|---|---|
| | Zeus | Megalopolis | S | 605 |
| | Temple at the port | Aegina | S | 580 |
| | Temple on the Akropol. (east. doorway) | Mykenae | R | 540 |
| | Later Erechtheum | Athens | R | 445, April 9. |
| | Metroon | Olympia | S | 360, Octo. 9. |
| | Temple of Bacchus (new) | Athens | R | 340, April 23. |
| β Lupi | Apollo (old temple) | Delphi | S | 970, March 1. |
| | Apollo (new temple) | Delphi | S | 630, March 1. |

It is evident that most of these dates are strikingly earlier than those usually accepted by archaeologists and deduced from literary sources and architectural elements (although it is known that temples are sometimes founded on remains of more ancient temples, whose orientation they may follow).

In 1907 Nissen published his *Orientation*, in which he studied the orientation of 113 Greek temples on the Greek mainland and the islands, as well as on the Ionian coast and in South Italy. Nissen based his study on Penrose's measurements of the bearing of the main axis of these temples but also on the available related philological evidence regarding festivals celebrated in honour of the deities to whom these temples were sacred. According to him, only 11 out of the 113 temples are *stellar*, namely the following:[17]

| STAR | TEMPLE | AREA | DATE |
|---|---|---|---|
| Castor | Apollo | Thera | 600 v. Chr. |
| | Persephone | Lokri | 500 |
| | Amphiaraos | Oropos | 200 |
| | Triptolemos | Eleusis | 200 |
| Pollux | Kabiren | Samothrake | 800 |
| | Apollo | Didyma | 1000 |
| | Hekate | Lagina | 1000 |
| Capella | Apollo | Metapont | 550 |
| | Dionysos | Athens | 750 |
| Arktur? | Aphrodite | Ancona | 300? |
| | Kabiren | Samothrake | 800 |

Nissen was of the opinion that the stellar orientation of certain Greek temples was possibly due to a foreign influence – probably Egyptian – on Greek cult. He thought that when there is a foreign influence either a foreign deity is nationalised, or it remains a foreigner, maintaining the special elements of its ritual, temple orientation included; consequently, by studying the orientation of the temples in relation to all the other evidence, we can deduce significant results for the travels of gods and their relations to the cult of any town. The rise of the ancient "polis" (= city) as a civil institution was, he thought, especially favourable to the solar orientation of the temples. He also pointed out that the deciphering and the estimation of the information hidden in ancient temple ruins must be the joint work of astronomy, archaeology, and oriental and classical philology.[18]

A comparison of Nissen's results with those of Penrose reveals a great difference both in the classification of the temples and their dating. Firstly, Penrose's 28 *stellar* temples are now classified as *solar* by Nissen. Secondly, the foundation dates according to Nissen differ greatly from those calculated by Penrose, sometimes by more than 800 years. Consequently, there are crucial questions to be answered. We must account for these differences, and we must explain how the same data can lead to very different conclusions, if we wish archaeoastronomy to be considered a rigorous science rather than a fancy.

This problem was posed by the American archaeologist William Bell Dinsmoor, who in 1939 published an article under the title *Astronomy and Archaeology*. As he states in the abstract of his paper,

> The astronomical calculations based on the data led in most cases to such fantastic dates for the temples that more recent opinion has tended to discredit or ignore the theory. But the fragments of literary evidence and the accumulated measurements of the actual orientations, taken together, furnish sufficient proof of the existence of the general principle. The following article endeavors to show that the error lay, not in the theory itself, but in the deduction of results from incomplete data. In the absence of detailed knowledge of the archaeological conditions, of the religious festivals, and of the various local lunar calendars, investigators felt obliged to check the observed orientations with reference to assumed astronomical factors such as the heliacal risings of stars – factors which could never have been employed by peoples observing lunar calendars – with results inevitably grotesque.[19]

Dinsmoor gives the following table showing the divergence of astronomical and archaeological evidence for Greek temple dates:[20]

| TEMPLE | PENROSE'S Astr. Date (B.C.) | ARCHAEOLOGICAL DATE OF BEGINNING | |
|---|---|---|---|
| | | Visible temple | Earlier Foundations |
| Parthenon | 1495, April 26 | 447 B.C. | c. 490 B.C. |
| Peisistratid | 2020, April 20 | c. 535-525 | ? |
| Erechtheum | 447, Sept. 2 | c. 421 | c. 479 |
| Nike temple | 1130, March 17 | c. 427 | c. 479 |
| Olympieum | 174, March 27 | 174 | c. 525-515 |
| "Theseum" | 466, Oct. 6 | 450-440 | ? |
| Old Nemesis Rhamn. | 1092, Sept. 17 | c. 530-490 | ? |
| New Nemesis Rhamn. | 747, Sept. 22 | c. 438-421 | - |
| Heraeum, Olympia | 1445, Sept. 12 | c. 610-600 | c. 610-600 |
| Zeus, Olympia | 790, April 6 | c. 472-460 | - |

According to Dinsmoor we may approach the problem by devising a pseudo-algebraic equation somewhat in the following form, $X=Ar+R+C+As$, wherein X represents the (usually unknown) date, Ar the archaeological evidence as to the date, R the religious evidence as to the cult, C the artificial astronomy of the local calendar, and As the natural astronomical observations. His opinion is that if more than one or two of the elements are unknown, there is little hope of finding a solution.[21]

Taking into consideration information concerning the Athenian lunar calendar in relation to the old "octaeteris" and the newly introduced metonic cycle of 19 years, Dinsmoor calculated the exact Athenian New Year Day for the years of the beginning of the construction (or reconstruction) of a number of Athenian temples for all of which he had the necessary information. I.e. by using the archaeological evidence and the Athenian sacred calendar (when Athenians celebrated festivals in honour of a deity), he calculated for a number of years the exact day of the festival of the deity to whom the temple was sacred, and then by checking the azimuth of the rising Sun on these dates against the orientation of the axis of the temple he decided the exact day of the foundation of the temple. In the following table "Preceding N.Y." (or "Following N.Y.") means the New Year Day of the year shown in the first column (or the following year) according to the available calendric data:[22]

| DATE (B.C.) | TEMPLE | NEW YEAR (=N.Y.) |
|---|---|---|
| 529, Sept. 7, | Peisistratid temple. | Preceding N.Y. = July 9 |
| 521, April 3, | Olympieum. | Following N.Y. = June 12 |
| 514, Sept. 26, | Old Nemesis Rhamnus | Preceding N.Y. = July 25 |
| 488, Aug. 31, | Older Parthenon. | Preceding N.Y. = Aug. 5 |
| 478, Sept. 15, | Erechtheum shrine. | Preceding N.Y. = July 16 |
| 449, Oct. 17, | "Theseum." | Preceding N.Y. = June 22 |
| 436, Sept. 30, | New Nemesis Rhamnus | Preceding N.Y. = July 29 |
| 427, Aug. 16 | Nike temple. | Preceding N.Y. = July 20 |

It is true that Nissen had likewise tried to explain the data on the basis of literary evidence regarding cult and calendar; but in this time the available evidence was unreliable. What I wish to draw attention to is Dinsmoor's statement at the very beginning of the main text of his article: "I know nothing of astronomy ... My only excuse for intruding upon such a field is the precedent established by Sir Norman Lockyer, who wrote upon archaeology from the astronomer's point of view, though his archaeology, like my astronomy, was purely empirical."[23] Dinsmoor thought that his astronomical knowledge, though "primitive", might lead him closer to the ideas of the ancients and also restrain him from the temptation to overshoot the mark.[24] Of course, his astronomical knowledge was sufficient for his calculations, in accordance with Lockyer's wish that every archaeologist should know a little astronomy.

Let us now comment on the views of the pioneers of Greek archaeoastronomy. Dinsmoor's "equation" is only a sum of various items of information taken from different sources. Although he actually means that "X" is a "function" of the other quantities, his sources are not like mathematical variables, and his equation is but a set of data. Consequently, X is not a "function" of these data, and "C" is important only in cases where we deal with classical Greek temples, especially the Athenian ones, for which we have sufficient calendric information. For this reason I think, in spite of his admirable approach to the problem, it is possible that Dinsmoor may have been wrong in his belief that heliacal risings of stars could never have been employed by peoples observing lunar calendars.

I incline to the opinion that it is not out of the question that ancient people might have used a stellar orientation for a sanctuary. But Lockyer and Penrose were misled by their own very accurate measurements (up to first and even second minutes of arc). By calculating the "exact" foundation date according to their actual measurements, they did not allow for the possibility of an "error" plus or minus several minutes of arc (possibly half a degree) which would correspond to several centuries, the actual interval depending on the methods and the inaccuracy of the instruments used in antiquity for the tracing of the desired direction, and the resultant deviation of the long axis of the temple from its "ideal" mathematical line of orientation. Moreover, the variations of the foundation dates of Greek temples due only to different values of the Sun's depression (the other factors considered as fixed) as given by Penrose, show his inability to determine the true date among his other "exact" dates, differing by 75 up to 400 years![25]

I would also like to point out, that the four pioneers of Greek archaeoastronomy only dealt with Classical Greece, ignoring Greek "Prehistory", which includes Neolithic Greece and the Cycladic,[26] Minoan and Mycenean civilizations with their splendid architectural remains. A possible excuse for Lockyer, Penrose and Nissen would be that many excavations of prehistoric sites in Greece were ongoing at the end of the 19th and the beginning of the 20th century. But when Dinsmoor wrote his article in 1938 the most important of these sites had been excavated and the results of these excavations were already published. It is possible that Dinsmoor did not consider these prehistoric civilisations as an early stage of Greek history, because it was only in 1952 that Michael Ventris deciphered the Linear B clay tablets, showing that the language of the texts written on them was Greek written in a syllabic system.[27] Taking into account the unbroken sequence of the cultural development of Greece from its Prehistory up to Classical times, together with the Minoan–Mycenean origin of Greek religion,[28] I am rather of the opinion that it is time for us likewise to trace the "prehistory" of archaic and classical temple orientation.[29]

I believe that the introduction of a new cult never eliminates the old one; on the contrary, the new usually adopts many ritualistic elements of the old, simply changing the names so as to make assimilation easier. What could be better proof of this than the official (or unofficial) survival of many ancient pagan rituals (mostly seasonal festivals closely linked with the annual path of the Sun) in the Greek orthodox church? The deep roots of religious beliefs and rituals in Greece reach back to its most remote past. I do not doubt that in the very remote past people were better observers of natural phenomena than we are today (apart from the specialists in the field) even though they lacked knowledge of spherical astronomy. Heliacal risings and settings of prominent stars and constellations are mentioned in Hesiod's *Works and days* as marking the beginning or the end of particular agricultural work.[20] This may mean that most calendric information about seasonal changes was provided by the annual path of the Sun in relation to prominent stars or constellations rather than by specific lunar or lunisolar calendars. Otherwise, why did not Hesiod mention them? He merely refers to superstitions related to the days (of the "age" of the Moon) of a lunar month as favourable or unfavourable for certain activities.[31]

I think in the case of Greece it must be recognised that two important conditions must be fulfilled if a rigorous and respected archaeo-astronomical research is to be possible. Firstly, we need measurements of the bearings of the long axes of as many as possible representative monuments of a given period, and then of different periods, in each region, together with those of the horizon profile in the direction of the axis. Secondly, we must recognise that while any statistical information to be taken from these measurements after the necessary calculations is factual, any possible astronomical explanation based on it must be carefully checked against the available related archaeological evidence. We must avoid overshooting the mark!

*University of Athens, Greece*

NOTES

[1] Nissen, H., 1885 "Über Tempel-Orientierung" (4th article), *Rheinisches Museum für Philologie*, N.F. 40, pp. 329–370. *1907 Orientation; Studien zur Geschichte der Religion*, Heft 2, Weidmann, Berlin.
[2] Lockyer, J.N., *1894 The Dawn of Astronomy; A Study of the Temple Worship and Mythology of the Ancient Egyptians*, Cassell & Co. [MIT Pr. 1964]. 1909 "The uses and dates of ancient temples", *Nature*, **80**, pp. 340–344.
[3] Penrose, F.C., 1892 "A preliminary statement on an investigation of the dates of some of the Greek temples as derived from their orientation", *Proc. Soc. Antiquaries* (2nd series) **14**, pp. 59–65. 1893 "On the results of an examination of the orientation of a number of Greek temples", *Phil. Trans. of the Royal Society* (Series A) **184**, pp. 379–384. 1897 "On the orientation of certain Greek temples and the dates of their foundation derived from astronomical considerations" (a supplement to the paper published in 1893), *Phil. Trans. of the Royal Society* (Series A) **190**, pp. 43–65.
[4] Lockyer, J.N., 1984, *op cit.* (Note 2), p. 78.

[5] The brightness of the Sun moving about 1° per day on the ecliptic prevents us from seeing its neighbouring stars. The first morning after a period of disappearance we can perceive a star in the dawn near the horizon before the Sun rises, is the day of its heliacal rising.
[6] Lockyer, J. N., 1894 *op. cit.* (Note 2), pp. 194, 305–313.
[7] *Ibid.*, pp. 416–424.
[8] *Ibid.*, pp. 46–47: "The amplitude of a body on the horizon is its distance north and south of the east and west points; it is always measured to the nearest of these two latter points, so that its greatest value can never exceed 90° ... We can say then that a star of a certain declination will rise or set at such an *azimuth*, if we reckon from the N. point of the horizon, or at such an amplitude if we reckon from the *equator*. This will apply to both north and south declinations."
[9] *Ibid.*, pp. 117–119. The change of the obliquity of the ecliptic is very small, so that in 6000 years the position of the Sun at sunrise and sunset on the horizon may vary by about 1° (pp. 123–124).
[10] Lockyer, J. N., 1894 *op. cit.* (Note 2), pp. 124, 126. Penrose, F. C., 1897 *op. cit.* (Note 3), p. 43.
[11] Penrose, F.C., 1897 *op. cit.* (Note 3), pp. 46, 53, 55. For example the long axis of the archaic Apollo temple at Didyma in Ionia (Asia Minor) destroyed in 494 B.C. by the Persians displays a difference of 1°45′ in relation to the long axis of the hellenistic temple founded in 280 B.C. on the remains of the archaic one. Hans Waltenberg and Wolfgang Gleissberg ["Das Rätsel von Didyma und seine astronomische Lösung" *Sterne und Weltraum*, 1968 8/9, pp. 217–20.] calculated that the archaic temple was founded in about 700 B.C. and its orientation was such that the heliacal rising of $\alpha$ and $\beta$ Geminorum could be seen in the inner sanctuary through the two central columns of the colonnade of the front side of the temple. Because of the precession of equinoxes, in 280 B.C. it would have been necessary for the long axis of the temple under construction to be tilted by 1°45′ if the same phenomenon was to be observed in the new inner sanctuary.
[12] Ptolemy used for stars of the first magnitude a Sun's depression below the true horizon of 11°, if the star and the Sun were on the same horizon (heliacal rising); if on opposite horizons (heliacal setting), a depression of 7°. For stars of the second magnitude his values of the Sun's depression were 14° and 8°30′ respectively [Ideler, I., 1820 "Mémoire sur le calendrier du Ptolemée", pp. 3–11 (here 10), in: Claude Ptolemée, *Apparitions des fixes et annonces* (trad. du grec. par M. Abbé Halma), Paris].
[13] Penrose, F. C., 1897 *op. cit.* (Note 3), p. 45.
[14] Lockyer, J. N., 1894 *op. cit.* (Note 2), pp. 129–131.
[15] Penrose, F. C., 1897 *op. cit.* (Note 3), pp. 46–47.
[16] *Ibid.*, pp. 44, 48–50, 65. Lockyer, J. N., 1897 *op. cit.* (Note 2), pp. 419–420.
[17] Nissen, H., 1907 *op. cit.* (Note 1), p. 158.
[18] *Ibid.*, p. 161. According to Nissen the orientation of other temples towards west, north and south, for which there is no early literary evidence, was also due to a foreign influence.
[19] Dinsmoor, W. Bell, 1939 "Archaeology and Astronomy" *Proceedings of the American Philosophical Society*, **80**, pp. 95–173 (here p. 95).
[20] *Ibid.*, table I, p. 104.
[21] *Ibid.*, p. 119.
[22] *Ibid.*, p. 168.
[23] *Ibid.*, p. 95.

[24] *Ibid.*, p. 96.
[25] Penrose, F. C., 1987 *op. cit.*, pp. 49–50, 52–54, 56–57, 59, 65.
[26] For example in the island of Naxos there are some rocks with very ancient engraved symbols – possibly of astronomical meaning [Bardanis, Michael, 1966–67 "Elements of the astronomical knowledge of the prehistoric people of Cyclades islands" (in Greek) *Deltion of the Library N. N. Glezou*, **2**, pp. 71–80]. It is pity that some of them are no longer in situ but have been transferred to the local Museum. They have thereby lost their importance, except as art objects, because we cannot now make measurements of their places, the distances between them, and their relations to the local horizon. As there is no official catalogue of the rocks and their original locations, the only possible way in which archaeoastronomers could discover this information would be by questioning the older inhabitants.
[27] Chadwick, John, 1970 *The decipherment of Linear B*, Cambridge, pp. 81–100. A support to Ventris's decipherment offers the thesis of the architect-archaeologist Tessi Sali-Axioti [1989 *The Mycenean Megaron according to Linear B tablets* (in Greek), Athens] in which she proves that the architectural terms on the tablets of Linear B can be fully identified with the architectural remains of Mycenean megaron (= palace) and she achieves a full representation of its construction; these architectural terms are the same as those we later find in the Homeric poems, Greek literature and architecture.
[28] Nilsson, Martin P., [2]1950 *The Minoan-Mycenean Religion and its Survival in Greek Religion*, Lund.
[29] The late S. Marinatos, an archaeologist, had pointed out the striking North-South orientation (with only small deviations) of the Minoan palaces in Crete [1st Greek Astron. Confer., Aug. 1971], but no one showed any interest.
[30] Hesiod, *Works and Days*, v. 383–7, 414–9, 479–80, 564–7, 571–2, 597–8, 609–11, 615–21, 663–5.
[31] *Ibid.*, v. 765–828.

# INDEX OF NAMES

14. R.S. Cohen and M.W. Wartofsky (eds.): *Methodological and Historical Essays in the Natural and Social Sciences.* Proceedings of the Boston Colloquium for the Philosophy of Science, 1969/72, Part II. [Synthese Library 60] 1974
ISBN 90-277-0392-2; Pb 90-277-0378-7
15. R.S. Cohen, J.J. Stachel and M.W. Wartofsky (eds.): *For Dirk Struik.* Scientific, Historical and Political Essays in Honor of Dirk J. Struik. [Synthese Library 61] 1974 ISBN 90-277-0393-0; Pb 90-277-0379-5
16. N. Geschwind: *Selected Papers on Language and the Brains.* [Synthese Library 68] 1974 ISBN 90-277-0262-4; Pb 90-277-0263-2
17. B.G. Kuznetsov: *Reason and Being.* Translated from Russian. Edited by C.R. Fawcett and R.S. Cohen. 1987 ISBN 90-277-2181-5
18. P. Mittelstaedt: *Philosophical Problems of Modern Physics.* Translated from the revised 4th German edition by W. Riemer and edited by R.S. Cohen. [Synthese Library 95] 1976 ISBN 90-277-0285-3; Pb 90-277-0506-2
19. H. Mehlberg: *Time, Causality, and the Quantum Theory.* Studies in the Philosophy of Science. Vol. I: *Essay on the Causal Theory of Time.* Vol. II: *Time in a Quantized Universe.* Translated from French. Edited by R.S. Cohen. 1980 Vol. I: ISBN 90-277-0721-9; Pb 90-277-1074-0
Vol. II: ISBN 90-277-1075-9; Pb 90-277-1076-7
20. K.F. Schaffner and R.S. Cohen (eds.): *PSA 1972.* Proceedings of the 3rd Biennial Meeting of the Philosophy of Science Association (Lansing, Michigan, Fall 1972). [Synthese Library 64] 1974
ISBN 90-277-0408-2; Pb 90-277-0409-0
21. R.S. Cohen and J.J. Stachel (eds.): *Selected Papers of Léon Rosenfeld.* [Synthese Library 100] 1979 ISBN 90-277-0651-4; Pb 90-277-0652-2
22. M. Čapek (ed.): *The Concepts of Space and Time.* Their Structure and Their Development. [Synthese Library 74] 1976
ISBN 90-277-0355-8; Pb 90-277-0375-2
23. M. Grene: *The Understanding of Nature.* Essays in the Philosophy of Biology. [Synthese Library 66] 1974 ISBN 90-277-0462-7; Pb 90-277-0463-5
24. D. Ihde: *Technics and Praxis.* A Philosophy of Technology. [Synthese Library 130] 1979 ISBN 90-277-0953-X; Pb 90-277-0954-8
25. J. Hintikka and U. Remes: *The Method of Analysis.* Its Geometrical Origin and Its General Significance. [Synthese Library 75] 1974
ISBN 90-277-0532-1; Pb 90-277-0543-7
26. J.E. Murdoch and E.D. Sylla (eds.): *The Cultural Context of Medieval Learning.* Proceedings of the First International Colloquium on Philosophy, Science, and Theology in the Middle Ages, 1973. [Synthese Library 76] 1975
ISBN 90-277-0560-7; Pb 90-277-0587-9
27. M. Grene and E. Mendelsohn (eds.): *Topics in the Philosophy of Biology.* [Synthese Library 84] 1976 ISBN 90-277-0595-X; Pb 90-277-0596-8
28. J. Agassi: *Science in Flux.* [Synthese Library 80] 1975
ISBN 90-277-0584-4; Pb 90-277-0612-3

29. J.J. Wiatr (ed.): *Polish Essays in the Methodology of the Social Sciences.* [Synthese Library 131] 1979 ISBN 90-277-0723-5; Pb 90-277-0956-4
30. P. Janich: *Protophysics of Time.* Constructive Foundation and History of Time Measurement. Translated from the 2nd German edition. 1985 ISBN 90-277-0724-3
31. R.S. Cohen and M.W. Wartofsky (eds.): *Language, Logic, and Method.* 1983 ISBN 90-277-0725-1
32. R.S. Cohen, C.A. Hooker, A.C. Michalos and J.W. van Evra (eds.): *PSA 1974.* Proceedings of the 4th Biennial Meeting of the Philosophy of Science Association. [Synthese Library 101] 1976 ISBN 90-277-0647-6; Pb 90-277-0648-4
33. G. Holton and W.A. Blanpied (eds.): *Science and Its Public.* The Changing Relationship. [Synthese Library 96] 1976 ISBN 90-277-0657-3; Pb 90-277-0658-1
34. M.D. Grmek, R.S. Cohen and G. Cimino (eds.): *On Scientific Discovery.* The 1977 Erice Lectures. 1981 ISBN 90-277-1122-4; Pb 90-277-1123-2
35. S. Amsterdamski: *Between Experience and Metaphysics.* Philosophical Problems of the Evolution of Science. Translated from Polish. [Synthese Library 77] 1975 ISBN 90-277-0568-2; Pb 90-277-0580-1
36. M. Marković and G. Petrović (eds.): *Praxis.* Yugoslav Essays in the Philosophy and Methodology of the Social Sciences. [Synthese Library 134] 1979 ISBN 90-277-0727-8; Pb 90-277-0968-8
37. H. von Helmholtz: *Epistemological Writings.* The Paul Hertz / Moritz Schlick Centenary Edition of 1921. Translated from German by M.F. Lowe. Edited with an Introduction and Bibliography by R.S. Cohen and Y. Elkana. [Synthese Library 79] 1977 ISBN 90-277-0290-X; Pb 90-277-0582-8
38. R.M. Martin: *Pragmatics, Truth and Language.* 1979 ISBN 90-277-0992-0; Pb 90-277-0993-9
39. R.S. Cohen, P.K. Feyerabend and M.W. Wartofsky (eds.): *Essays in Memory of Imre Lakatos.* [Synthese Library 99] 1976 ISBN 90-277-0654-9; Pb 90-277-0655-7
40. B.M Kedrov and V. Sadovsky (eds.): Current Soviet Studies in the Philosophy of Science. (In prep.) ISBN 90-277-0729-4
41. M. Raphael: Theorie des geistigen Schaffens aus marxistischer Grundlage. (In prep.) ISBN 90-277-0730-8
42. H.R. Maturana and F.J. Varela: *Autopoiesis and Cognition.* The Realization of the Living. With a Preface to 'Autopoiesis' by S. Beer. 1980 ISBN 90-277-1015-5; Pb 90-277-1016-3
43. A. Kasher (ed.): *Language in Focus: Foundations, Methods and Systems.* Essays in Memory of Yehoshua Bar-Hillel. [Synthese Library 89] 1976 ISBN 90-277-0644-1; Pb 90-277-0645-X
44. T.D. Thao: *Investigations into the Origin of Language and Consciousness.* 1984 ISBN 90-277-0827-4

45. A. Ishimoto (ed.): Japanese Studies in the History and Philosophy of Science. (In prep.) ISBN 90-277-0733-3
46. P.L. Kapitza: *Experiment, Theory, Practice.* Articles and Addresses. Edited by R.S. Cohen. 1980 ISBN 90-277-1061-9; Pb 90-277-1062-7
47. M.L. Dalla Chiara (ed.): *Italian Studies in the Philosophy of Science.* 1981 ISBN 90-277-0735-9; Pb 90-277-1073-2
48. M.W. Wartofsky: *Models.* Representation and the Scientific Understanding. [Synthese Library 129] 1979 ISBN 90-277-0736-7; Pb 90-277-0947-5
49. T.D. Thao: *Phenomenology and Dialectical Materialism.* Edited by R.S. Cohen. 1986 ISBN 90-277-0737-5
50. Y. Fried and J. Agassi: *Paranoia.* A Study in Diagnosis. [Synthese Library 102] 1976 ISBN 90-277-0704-9; Pb 90-277-0705-7
51. K.H. Wolff: *Surrender and Cath.* Experience and Inquiry Today. [Synthese Library 105] 1976 ISBN 90-277-0758-8; Pb 90-277-0765-0
52. K. Kosík: *Dialectics of the Concrete.* A Study on Problems of Man and World. 1976 ISBN 90-277-0761-8; Pb 90-277-0764-2
53. N. Goodman: *The Structure of Appearance.* [Synthese Library 107] 1977 ISBN 90-277-0773-1; Pb 90-277-0774-X
54. H.A. Simon: *Models of Discovery* and Other Topics in the Methods of Science. [Synthese Library 114] 1977 ISBN 90-277-0812-6; Pb 90-277-0858-4
55. M. Lazerowitz: *The Language of Philosophy.* Freud and Wittgenstein. [Synthese Library 117] 1977 ISBN 90-277-0826-6; Pb 90-277-0862-2
56. T. Nickles (ed.): *Scientific Discovery, Logic, and Rationality.* 1980 ISBN 90-277-1069-4; Pb 90-277-1070-8
57. J. Margolis: *Persons and Mind.* The Prospects of Nonreductive Materialism. [Synthese Library 121] 1978 ISBN 90-277-0854-1; Pb 90-277-0863-0
58. G. Radnitzky and G. Andersson (eds.): *Progress and Rationality in Science.* [Synthese Library 125] 1978 ISBN 90-277-0921-1; Pb 90-277-0922-X
59. G. Radnitzky and G. Andersson (eds.): *The Structure and Development of Science.* [Synthese Library 136] 1979 ISBN 90-277-0994-7; Pb 90-277-0995-5
60. T. Nickles (ed.): *Scientific Discovery.* Case Studies. 1980 ISBN 90-277-1092-9; Pb 90-277-1093-7
61. M.A. Finocchiaro: *Galileo and the Art of Reasoning.* Rhetorical Foundation of Logic and Scientific Method. 1980 ISBN 90-277-1094-5; Pb 90-277-1095-3
62. W.A. Wallace: *Prelude to Galileo.* Essays on Medieval and 16th-Century Sources of Galileo's Thought. 1981 ISBN 90-277-1215-8; Pb 90-277-1216-6
63. F. Rapp: *Analytical Philosophy of Technology.* Translated from German. 1981 ISBN 90-277-1221-2; Pb 90-277-1222-0
64. R.S. Cohen and M.W. Wartofsky (eds.): *Hegel and the Sciences.* 1984 ISBN 90-277-0726-X
65. J. Agassi: *Science and Society.* Studies in the Sociology of Science. 1981 ISBN 90-277-1244-1; Pb 90-277-1245-X

66. L. Tondl: *Problems of Semantics.* A Contribution to the Analysis of the Language of Science. Translated from Czech. 1981 ISBN 90-277-0148-2; Pb 90-277-0316-7
67. J. Agassi and R.S. Cohen (eds.): *Scientific Philosophy Today.* Essays in Honor of Mario Bunge. 1982 ISBN 90-277-1262-X; Pb 90-277-1263-8
68. W. Krajewski (ed.): *Polish Essays in the Philosophy of the Natural Sciences.* Translated from Polish and edited by R.S. Cohen and C.R. Fawcett. 1982 ISBN 90-277-1286-7; Pb 90-277-1287-5
69. J.H. Fetzer: *Scientific Knowledge.* Causation, Explanation and Corroboration. 1981 ISBN 90-277-1335-9; Pb 90-277-1336-7
70. S. Grossberg: *Studies of Mind and Brain.* Neural Principles of Learning, Perception, Development, Cognition, and Motor Control. 1982 ISBN 90-277-1359-6; Pb 90-277-1360-X
71. R.S. Cohen and M.W. Wartofsky (eds.): *Epistemology, Methodology, and the Social Sciences.* 1983. ISBN 90-277-1454-1
72. K. Berka: *Measurement.* Its Concepts, Theories and Problems. Translated from Czech. 1983 ISBN 90-277-1416-9
73. G.L. Pandit: *The Structure and Growth of Scientific Knowledge.* A Study in the Methodology of Epistemic Appraisal. 1983 ISBN 90-277-1434-7
74. A.A. Zinov'ev: *Logical Physics.* Translated from Russian. Edited by R.S. Cohen. 1983 ISBN 90-277-0734-0
    *See also* Volume 9.
75. G-G. Granger: *Formal Thought and the Sciences of Man.* Translated from French. With and Introduction by A. Rosenberg. 1983 ISBN 90-277-1524-6
76. R.S. Cohen and L. Laudan (eds.): *Physics, Philosophy and Psychoanalysis.* Essays in Honor of Adolf Grünbaum. 1983 ISBN 90-277-1533-5
77. G. Böhme, W. van den Daele, R. Hohlfeld, W. Krohn and W. Schäfer: *Finalization in Science.* The Social Orientation of Scientific Progress. Translated from German. Edited by W. Schäfer. 1983 ISBN 90-277-1549-1
78. D. Shapere: *Reason and the Search for Knowledge.* Investigations in the Philosophy of Science. 1984 ISBN 90-277-1551-3; Pb 90-277-1641-2
79. G. Andersson (ed.): *Rationality in Science and Politics.* Translated from German. 1984 ISBN 90-277-1575-0; Pb 90-277-1953-5
80. P.T. Durbin and F. Rapp (eds.): *Philosophy and Technology.* [*Also* Philosophy and Technology Series, Vol. 1] 1983 ISBN 90-277-1576-9
81. M. Marković: *Dialectical Theory of Meaning.* Translated from Serbo-Croat. 1984 ISBN 90-277-1596-3
82. R.S. Cohen and M.W. Wartofsky (eds.): *Physical Sciences and History of Physics.* 1984. ISBN 90-277-1615-3
83. É. Meyerson: *The Relativistic Deduction.* Epistemological Implications of the Theory of Relativity. Translated from French. With a Review by Albert Einstein and an Introduction by Milić Čapek. 1985 ISBN 90-277-1699-4

84. R.S. Cohen and M.W. Wartofsky (eds.): *Methodology, Metaphysics and the History of Science.* In Memory of Benjamin Nelson. 1984 ISBN 90-277-1711-7
85. G. Tamás: *The Logic of Categories.* Translated from Hungarian. Edited by R.S. Cohen. 1986 ISBN 90-277-1742-7
86. S.L. de C. Fernandes: *Foundations of Objective Knowledge.* The Relations of Popper's Theory of Knowledge to That of Kant. 1985 ISBN 90-277-1809-1
87. R.S. Cohen and T. Schnelle (eds.): *Cognition and Fact.* Materials on Ludwik Fleck. 1986 ISBN 90-277-1902-0
88. G. Freudenthal: *Atom and Individual in the Age of Newton.* On the Genesis of the Mechanistic World View. Translated from German. 1986 ISBN 90-277-1905-5
89. A. Donagan, A.N. Perovich Jr and M.V. Wedin (eds.): *Human Nature and Natural Knowledge.* Essays presented to Marjorie Grene on the Occasion of Her 75th Birthday. 1986 ISBN 90-277-1974-8
90. C. Mitcham and A. Hunning (eds.): *Philosophy and Technology II.* Information Technology and Computers in Theory and Practice. [*Also* Philosophy and Technology Series, Vol. 2] 1986 ISBN 90-277-1975-6
91. M. Grene and D. Nails (eds.): *Spinoza and the Sciences.* 1986 ISBN 90-277-1976-4
92. S.P. Turner: *The Search for a Methodology of Social Science.* Durkheim, Weber, and the 19th-Century Problem of Cause, Probability, and Action. 1986. ISBN 90-277-2067-3
93. I.C. Jarvie: *Thinking about Society.* Theory and Practice. 1986 ISBN 90-277-2068-1
94. E. Ullmann-Margalit (ed.): *The Kaleidoscope of Science.* The Israel Colloquium: Studies in History, Philosophy, and Sociology of Science, Vol. 1. 1986 ISBN 90-277-2158-0; Pb 90-277-2159-9
95. E. Ullmann-Margalit (ed.): *The Prism of Science.* The Israel Colloquium: Studies in History, Philosophy, and Sociology of Science, Vol. 2. 1986 ISBN 90-277-2160-2; Pb 90-277-2161-0
96. G. Márkus: *Language and Production.* A Critique of the Paradigms. Translated from French. 1986 ISBN 90-277-2169-6
97. F. Amrine, F.J. Zucker and H. Wheeler (eds.): *Goethe and the Sciences: A Reappraisal.* 1987 ISBN 90-277-2265-X; Pb 90-277-2400-8
98. J.C. Pitt and M. Pera (eds.): *Rational Changes in Science.* Essays on Scientific Reasoning. Translated from Italian. 1987 ISBN 90-277-2417-2
99. O. Costa de Beauregard: *Time, the Physical Magnitude.* 1987 ISBN 90-277-2444-X
100. A. Shimony and D. Nails (eds.): *Naturalistic Epistemology.* A Symposium of Two Decades. 1987 ISBN 90-277-2337-0
101. N. Rotenstreich: *Time and Meaning in History.* 1987 ISBN 90-277-2467-9
102. D.B. Zilberman: *The Birth of Meaning in Hindu Thought.* Edited by R.S. Cohen. 1988 ISBN 90-277-2497-0

103. T.F. Glick (ed.): *The Comparative Reception of Relativity.* 1987
ISBN 90-277-2498-9
104. Z. Harris, M. Gottfried, T. Ryckman, P. Mattick Jr, A. Daladier, T.N. Harris and S. Harris: *The Form of Information in Science.* Analysis of an Immunology Sublanguage. With a Preface by Hilary Putnam. 1989 ISBN 90-277-2516-0
105. F. Burwick (ed.): *Approaches to Organic Form.* Permutations in Science and Culture. 1987 ISBN 90-277-2541-1
106. M. Almási: *The Philosophy of Appearances.* Translated from Hungarian. 1989
ISBN 90-277-2150-5
107. S. Hook, W.L. O'Neill and R. O'Toole (eds.): *Philosophy, History and Social Action.* Essays in Honor of Lewis Feuer. With an Autobiographical Essay by L. Feuer. 1988 ISBN 90-277-2644-2
108. I. Hronszky, M. Fehér and B. Dajka: *Scientific Knowledge Socialized.* Selected Proceedings of the 5th Joint International Conference on the History and Philosophy of Science organized by the IUHPS (Veszprém, Hungary, 1984). 1988 ISBN 90-277-2284-6
109. P. Tillers and E.D. Green (eds.): *Probability and Inference in the Law of Evidence.* The Uses and Limits of Bayesianism. 1988 ISBN 90-277-2689-2
110. E. Ullmann-Margalit (ed.): *Science in Reflection.* The Israel Colloquium: Studies in History, Philosophy, and Sociology of Science, Vol. 3. 1988
ISBN 90-277-2712-0; Pb 90-277-2713-9
111. K. Gavroglu, Y. Goudaroulis and P. Nicolacopoulos (eds.): *Imre Lakatos and Theories of Scientific Change.* 1989 ISBN 90-277-2766-X
112. B. Glassner and J.D. Moreno (eds.): *The Qualitative-Quantitative Distinction in the Social Sciences.* 1989 ISBN 90-277-2829-1
113. K. Arens: *Structures of Knowing.* Psychologies of the 19th Century. 1989
ISBN 0-7923-0009-2
114. A. Janik: *Style, Politics and the Future of Philosophy.* 1989
ISBN 0-7923-0056-4
115. F. Amrine (ed.): *Literature and Science as Modes of Expression.* With an Introduction by S. Weininger. 1989 ISBN 0-7923-0133-1
116. J.R. Brown and J. Mittelstrass (eds.): *An Intimate Relation.* Studies in the History and Philosophy of Science. Presented to Robert E. Butts on His 60th Birthday. 1989 ISBN 0-7923-0169-2
117. F. D'Agostino and I.C. Jarvie (eds.): *Freedom and Rationality.* Essays in Honor of John Watkins. 1989 ISBN 0-7923-0264-8
118. D. Zolo: *Reflexive Epistemology.* The Philosophical Legacy of Otto Neurath. 1989 ISBN 0-7923-0320-2
119. M. Kearn, B.S. Philips and R.S. Cohen (eds.): *Georg Simmel and Contemporary Sociology.* 1989 ISBN 0-7923-0407-1
120. T.H. Levere and W.R. Shea (eds.): *Nature, Experiment and the Science.* Essays on Galileo and the Nature of Science. In Honour of Stillman Drake. 1989
ISBN 0-7923-0420-9

# Boston Studies in the Philosophy of Science

121. P. Nicolacopoulos (ed.): *Greek Studies in the Philosophy and History of Science.* 1990 ISBN 0-7923-0717-8
122. R. Cooke and D. Costantini (eds.): *Statistics in Science.* The Foundations of Statistical Methods in Biology, Physics and Economics. 1990 ISBN 0-7923-0797-6
123. P. Duhem: *The Origins of Statics.* Translated from French by G.F. Leneaux, V.N. Vagliente and G.H. Wagner. With an Introduction by S.L. Jaki. 1991 ISBN 0-7923-0898-0
124. H. Kamerlingh Onnes: *Through Measurement to Knowledge.* The Selected Papers, 1853-1926. Edited and with an Introduction by K. Gavroglu and Y. Goudaroulis. 1991 ISBN 0-7923-0825-5
125. M. Čapek: *The New Aspects of Time: Its Continuity and Novelties.* Selected Papers in the Philosophy of Science. 1991 ISBN 0-7923-0911-1
126. S. Unguru (ed.): *Physics, Cosmology and Astronomy, 1300-1700.* Tension and Accommodation. 1991 ISBN 0-7923-1022-5
127. Z. Bechler: *Newton's Physics on the Conceptual Structure of the Scientific Revolution.* 1991 ISBN 0-7923-1054-3
128. É. Meyerson: *Explanation in the Sciences.* Translated from French by M-A. Siple and D.A. Siple. 1991 ISBN 0-7923-1129-9
129. A.I. Tauber (ed.): *Organism and the Origins of Self.* 1991 ISBN 0-7923-1185-X
130. F.J. Varela and J-P. Dupuy (eds.): *Understanding Origins.* Contemporary Views on the Origin of Life, Mind and Society. 1992 ISBN 0-7923-1251-1
131. G.L. Pandit: *Methodological Variance.* Essays in Epistemological Ontology and the Methodology of Science. 1991 ISBN 0-7923-1263-5
132. G. Munévar (ed.): *Beyond Reason.* Essays on the Philosophy of Paul Feyerabend. 1991 ISBN 0-7923-1272-4
133. T.E. Uebel (ed.): *Rediscovering the Forgotten Vienna Circle.* Austrian Studies on Otto Neurath and the Vienna Circle. Partly translated from German. 1991 ISBN 0-7923-1276-7
134. W.R. Woodward and R.S. Cohen (eds.): *World Views and Scientific Discipline Formation.* Science Studies in the [former] German Democratic Republic. Partly translated from German by W.R. Woodward. 1991 ISBN 0-7923-1286-4
135. P. Zambelli: *The Speculum Astronomiae and Its Enigma.* Astrology, Theology and Science in Albertus Magnus and His Contemporaries. 1992 ISBN 0-7923-1380-1
136. P. Petitjean, C. Jami and A.M. Moulin (eds.): *Science and Empires.* Historical Studies about Scientific Development and European Expansion. ISBN 0-7923-1518-9
137. W.A. Wallace: *Galileo's Logic of Discovery and Proof.* The Background, Content, and Use of His Appropriated Treatises on Aristotle's *Posterior Analytics.* 1992 ISBN 0-7923-1577-4

## Boston Studies in the Philosophy of Science

138. W.A. Wallace: *Galileo's Logical Treatises.* A Translation, with Notes and Commentary, of His Appropriated Latin Questions on Aristotle's *Posterior Analytics.* 1992 ISBN 0-7923-1578-2
Set (137 + 138) ISBN 0-7923-1579-0
139. M.J. Nye, J.L. Richards and R.H. Stuewer (eds.): *The Invention of Physical Science.* Intersections of Mathematics, Theology and Natural Philosophy since the Seventeenth Century. Essays in Honor of Erwin N. Hiebert. 1992
ISBN 0-7923-1753-X
140. G. Corsi, M.L. dalla Chiara and G.C. Ghirardi (eds.): *Bridging the Gap: Philosophy, Mathematics and Physics.* Lectures on the Foundations of Science. 1992 ISBN 0-7923-1761-0
141. C.-H. Lin and D. Fu (eds.): *Philosophy and Conceptual History of Science in Taiwan.* 1992 ISBN 0-7923-1766-1
142. S. Sarkar (ed.): *The Founders of Evolutionary Genetics.* A Centenary Reappraisal. 1992 ISBN 0-7923-1777-7
143. J. Blackmore (ed.): *Ernst Mach – A Deeper Look.* Documents and New Perspectives. 1992 ISBN 0-7923-1853-6
144. P. Kroes and M. Bakker (eds.): *Technological Development and Science in the Industrial Age.* New Perspectives on the Science–Technology Relationship. 1992 ISBN 0-7923-1898-6
145. S. Amsterdamski: *Between History and Method.* Disputes about the Rationality of Science. 1992 ISBN 0-7923-1941-9
146. E. Ullmann-Margalit (ed.): *The Scientific Enterprise.* The Bar-Hillel Colloquium: Studies in History, Philosophy, and Sociology of Science, Volume 4. 1992 ISBN 0-7923-1992-3
147. L. Embree (ed.): *Metaarchaeology.* Reflections by Archaeologists and Philosophers. 1992 ISBN 0-7923-2023-9
148. S. French and H. Kamminga (eds.): *Correspondence, Invariance and Heuristics.* Essays in Honour of Heinz Post. 1993 ISBN 0-7923-2085-9
149. M. Bunzl: *The Context of Explanation.* 1993 ISBN 0-7923-2153-7
150. I.B. Cohen (ed.): *The Natural Sciences and the Social Sciences.* Some Critical and Historical Perspectives. 1993 ISBN 0-7923-2223-1
151. K. Gavroglu, Y. Christianidis and E. Nicolaidis (eds.): *Trends in the Historiography of Science.* 1993 ISBN 0-7923-2255-X

*Also of interest:*
R.S. Cohen and M.W. Wartofsky (eds.): *A Portrait of Twenty-Five Years Boston Colloquia for the Philosophy of Science, 1960-1985.* 1985 ISBN Pb 90-277-1971-3

*Previous volumes are still available.*

KLUWER ACADEMIC PUBLISHERS – DORDRECHT / BOSTON / LONDON