Studies in History and Philosophy of Science

Stud. Hist. Phil. Sci. 33 (2002) 487–509

www.elsevier.com/locate/shpsa

Reconstructing Lakatos: a reassessment of Lakatos' epistemological project in the light of the Lakatos Archive

Matteo Motterlini [a,b]

[a] *Department of Economics and Laboratory of Cognitive Science, University of Trento, Italy*
[b] *Centre for Philosophy of Natural and Social Science, The London School of Economics, UK*

Received 23 October 2000; received in revised form 20 February 2002

Abstract

Based on the material in the Lakatos Archive, this paper reconstructs, and then re-assesses, Lakatos' epistemological project by placing it in the context of the debate on the role of reason in the history of science, and of the justification of rationality as a normative notion. It is claimed that Lakatos' most fruitful ideas come from a peculiar philosophical combination of Hegelian historicism and Popperian fallibilism. The original tension, however, cannot be ultimately resolved. As a consequence, the problems that Lakatos has to deal with in his attempt to justify a set of genuinely epistemological canons of scientific rationality that are not reducible to psychology or sociology of knowledge stand as a warning for any normative philosophy of science that takes history at its face value.

Keywords: Lakatos; History; Method; Rationality

> Devil chooses strange shapes to confound the innocent. For example, he may choose the shape of a rationalist from London with a Hungarian accent.
>
> Paul Feyerabend

E-mail address: mmotterl@gelso.unitn.it (M. Motterlini).

0039-3681/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0039-3681(02)00024-9

1. A philosophical bastard

Throughout his daily and life-long correspondence with his friend Lakatos, Paul Feyerabend often claimed that his fellow was a philosophical bastard: a 'Pop-Hegelian' born from a Popperian father and a Hegelian mother ('Lakatos-Feyerabend Correspondence', in Motterlini, 1999, pp. 184, 194). In fact, some of Lakatos' most fruitful contributions—such as his dialectical conception of mathematical heuristic, the idea of rational reconstruction and, more generally, his emphasis on the role of criticism in the progress of knowledge—come directly from a peculiar philosophical conflation of Hegelian and Popperian ideas. I shall argue, however, that the original tension between conflicting parents cannot be ultimately resolved. Lakatos' project remains somehow trapped between Hegelian historicism and Popperian fallibilism. The reconstruction and re-assessment of Lakatos' syncretic project will be largely based on the material in the Archive of Professor Imre Lakatos (British Library at the London School of Economics for Political and Economic Science, henceforth the Archive).[1] I shall focus on those items in the Archive which illustrate the central dilemma in Lakatos' programme of providing genuinely epistemological canons of scientific rationality that are not historically relative, and that are not reducible to psychology and sociology of knowledge. I shall therefore deal in turn with (i) Lakatos' attempt to capture the rationality of the development of mathematics through the method of 'Proofs and Refutations', and the related notions of dialectics and heuristic; (ii) Lakatos' goal of grasping the unfolding rationality within the history of science, and the related revision of Popper's static fasificationist account; (iii) Lakatos' ever-shifting concept of methodology from his early writings to his latest declaration of intentions of a 'Changing Logic of Scientific Discovery'. The paper concludes by suggesting that the problems at stake do not just affect Lakatos' ingenious proposal, but also the very attempt to provide a normative theory of rationality that takes history at its face value.

2. Dialectics, Marxism and rationality: from Budapest to Cambridge

Lakatos' career as a professor of philosophy accounts for less than half of his adult life. In the 1940s and 1950s Lakatos was deeply involved in the political and intellectual debates of his native country.[2] His earliest published writings date back

[1] Lakatos' Archive is arranged in the following thirteen sections: 1. Papers published in Hungary. 2. Early Notes on Mathematics and the Philosophy of Mathematics. 3. Essays in the Logic of Mathematical Discovery. 4. Other Papers in the Philosophy of Mathematics. 5. Early Papers in the Philosophy of Science. 6. Middle-Period Papers in the Philosophy of Science. 7. Later Papers in the Philosophy of Science. 8. Papers on General Philosophy. 9. Lectures. 10. Notes on Feyerabend, Kuhn and Popper and Miscellaneous Notes. 11. Miscellaneous. 12. Selected Correspondence (includes exchanges with George Pólya, Victor Kraft, Rudolf Carnap, Richard Popkin, Paul Bernays, Alan Musgrave, Adolf Grünbaum, Thomas Kuhn and Karl Popper). 13. General correspondence.

[2] Lakatos had worked at the Ministry of Education (Culture and Religion) between 1945 and 1948 dealing with education reform. See Long (1997).

to the post-war period, and possess *in nuce* some of the characteristic traits of his mature philosophical works. Most of these writings are devoted to socio-political and pedagogic issues, and are heavily loaded with Hegelian-Marxist language and theory. In '*Le Citoyen* and the working class' (Archive 1.1), Lakatos uses an interesting device that would characterise his later work in the philosophy of science: the abstract principles of the philosophy of science are contrasted with the real scientific practice of the working scientist, just as the abstract figure of the *Citoyen* is contrasted with the reality of the working class. As a Marxist, Lakatos argues that the real principles of scientific and political development unfold at the material level, rather than in the realm of ideas.

In the years to follow, scientific method was to become more important in Lakatos' ideas, and Marxism less and less orthodox. After his purge from the Communist Party, he participated in the anti-Stalinist movement, and in particular co-authored a plea for the 'freedom of Hungarian scientific life from its Stalinist shackles.'[3] Critical attitude, science and absence of censorship must guide the party, instead of the other way around. Indeed, Lakatos came to believe that the positive aspects of Marxism coincided with the principles of scientific method.[4] Yet he still found a positive role for dialectics as an instrument of criticism. Dialectics—in the genuine Hegelian sense—was according to Lakatos to be contrasted with 'irrationalistic mystification', decadentism, vulgar Marxism and pedestrian mechanistic materialism. Curiously enough, Lakatos' dialectics looks somehow similar to the sort of fallibilism that he was to endorse in his later Popperian period: it is more an instrument of criticism than the standard attempt to vindicate the *status quo* by explaining the necessary (rational) realisation of present society in an age of totalitarian systems. Here, for instance, are Lakatos' views on the education of young scholars:

> The history of science indicates that we ought to teach the future scholar to be modest, to be humble in his scientific claims, to be averse to all kinds of fanaticism. He ought to learn that what he does not understand, or disapproves of, still has a right to exist, and that no scientific theory, no theorem can conclude anything finally, in the history of science. . . . New, hitherto unfamiliar chapters ought to be included in pedagogical textbooks, such as 'Methods for stimulating curiosity and developing it into interest', 'How to teach to think scientifically', 'How to teach people respect for facts' and—God forbid!—'How to teach people to doubt'. ('On rearing scholars', Archive, 1. 9; now in Motterlini, 1999, Appendix A)

The peculiar mix that makes up Lakatos' philosophy is already present. Such a mix, however, would take different forms in different periods of his intellectual career, depending on whether the Hegelian-Marxist component was to be corrected by a whiff of fallibilism, or Popperian fallibilism was to be ameliorated by means

[3] Declaration of the National Committee of the Hungarian Academy of Science, Archive 1.10.

[4] In the same period, during a discussion with friends, he is reported to have lost his temper when the question turned to defending Marxism: 'You are talking about scientific method, why do you keep calling it Marxism?' (Archive 1).

of an injection of Hegelian historicism. The first product of the Pop-Hegelian philosopher was the PhD thesis Lakatos completed in Cambridge, where he took refuge after the Hungarian Uprising in late 1956. In the acknowledgements, Lakatos indicates his polemical target in the conception of mathematics as static and authoritarian. On the contrary, in his view, 'mathematics is dialectics' and 'cannot exist without criticism'. Lakatos recognises 'three major—and apparently quite incompatible—"ideological" sources' of his work: Pólya's mathematical heuristics, Hegel's dialectic, and Popper's fallibilism (PhD thesis, Archive, 3.4). The emphasis on the dynamic unfolding of mathematical knowledge as a product largely independent of the producer's psychology is a clear reference to Hegel; whereas when Lakatos refers to Popper he takes a position against any account of mathematics as certain and definitive knowledge. Combined with this is the reference to Pólya: mathematics is a problem-solving activity. The next section is devoted to illustrating in details such a peculiar philosophical combination.

3. Hegel, Pólya and Popper as 'ideological sources'

Regarding Lakatos' intellectual education during the Hungarian period, Arpad Szabó later claimed that Lakatos attended his seminars on ancient Greek mathematics, but apparently was more interested in Marxism than anything else. We know that he attended Lukács' lectures on Hegel's *Phenomenology of Spirit* at Budapest University, but we do not know exactly what sort of impact, if any, these had on his philosophical upbringing. Lakatosian scholars have recently tried to highlight the Hegelian component of his later work, both in the philosophy of the natural sciences (Hacking, 1979) and in the philosophy of mathematics (Larvor, 1999; Kadvany, 2001). We shall briefly discuss Hacking's views later—for the time being let us just note that the influence cannot have sprung directly from Hegel's philosophy of mathematics as outlined in the *Phenomenology of Spirit* (1977). Hegel uses mathematics precisely as a (negative) example of non-dialectical or 'static' rationality. Philosophy, on the contrary, is characterised by dynamic propositions in constant development, according to the dialectical scheme of thesis-antithesis-synthesis. Lakatos' mathematics, therefore, looks much more similar to Hegel's philosophy than to Hegel's mathematics.

Lakatos' contribution, however, is much more than a mere application of the dialectical approach to a new domain, for he adds a crucial component: empirical heuristics. The standard view of mathematics, from Euclid onwards, but especially after the logicist revolution of the early twentieth century, is dominated by proof. It is formal proof that provides a mathematical result with reliability and rigour. Lakatos separates reliability from certainty. The key move is to consider mathematical knowledge 'in the making', i.e. in pre-axiomatic form. The process of mathematical discovery, not unlike discovery in empirical science, is a mixture of conjectures, 'experiments' and refutation. Although the language we have used (following Lakatos) is explicitly Popperian, the origin of these ideas is not Popper himself. It is well estab-

lished that Lakatos took inspiration from his fellow Hungarian Pólya, whose book *How to solve it* he translated from English into Hungarian.[5]

According to Pólya (1954) mathematics as a problem-solving activity—that is, mathematics in the making—is in most respects an inductive science, making heavy use of examples, generalisations, counterexamples and refutations. Proof enters the stage only much later. The engine of discovery is fuelled precisely by counterexamples, i.e. thought-experiments aimed at invalidating previous generalisations.[6] Counterexamples in mathematics raise pretty much the same problems as anomalous experiments in the natural sciences. Whenever a set of anomalous data are observed, the scientist must investigate which part of the system under test is responsible for the inconsistency. Is the original theory to be rejected? Or is part of the apparatus at fault? Or maybe some peripheral assumption has been violated? Similarly, it is not easy to deal with a mathematical anomaly. The arrow of refutation must be directed towards some part of the theoretical system. Most often, some auxiliary lemma is criticised in order to save from refutation the 'hard core' of our research. Feedback from counter-examples is particularly crucial in mathematics because, in calling for a further analysis of the primitive conjecture and of the proof, it suggests where the amendments have to be made and which (no longer hidden) lemma has to be replaced.

It is well known that according to Popper (and Reichenbach) there is no logic of discovery—in the sense of a rigorous set of rules to be followed in the construction of scientific theories. Rationality is limited to the realm of testing, discovery is to be explicated by psychologists. Pólya, similarly, spoke of heuristic as a series of 'mental operations typically useful for the solution of problems' (Pólya, 1973, p. 2).[7] Lakatos instead envisages room for a non-psychologistic heuristic, providing instructions for the construction and improvement of mathematical conjectures. In this sense, there is a close connection between discovery and testing, for proofs allow the laying out of the target for refutation (the axioms and lemmas), and proof-analysis allows the identification of the weak ones, and their replacement according to a set of identifiable strategies.

The idea that mathematical discovery follows some patterns that can be rigorously

[5] See for example, Pólya (1973): 'Studying the methods of solving problems, we perceive another face of mathematics. Yes, mathematics has two faces; it is the rigorous science of Euclid, but it is also something else. Mathematics presented in the Euclidean way appears as a systematic, deductive science, but mathematics in the making appears an experimental, inductive science' (p. vi).

[6] It is worth mentioning the influence of Lakatos' former teacher at Debrecen in this respect. In particular, Lakatos refers to Szabó (1958) when he considers 'thought experiments' as the pattern of mathematical proofs in the pre-Euclidean Greek mathematics. From Szabó (1960) Lakatos takes the idea that, in Euclid's time, 'postulates' and 'axioms' meant propositions in the critical (dialectic) dialogue put forward to be tested for consequences without being admitted as true by the discussants.

[7] Another major difference between Pólya's and Lakatos' work lies in the former's lack of interest in epistemological issues. In a letter, Pólya claims that he 'would hardly be able to say anything on "epistemology" that would deserve the attention of the public' (Pólya to Lakatos, 15 December 1965, Archive 12.9; for the full letter see Motterlini, 2002, p. 30). For an extensive comparison of Pólya's and Lakatos' heuristic, see Feferman (1978).

analysed comes from Pólya; the idea that it is a matter of conjectures, (proofs) and refutations comes from Popper. However, it is mainly the Hegelian ingredient that provided Lakatos with a different perspective from both Pólya and Popper. The Hegelian ingredient I am referring to is the identification of pure growth with the *essence* of mathematics—an essence, crucially, which is independent of the activity of the mathematicians, which 'alienates itself'. Mathematics 'becomes a living, growing organism, that acquires certain autonomy from the activity which has produced it; develops its own autonomous laws, its own dialectics' (Lakatos, 1976, pp. 145–6).[8]

Since this aspect of Lakatos' philosophy of mathematics has already been discussed elsewhere (Larvor, 1999; Kadvany, 2001; Motterlini, 2002), I shall just mention one important implication of the Hegelian move here. In a footnote to *Proofs and Refutations*, Lakatos notices that despite Marx's failure to notice it, 'there is nothing inexorable in the realisation of this process [of personification of mathematics in the concrete mathematical activity of the mathematicians]. On the contrary, human activity can always suppress or distort the autonomy of the alienated process and can give rise to new ones' (Lakatos, 1976, p. 148 n. 1). The appearance of inexorability comes from the temptation of justifying any historical event with the Hegelian 'cunning of reason': everything that takes place must have taken place and thus must also be 'right'. But Lakatos here corrects Hegel with a dose of Popperian fallibilism: the process of scientific growth is never-ending, and our syntheses of today will be our theses tomorrow. The same with heuristics: we cannot be sure that everything we observe is a direct manifestation of the pure (objective) logic of mathematical discovery. There is nothing inexorable in the instantiation of pure rationality in human activity.

To sum up: Lakatos' philosophy of mathematics combines pieces of Hegel's, Pólya's and Popper's ideas in a delicate but exciting equilibrium, where the shortcomings of one approach are mitigated or compensated for by the strengths of another one. Popper's fallibilism mitigates Hegel's authoritarianism. The irrationality of discovery is replaced by the idea of a dynamically unfolding rationality and well structured rules of discovery. Pólya's subjectivism is replaced by Hegel's 'alienated', objectified process. It is important to stress that this is not merely a philosophical patchwork: as we shall see in the rest of this paper, Lakatos was using powerful instruments to tackle an all-important problem—that of the rationality of science and theory change—which still defies solution nowadays. In the next section we shall see how the method first outlined in *Proofs and Refutations* has deep philosophical implications that would be carried on to Lakatos' work in the philosophy of natural science.

[8] In 1966 'Proofs and Refutations' was translated into Russian in 70,000 copies. Lakatos himself was very proud of the success it met. Some handwritten notes in the Archive are useful in explaining this success: the Hegelian dialectical triad lurks behind the historical reconstruction of Euler's theorem: 'Thesis: Primitive conjecture, Antithesis: Counterexample, Synthesis: Theorem and proof generated concept (+ lemma incorporation)' (Archive 3). Russian readers of Lakatos' *Dokatatelstva i Oprovershenia* encountered no problem in detecting a certain familiarity linking mathematical heuristic and dialectic.

4. The philosophical implication of the method of *Proofs and Refutations*[9]

Lakatos' work in mathematics was mainly carried out in Cambridge in the late 1950s and early 1960s, whereas his philosophy of science was developed at LSE initially under Karl Popper. We have seen that Lakatos' philosophy of mathematics is a product of many influences, one of which was certainly Popper's falsificationism. So, on the one hand Popper's theory of scientific method influenced Lakatos' philosophy of mathematics. On the other, Lakatos' ideas in philosophy of science, which improve on Popper's approach, are clearly based on his previous research into the growth of mathematical knowledge. Lakatos' programme therefore develops along a single route, but in two ways: from the philosophy of science to the philosophy of mathematics, and then back from the philosophy of mathematics to the philosophy of science.[10] As we shall see, however, the goal of these enterprises is quite different in one respect. Lakatos' philosophy of mathematics aims to challenge dogmatism (i.e. Euclidianism and formalism) rather than to defend fallibilism from the attack of sceptical irrationalism. Within an Hegelian framework, the rationality of the development of mathematics (and science) being the very premise does not need to be argued for. Lakatos' concern here was to show that knowledge in mathematics is not static but dynamic, that it cannot exist without criticism (i.e. dialectic). On the other hand, Lakatos' philosophy of science aims to grasp the unfolding rationality within the history of science and to defend it from any attack, these being anarchic (Feyerabend), elitist (Polanyi, Kuhn) or even Hegelian in disguise (Toulmin).

The Archive allows us to reconstruct Lakatos' own pattern of discovery in detail. I refer in particular to a couple of files catalogued by Lakatos himself under the significant titles 'Some Philosophical Implications of the Method of Proofs and Refutations' and 'Research Programmes as a Continuation of the Method of Proofs and Refutations'. Here is how Lakatos extended the method of *Proofs and Refutations* to the empirical sciences:

> In science the same heuristic pattern [one frequently comes across in mathematics] has an important role: 1. A problem is proposed. 2. A solution is put forward in the form of a naive conjecture. 3. The naive conjecture is explained and refuted. 4. The explanation is analysed into lemmas, and lemmas are incorporated in the naive conjecture. The result is an irrefutable theorem. 5. The global counterexamples are tentatively traced to lemmas. 6. The lemmas—in particular those pointed out in 5—are refuted (local counterexample) and replaced by more general ones, and the theorem is correspondingly generalised. Refutations lead to rival theories. 7. Total local refutations lead to rival theories. 8. After saturation point: rejection. (Handwritten notes, 'The logic of Explanations and Refutations', Archive 5.8)

[9] This section is largely based on the Lakatos Archive and it is mainly exegetical. A different version has already been published in Motterlini (2002), Sections 5 and 6.

[10] For the analogies between Lakatos' logic of mathematical discovery and his MSRP, see Zheng (1990).

These notes should be regarded in the context of Lakatos' struggle with the 'Duhem problem' and the related difficulties this implies for a falsificationist account of science.[11] As is well known, Duhem had claimed that physics, far from being a 'machine which lets itself be taken apart' is rather 'an organism in which one part cannot be made to function except when the parts that are most remote from it are called into play, some more so than others, but all to some degree' (Duhem, 1954, pp. 187–188). In this regard, Lakatos noted that

> In physics, if you have a global counter example you do not know what you have refuted (Poincaré, Duhem, Quine). Deductive model—sorting out possible lemmas—devise pinpointing tests. (Handwritten notes, Archive 5)

There follow a few notes in which Lakatos shows his intention of articulating the method of *Proofs and Refutations* in order to give an explanation of instances taken from the history of science such as, for example, Bohr's discussion of the structure of the atom—see Lakatos (1970). Among the papers present in the section there are also other draft notes on the case study of the alleged 'deduction' of Newton's theory from Kepler's laws. Lakatos meant to present this in dialogue form, in analogy with his doctoral dissertation with the title 'From Facts to Empirical Law'.[12] Lakatos moves on from his consideration of the role of criticism in mathematics to the methodology of scientific research programmes (MSRP) as a solution of the 'Duhem problem'. It is peculiar that Duhem had always set aside mathematics from the context of his challenge—and yet the passage in *La théorie physique* in which this occurs 'is the only one in Duhem's great book that has aged since it was written' (Truesdell, 1984, p. 490). It remains valid for mathematics too that one should examine a series of propositions rather than a single one. Since 1961, Lakatos had claimed that when handling a counter-example to the initial conjecture ('global counter-example') or to any of the lemmas ('local counter-examples') one has to choose between various lines of behaviour each characterised by promises and risks peculiar to itself. The main case study of *Proofs and Refutations* illustrates the different possible strategies available in dealing with polyhedra that are exceptions to Euler's conjecture. A possible choice involves 'eliminating monstrosities' by refining the concept of polyhedron; another consists in 'surrendering to the counterexamples', thereby declaring the conjecture to be false; yet another in making certain 'hidden lemmas' from background knowledge explicit and inserting them into the theorem enunciation, and so on. The first kind of tactic reduces the content of Euler's conjecture and turns it into

[11] The Archive (5.4) includes three different contributions to Popper's philosophy: 'The Popperian historiography', 'The so-called deductive model of explanation' and 'The metaphysical presuppositions of Popperian methodology'. Lakatos was probably aiming to write a single article, which however was never put together. This material eventually led to Lakatos (1970, 1974).

[12] The same problem has been tackled from a Lakatosian point of view by Worrall (mimeo) and Zahar (1983).

a 'miserable convention'.[13] The second move is that of 'naive' falsificationism in mathematics. The point is, of course, that one must recognise the element of decision inherent both in restricting the meaning of basic terms (such as polyhedron, vertex, edge, face etc.) and in extending it. Growth of mathematical knowledge results from those same 'adjustments' that allow what in the empirical sciences constitutes an increase in (corroborated) content—i.e. those 'stratagems' that allow one to invent/discover new problems, solve some of them and give a more elegant formulation of a simple conjunction of restrictive clauses. Only this third way corresponds to 'sophisticated falsificationism' in the empirical sciences.

The following is Lakatos' heuristic model of explanation as an attempt to meet Duhem's challenge:

> [O]f course, I take the 'Duhem-Quine thesis' for granted in the sense that any refutation undermines a large bulk of our knowledge and not a uniquely specified part of it. But Duhem and Quine give no sufficient indication of how to make a reasoned guess *which* part of our knowledge is responsible for the inconsistency; *indeed, they insinuate that no such reasoned guess can be proposed*. This is the variant of the Duhem-Quine thesis that Popperians reject. Also, the Duhem-Quine thesis does not lay sufficient emphasis on the 'implicit' or 'hidden' character of the bulk of the background knowledge. . . . The concrete, positive version of a hidden lemma is always invented/discovered under critical pressure. Imagination and criticism unfolds—slowly and with frequent hitches—ever more of the deductive structure. One may say that *the target of the arrow of refutation is shaped while the arrow is already in the air*. Criticism does not *assume* a fully articulated deductive structure: it creates it. The true deductive model of explanation is an ever-changing one; one in which propositions keep being added and deleted. *One may not explain what one has set out to explain; one may not refute what one has set out to refute*. ('On the so-called "deductive" model of explanation', Archive 5.4; emphasis added)

Since theories are not fully fledged deductive systems, counter examples are not entirely negative in so far as they help to reveal the hidden assumptions and unfold thereby a new deductive structure. In the very spirit of *Proofs and Refutations*, this long passage is nonetheless only a declaration of intention. In fact, in promoting the MSRP in analogy to the method of *Proofs and Refutations*, Lakatos found himself giving too much credit to the theoretical autonomy of research programmes, thereby undermining the role of counter-examples and also the intrinsic unity of the context of discovery and the context of justification. He thus lost sight of the main question—

[13] One can read into it Poincaré's conservative stand on relevant cases taken from the history of physics: see Poincaré (1982).

that of how to improve one's conjecture—and did not fully exploit the possibilities he had outlined in his programme.[14]

From the method of *Proofs and Refutations* to the MSRP an important shift in the concept of heuristic occurs.[15] Lakatos the philosopher of mathematics claims that he uses the 'word "methodology" in a sense akin to Pólya's "heuristic" and Popper's (1959) "logic of discovery"' (Lakatos, 1976, p. 3). Recall that for Pólya heuristic is a set of strategies for solving mathematical problems in order to learn, to teach and to reconstruct mathematics. Discovery and invention are mainly considered in their psychological aspects. According to Popper, the logic of discovery (or, better, the 'Logik der Forschung'), in the sense of a theory of scientific method, is both descriptive and normative. Popper's demarcationist project not only evaluates scientific products but also offers standards of intellectual honesty that scientists have to meet in order to fulfil the aim of science. According to the author of *Proofs and Refutations*, the role of 'heuristic-methodology' is strictly related to its object of inquiry. As we have seen, mathematics is a product of human activity that 'alienates itself': it is autonomous and objective. So the purpose of heuristic-methodology is to grasp the logic of the development of mathematics, the dialectical pattern of growth, the rationality of mathematics in the making. From this point of view, the heuristic-methodology looks *backward* to identify the rules that made such a growth possible in the past, and at the same time it looks *forward* to advise on how to obtain progress in the future. Heuristic-methodology, although fallible, is both evaluative and normative. Generally speaking, from the Hegelian–Marxist view the aim of philosophy is not a contemplation of eternal truths, but rather an effort to interpret the present in the light of the past with a view to shaping the present for a better (utopian) future (see Larvor, 1998, p. 1). Thus, for the Hegelian philosopher of mathematics, methodology, heuristic and logic of discovery are synonymous. But this is not the case for the author of the MSRP. For the LSE Professor, the hope that methodology

> would provide scientists with a mechanical book of rules for solving problems has been given up: modern methodologies or logics of discovery consist merely of a set of . . . rules for the appraisal of ready articulated theories. . . . The term 'normative' no longer means rules for arriving at solutions, but merely directions for the appraisal of solutions already there . . .

It follows that 'methodology is separated from heuristic rather as value statements are from "ought" statements' (Lakatos, 1971a, p. 103 n. 1). Thus, methodology no longer concerns the set of rules and strategies to be adopted in the context of discovery. Only heuristic does. But it does so in a different way than before. Heuristic principles (as distinct from methodological ones) are not 'objective' and 'auton-

[14] A similar point has also been made by Fine (1978), who has stressed the difference between playing the game 'save the proof' (by stretching the concept) and the game of 'advancing our science by developing new theories'. See also Forrai (1993), pp. 170–174.

[15] Forrai (1993) has already raised this point and has also suggested the importance of the Hegelian inheritance in this context.

omous'. They instruct scientists how to proceed within a particular research programme. They are therefore subjected to change alongside changes in science. (For example, we have an Aristotelian Ptolemaic heuristic, a Copernican heuristic, a Newtonian heuristic, an Einsteinian heuristic, and so on.) The last nail in the coffin of the early Lakatos is that methodology (divorced from normative heuristic) aims to advise scientists neither about how to arrive at good theories nor even about which of two rival programmes they should work on: 'Whatever scientists have done', Lakatos claims in his 'Replies to Critics', 'I can judge: I can say whether they have made progress or not. But I cannot advise them—and I do not wish to advise them—about exactly what to worry and in which direction they should seek progress' (Lakatos, 1971b, p. 178).

Moreover, the early Lakatos seems to take for granted that 'After a saturation point: we reject the theory' (see clause 8 of the passage cited above), while elsewhere he states, on the contrary, that there is 'no such a thing as a "natural" saturation point' for a research programme.[16] His standards of appraisal do not in fact fix any time limit for the final assessment of the empirical progressiveness or degeneration of a programme. At the dawn of a new and ambitious scientific idea a certain methodological tolerance is called for, and this applies to research programmes whose heuristic has 'run out of steam'. There is nothing 'irrational' in supporters of a theory defending it with ingenious *ad hoc* stratagems or holding to it despite long periods without empirical success. 'Crucial' experiments are hence seen to be crucial only decades later, 'after long hindsight', and 'rationality works much more slowly than most people tend to think, and even then fallibly. Minerva's owl flies at dusk' (Lakatos, 1970, pp. 72, 87). Curiously enough, even in the later Lakatos, Hegel lives on. To put it in Hegelian jargon, 'absolute knowledge' in the form of complete 'self-consciousness' and 'self-possession of spirit' is only available at the end-point of the process:

> Philosophy always comes to the scene too late to give instruction as to what the world ought to be. As the *thought* of the world, it appears only when actuality is already there, cut and dried, after its process of formation has been completed . . . It is only with the fall of the dusk that the owl of Minerva spreads its wings. (Hegel, 1991, 'Preface', p. 23)

It may be the case that (the rationality of) science can be understood backwards, but in practice, of course, it has to be done forwards, as Kierkegaard had noted with reference to Hegel (simply substitute 'life' for 'science'). Having set aside Popper's 'instant rationality' in order to bestow on methodology the role of the owl of Minerva, Lakatos has been accused by Feyerabend of oscillating between a conservative use of the methodological standards of appraisal, which would eventually lead

[16] See Lakatos (1970): 'in my [*Proofs and Refutations*] I was more of a Hegelian, and I thought there was [a natural 'saturation point']; now I use the expression with ironical emphasis. There is no predictable or ascertainable limitation on human imagination in inventing new, content-increasing theories' (p. 72 n. 1).

to benefits for the *status quo*, and a revolutionary use which amounts simply to 'anything goes'.[17] Lakatos betrays his authoritarian nature by strengthening his methodological standards not on an argumentative level, but by shaping a historical and social situation that renders it difficult, *in practice*, to cultivate a degenerating programme. (For example, he claims that people who support a degenerating research programme should not enjoy as much freedom as might appear: 'they can do this mostly in private. Editors of scientific journals should refuse to publish their papers, which will contain either solemn reassertion of their positions, or absorption of counter-evidence by *ad hoc* linguistic adjustments. Research foundations, too, should refuse money (Lakatos, 1971a, p. 117). On the other hand, he also leaves space for ungoverned freedom since, taken by themselves, these standards are, *in principle*, incapable of bearing even the most outrageous pseudoscientific behaviour. (For example, Lakatos claims that 'there is freedom . . . over which programme to work on . . .. *Appraisal* does not imply *advice*'—Lakatos, 1973, p. 110.) Whether or not Lakatos has really gone a long way towards epistemological anarchism[18] is a problem we can better solve in the more general framework of Lakatos' late struggle against the 'modern intellectuals' betrayal of reason'.

5. Change, progress and truth: the Hegelian cunning of reason *versus* Popperian fallibilism

In the Summer of 1973 Lakatos was working on a review of Stephen Toulmin's *Human Understanding* (Toulmin, 1972). His intention was to place Toulmin's approach in the more general framework of the great traditions in the history of epistemology: scepticism, demarcationism and elitism. Lakatos assimilates the enquiry into 'human understanding' with the teaching of the 'second Wittgenstein' from whom Toulmin clearly descends, were it not for an important distinction. From the point of view of Lakatos, the Wittgenstein of the *Philosophical Investigations* (1958) is an intellectual defender of the *status quo*. The mission of new Wittgensteinian philosophers is to discourage every incursion from the outside and overthrow a 'linguistic game' or a 'form of life'. Toulmin, however, also supports the usefulness of change, but at the cost of appealing to the 'Hegelian Cunning of Reason', which justifies *change* in the name of *progress*. If progress is guaranteed by the cunning of reason, the description of change is the description of progress:

> [The elitists] believe that science can make real progress, but since they claim that there is no universal criterion of progress, they are bound to claim that any change in science means by the Hegelian Cunning of Reason, progress in science.

[17] Feyerabend's criticisms of Lakatos are freely taken from Feyerabend (1975), Ch. 17, and Feyerabend (1976).

[18] On the alleged existence of 'two Lakatos''—the supporter of a more or less close relation between appraisal and advice and the one who, denying that MSRP does not give any advice at all, has gone 'a long way towards epistemological anarchism'—see Musgrave (1976) and Motterlini (1995).

> Might is right—at least among genuine scientific communities; selective survival is the criterion of progress. (Lakatos, 1973/1976, p. 228)

Like a seaman in the famous pirates' ballad, caught 'between the devil and the deep blue sea' (that is, between his captain who held near-dictatorial powers, on the one side, and the dangerous boundless forces of nature on the other),[19] Toulmin recurrently finds himself 'between the Wittgensteinian devil and the Hegelian deep blue sea'—which is precisely the title of one of Lakatos' three increasingly detailed versions of the same paper on Toulmin (Archive 8.5). Lakatos had surely thought about this, since, in another version, Toulmin was caught 'between the Hegelian devil and the Wittgensteinian deep blue sea'. The uncertainty here is understandable. The outcomes of both of Toulmin's stands were diabolical to Lakatos, since they invariably led to a 'closed society'.[20]

Lakatos' scathing sarcasm, however, may easily backfire at this point, since he would *himself* remain trapped between the Hegelian devil and Popperian deep blue sea.[21] In particular, according to Hacking's influential interpretation, due to his Hegelian-Marxist education, Lakatos 'dispenses with the notion of truth as correspondence to the facts and *replaces truth by method*' (Hacking, 1979). Hence, Lakatos too justifies change in the name of progress.

Apart from the instigation of rhetoric, we need to address the problem of whether Lakatos succeeded in distinguishing his position as opposed to authoritarian elitism, scepticism and the Hegelian cunning of reason.

To start with, Lakatos maintains that any demarcation criterion is fallible. However, collapse into historical relativism can be avoided only should 'rational appraisal *precede* and not *follow* full-scale empirical history' (i.e. 'internal' normative history is primary and external descriptive history is secondary). Similarly, the regressive problem-shift from the problem of demarcation to the problem of knowledge expressed 'in the form of skills and activities' that only the experts (the elite) can judge (the latter leading to the Orwellian world where history is written by the

---

[19] On the seaman's dilemma, see Rediker (1987), p. 5.

[20] See Lakatos (1973/1976): 'Following the tradition of Wittgenstein, Polanyi and Kuhn . . . the picture described [by Toulmin] is that of a society without radical alternatives . . . a society whose membership depends on oaths of loyalty to specific doctrines ("commitment to collective ideas") and where only "professional forums" can judge the implications of these doctrines for specific cases. In this closed society critical reappraisal and modification are allowed only if done by "qualified judges". The layman is powerless, the elite self-perpetuating' (p. 241).

[21] See Motterlini (2002). Moreover, since in his late years Lakatos became more and more disappointed about Popper's achievements, *in this case too*, reversing the terms—i.e. 'Lakatos between the Popperian devil and the Hegelian deep blue sea—could work just as well. To Lakatos, Popper's lack of tolerance regarding the criticism directed towards his own falsificationism must have appeared as diabolical as the old Hegelian authoritarianism. An example is how Lakatos taught his students about Popper's philosophy while lecturing at LSE in 1973: 'Popper's three major contributions to philosophy were: 1) his falsifaibility criterion—I think this is a step back from Duhem; 2) his solution to the problem of induction—where I think he is a step back from Hume, and 3) his literary masterpiece "The Open Society by one of its enemies"' (Motterlini, 1999, p. 189). See also Lakatos–Feyerabend Correspondence, in particular the letter dated 26.6.1972, where Popper is mentioned as 'Al-Poppuni', 'the great tyrant of Reason'.

winners) can be avoided only by appealing to a 'universal criterion of progress'. Such a criterion is provided by the MSRP in the form of principles of theory-appraisal.[22] Moreover, as we appraise scientific theories, we should be able to appraise the standards of appraisals. The tool for this task is history. The meta-criterion is provided by the methodology of historical research programmes (MHRP). In particular, we accept a methodological proposal if it can be shown that it was effective in paradigmatic cases of growth of knowledge. In fact, even though there has been no general agreement concerning a methodological criterion, Lakatos maintains that there has been considerable agreement about whether a particular step in the game was scientific or crankish, or whether a particular gambit was played correctly or not. Thus Lakatos' meta-criterion appraises methodologies on their ability to provide historical reconstructions minimising the influence of 'external' factors and maximising 'internal' explanations. Last but not least, it is my view that Hacking's interpretation is misleading. In fact, the late Lakatos could not have regarded methodology as a substitute for truth because he explicitly demanded a connection between the game of science (i.e. method) and its aim (i.e. truth).[23] Popperian fallibilism rather than the post-Kantian demolition of the copy theory of truth is therefore the correct key to understand Lakatos's late view in this context.[24] Of course, for a fully coherent fallibilist, 'there is no ultimate proof that . . . we have been heading towards the Truth. We can only (non-rationally) believe, or rather hope, that we have been. Unless hope is a "solution", there is no solution to Hume's problem' (Lakatos, 1974, p. 213). As we shall see, having disentangled himself from the Hegelian cunning of reason Lakatos finds the Popperian fallibilism at the end of the road he is travelling.

6. Uses and misuses of history: rational reconstruction or mystification?

Feyerabend's challenge to Lakatos' rationalism takes different forms. In general he objected that Lakatos' representation of science is just a misleading reconstruction: philosophy of science should be based on real science, rather than on a convenient

[22] These include 'the basic tenets of deductive logic and intuitive rules for weighing evidence (especially the principle that special weight is to be given to a theory's predictive success)' (Worrall, 1989, p. 377). On the value of 'novelty' for theory confirmation, see Worrall (1985, 1989) and Zahar (1989). Lakatos' criteria of progress have been applied in many different case studies both in the natural sciences and social sciences. It goes beyond the scope of this paper to assess them. The question of course remains whether these principles are stable or change alongside changes in substantive theories. I shall deal with this problem in the next section.

[23] See, for example, the exhortation 'to posit some extramethodological inductive principle to relate—even if tenuously—the scientific gambit of pragmatic acceptances and rejections to verisimilitude. Only such "inductive principle" can turn science from a mere game into an epistemologically rational exercise; from a set of light-hearted sceptical gambits pursued for intellectual fun into a—more serious—fallibilist venture of approximating the Truth about the Universe' (Lakatos, 1971a, p. 101; see also ibid., pp. 97, 108–109 and Lakatos, 1974, pp. 154–159).

[24] I have argued this point extensively in Motterlini (1995). See also Larvor (1998), pp. 62–65.

myth. The very Popperian project is prejudiced from the start, in so far as it posits the existence of accepted (and presumably universal) 'basic' value judgements on singular scientific achievements. Assessment criteria are not fixed, and do change along with changes at the theoretical level. A scientific revolution is truly revolutionary precisely because it leaves nothing unchanged. Thus, in a sense, the better a methodology seems to capture the allegedly unchanging rationality of science, the greater its mystification.

Feyerabend also questions the usual, 'whiggish' reconstruction of some crucial scientific episodes. The new astronomy of Copernicus, Kepler and Galileo in fact 'superseded' the Ptolemaic approach in a way that would seriously embarrass a rationalist philosopher. The history of 'successful', progressive science is replete with scientists defending obsolete theories in spite of anomalous evidence and indeed in violation of all the normative rules invoked to capture the rationality of science. Why should we even assume that progress has occurred? Why move from the assumption that science has de facto developed to some given methodology in the last three hundred years?

There are various ways of rebutting Feyerabend's challenge. First, we need to clarify the concept and scope of 'rational reconstructions'. Secondly, we need to assess Feyerabend's claims regarding the (non-)stability of methodological principles (which principles should we expect to be stable? And at which level of generality should we expect to locate them?). Regarding the first issue, Lakatos is well known for his parody of Kant's maxim: 'philosophy of science without history of science is empty, history of science without philosophy of science is blind'. Following Hegel, Lakatos starts from the premise that knowledge does grow, suggesting the task for the philosopher of extracting rationality from the historical development. But following Popper, he takes a fallibilist stance on the issue of what constitutes the rationality of science (or, which is the same, the principles that govern its growth). This is why we need to test our theories of rationality against the history of science. Rational reconstructions, in the context of such a project, are not 'mystifying' attempts at appraising past beliefs according to new standards. They are rather to be taken as heuristic devices that lead us to find historical data and sketch a (fallible) explanation of the whole process. The historian and the philosopher will both learn from such an exercise: the former by discovering 'new' facts, the latter by testing his own theories of rationality. Hence, where Feyerabend sees a vicious circle of justificaton, Lakatos envisages a virtuous interaction between philosophy and history of science. In a letter to Lakatos, Feyerabend implicitly recognised this by praising the methodology of scientific research programmes (and criticising Kuhn) precisely on the ground of its fertility as a historiographical approach (cf. Lakatos & Feyerabend, 1999, p. 325). Lakatos, in turn, conceded that the virtuous circle can start to spin only if enough pluralism is tolerated. The proliferation of different points of view, the comparison between different reconstructions are necessary to learn from history and escape the influence of 'bad' philosophies (i.e. theories of rationality). But here again perhaps Lakatos is dangerously unbalanced between the Hegelian and the Popperian stand. In particular, the little trace of historicism that Lakatos slipped into the static and ahistorical Popperian account of science may turn into a very dangerous

explosive that must be handled with great care, as Lakatos himself acknowledged in 1961.[25]

As is well known, Feyerabend did not care very much: history ends up blowing up all reference points, theories and criteria. Lakatos did not want to go as far as that. In particular, he wanted to rescue a set of immutable principles that would allow the rationalisation of theory change and perhaps also methodology change.

As is also well known, Lakatos intended to write a history of the interaction between methodological standards and scientific achievements. The book was to be called *The Changing Logic of Scientific Discovery*. When he died of a heart attack on the 2nd of February 1974 , only one chapter of that book was written, which was later published as 'Newton's Effect on Scientific Standards'.[26] The core of the paper deals with the battle between Cartesians and Newtonians over the standards of scientific proofs and criticism. In a somehow cynical unpublished note from the same period, Lakatos draws an analogy between his own project and the project of a dogmatic philosopher dealing with a revolution in moral standards. In the end, he argues, the revolutionaries win by shifting the very notion of what constitutes moral behaviour. Eventually, they put forward 'a theory in the light of which their revolution was not just a change in subjective fashion, but progress towards moral truth'.[27]

The revolutionaries here are the Newtonians, who violated the standards of the scientific community of their time. Lakatos points out how Newton's theory of method (i.e. his *Regulae philosophandi*) is at odds with Newton's own scientific practice (i.e. his method). Thanks to a historical paradox, Lakatos claims, that Newtonian professed methodology was to create philosophy of science. 'Philosophy unaware of the split between the wonderful Newtonian method as practised and the

[25] At that time he wrote: 'I am afraid that some ardent Popperite may already be rejecting all that I am about to say [but] I am quite convinced that *even the poverty of historicism is better than the complete absence of it*—always providing of course that it is handled with the care necessary in dealing with any explosives' (Lakatos, 1959–1961, p. 61; emphasis added).

[26] Lakatos (1978a), Ch. 5. Early drafts of this paper were written in 1963 and1964, but Lakatos returned to it several times in the following years.

[27] Here is the passage in full: 'Let us imagine that mankind has set up moral standards and maintained them for centuries without anybody being able to live up to them. Some claimed, hypocritically, that moral acts—and, indeed, moral men, did exist; others kept unmasking these reports and either claimed that the high moral standards will necessarily remain utopian forever, or set up programme after programme for making man moral and noble. Then, suddenly, some actually started behaving in a new way which soon struck most people as truly moral. First they hailed them as having finally realised the old dreams. But then slowly it dawned on them that although the new men's acts were undoubtedly virtuous, they did not in fact comply with the old standards. They tried gradually to weaken the standards, but no matter how one weakened the standards they remained unsatisfied . . . People had two ways out. The sceptics celebrated the end of all morality. But a new school of dogmatists accepted the *new moral facts* and devised *new moral standards* in the light of which their morality—a new morality—could be seen. This was then a *revolution*—first in morality, then in moral theory. Not necessarily the last one: inconsistencies remained, new types of moral acts (expressed in "moral basic propositions") further upset the new standards. The problem of morality was shifted. Moreover, the revolutionaries set up a theory of moral progress which sets up standards for standard change; a theory in the light of which their revolution was not just a change in subjective fashion, but progress towards moral truth' ('The Intellectuals' betrayal of Reason', Archive 8.2; now in Motterlini (Ed.), 1999, Appendix C, pp. 396–397).

mad Newtonian method as professed, tried to clarify the professed method, and thus turned into a mad study of madness' ('The Rise of Defensive Positivism', Archive 5.5).

Underlying the paradox is the fact that Newton took his methodological standards from past centuries dominated by the all-pervading idea that religious knowledge was certain and indubitable; by analogy, science was expected to respect similar standards. In fact, the split brought on by the Reformation (though against the original intentions of the founding fathers) raised the problem of fallibilism: how is the Church expected to be the authority on its own infallibility, given that the point in question is precisely whether the Church is the 'true' authority on religious matters? This question was to unleash a sceptical crisis not only in theology but also, shortly thereafter, in the sciences and in all other areas of human knowledge.[28] A century later, the stunning success of modern natural sciences, and primarily of Newton's research programme, was to breathe new life into the battle between the dogmatics and the sceptics.[29] Lakatos refuses to side with either the radical dogmatists or the radical sceptics. Instead, he points to a new appraisal of this controversy, by recognising 'the basic unity of opposites (dogmatism-scepticism) and the possibility of their dialectical "Aufheben" [superseding]'.[30] Yet Lakatos regards the 'modern betrayal of reason', consisting in the intellectual attack on the objective epistemological value of the exact sciences, as 'criminal'.[31] If, on the one hand, Feyerabend had shown many different ways in which any principle of rationality can turn into a prison, and has consequently claimed along with the sceptics that the betrayal of 'reason' by man is preferable to the betrayal of man by reason, on the other hand, Lakatos challenges his rival's anarchism insisting on the bias underlying the sceptico-dogmatist's point of view, by showing how radical relativism ultimately leads to forms of intellectual surrender which maintain that *might is right*.

Worthy theologian of rationality, in order to propose his (fallibilist) stand as an alternative to dogmatism and scepticism, Lakatos himself had to face the challenge of the Greek Pyrrhonians, whose arguments, after lying forgotten for centuries, had suddenly come to the fore, especially in regard to the primary epistemological problem brought up by the Reformation. Luther's denial of the authority of the Church and his assertion of a new 'rule of faith' (i.e. 'true' is what conscience is compelled to believe when reading Scripture) for determining religious truths constitutes a rather neat example of the 'problem of the criterion' as it appeared in Sextus Empiricus:

---

[28] Consider, for instance, the shift in the criterion of truth from Erasmus to Descartes described by Popkin (1979), Ch. 1.

[29] On the one hand, 'some radical dogmatists, undeterred by the long series of defeats of reason in human affairs, have been trying to generalise Newtonian method to social, ethical and political problems'. On the other, 'some radical sceptics, undeterred by the long series of successes of Newtonian science, have been trying to show that all these successes were sham successes and even the best theories of the exact sciences were nothing more than irrational (if possibly "great") beliefs' ('The Intellectuals' betrayal of Reason', in Motterlini (Ed.), 1999, p. 397).

[30] Lakatos conversation with Popkin, in Breck and Yourgrau (1970), p. 22.

[31] 'The Intellectuals' betrayal of Reason', in Motterlini (Ed.), 1999, p. 397.

> In order to decide the dispute which has arisen about the criterion, we must possess an accepted criterion by which we shall be able to judge the dispute; and in order to possess an accepted criterion, the dispute about the criterion must be first decided. And when the argument thus reduces itself to the form of circular reasoning the discovery of the criterion becomes impracticable, since we do not allow the Dogmatic philosophers to adopt a criterion by assumption, while if they offer to judge the criterion by a criterion we force them to a regress ad infinitum. (Sextus Empiricus, *Outlines of Pyrrhonism*, in Popkin, 1979, p. 3)

Feyerabend's criticism of Lakatos is clearly a revival of this classical Pyrrhonian challenge. Despite the paradoxical tone of these passages, according to Feyerabend, history reveals diversity in method as well as diversity in science. Hence, change occurs at the same time at both levels. Once such a holistic view of change is held, relativism is unavoidable—see Worrall (1988). But according to Lakatos method only *appears* to change, where in fact it does not. The distinction between 'professed' and 'practised' method is thus crucial for the Lakatosian project. In the Newtonian case, we should focus not on what the Newtonians (or, for that matter, the Cartesians) *said*, for there is surely a disagreement at that level. We should focus instead on what they *did*—and it is a most important task of the historian to articulate exactly what it is that they did.[32] Feyerabend, according to this view, stretches the term 'methodology' to such an extent that the dispute on the criterion can never be decided, because no criterion is fixed. Lakatos allegedly works with a narrower concept of methodology and therefore rationality. The idea is that even though we have 'learned to do science better alongside doing better science', we maintain that some set of standards are not themselves historical (Worrall, 1988, p. 272). The task of Lakatos' 'Changing Logic of Scientific Discovery', then, is (again quite paradoxically) to articulate the *unchanging* standards. Lakatos can thus consistently maintain that we should be able to rank Copernicus' research programme better than Ptolemy's and Einsten's better than Newton's with respect to some methodological standards implicitly endorsed by the scientific community which are independent from professed methodologies. In this way only he can defend rationality by its collapse into either sceptical relativism (disguised as anarchism) or the cunning of reason (disguised as authoritarian elitism).

To accomplish such a task would be an empirical, rather than a purely *a priori*, matter. Of course, the burden of proof is with the rationalist, because he has to show that an invariant core really exists. Here is the rationalist enterprise laid bare in front of us. The 'Changing Logic', however, was never written and we cannot close the gap between the present 'Lakatos Reconstruction' and the 'real thing'. For Lakatos clearly saw what the task was, and also what had to be done in order to achieve his goal. But, in light of the historical case studies he and his students had examined,

[32] On the plausibility of Lakatos' resort to scientists' 'false (methodological) consciousness' in accounting for the rationality of theory change, see Newton-Smith (1981), p. 71, Laudan (1989), p. 321; for a reply, see Worrall (1985, 1989).

Lakatos possibly had not made up his mind as to whether the odds were for or against him. Or, perhaps, he made up his mind only near his end. In a letter of 10th January 1974 to Feyerabend he wrote:

> Dear Paul, I was amused by your suggestion that the scientific revolution was a revolution in the standards. This is of course the story I encapsulated in the announced title: *The Changing Logic of Scientific Discovery*. A chapter of this book would have been my Newton paper in which I discuss this change in standards in detail. The trouble was that Worrall and Zahar persuaded me that the *standards* which I ascribed to the seventeenth century *were already there* at the end of the sixteenth. Bashi Sabra almost beat me up since according to him these standards were there in the age of Ptolemy. Now your letter caused a conversion effect, and now I think that my Newton paper is perfectly correct and can be made consistent with the Copernicus paper, and now I am going to publish it. (Motterlini, 1999, p. 355; see also p. 357; emphasis added)

7. The dark side of the truth

Lakatos' syncretic programme, and the forms it took in the different phases of his intellectual life, has constituted the main theme of this paper. Lakatos walks on a thin edge. The balance between the Hegelian and the Popperian elements in his philosophy is never a stable one. Such a precarious equilibrium is in the very nature of his rationalistic project. The difficulty arises first in Lakatos' philosophy of mathematics, in particular from his attempt to claim a normative role for heuristics, to be grounded on the objective (and effective) pattern of the growth of knowledge. This project will appear to fail as, many years later, Lakatos (the philosopher of natural science) restricts the impact of his methodological standards to retrospective descriptions. A similar apparent contrast characterises his attempts to play with fallibilism: Lakatos' positions range from extreme liberalism (when he stresses the importance of the proliferation of different research programmes for criticism and the growth of knowledge) to conservatism (when he claims that degenerating research programmes should not be funded). Sometimes the attempt to join in a synthesis the best aspects of the two traditions is clearly fruitful, for example when Popperian fallibilism is used to bash the misuses of history in its 'cunning of reason' version by the Hegelians, and Hegelianism to criticise Popper for disregarding the dynamic dimension of scientific rationality. At other times, Lakatos seems to oscillate between mutually incompatible standpoints, on the verge of internal inconsistency. He plays the sceptic who claims that there is no immutable statute law by which one can demarcate good and bad science. But in the same period he is involved in the project of distinguishing between a historically invariable method of science and historically variable methodologies, in order to save the very concept of a 'fixed' canon of scientific rationality. Lakatos 'the friend of Reason' was perhaps a *sui generis* dogmatist who owed a lot to sceptical arguments, in philosophy of science as well as political philosophy and metaphysics—reviving a glorious theological image, for instance,

Lakatos claims that 'the sentences uttered by God in the Book of Nature were of an infinite length', and therefore no human statements (obviously of a finite length) can express a Natural law.

The gap between all these stands can be bridged only partially. Perhaps, as Naturalized Epistemology suggests, what is questionable is the very object of Lakatos' project:[33] the justification of objective canons of evaluation of scientific claims that are genuinely epistemological and not reducible to psychological, biological, neurocomputational or social factors.[34] Paraphrasing the opening words of Kuhn (1962), history, viewed as a repository for more than anecdote or chronology, has produced a decisive transformation of the image of science by which we were possessed. In particular, history reveals that a theory of scientific rationality should be richer, more articulated and dynamic than the one fossilized by the 'old fashioned Popperian-type' philosophy of science. Science, as historically practised, indicates that methodologies, historiographies, ideologies, scientists' value judgments and theological and scientific standards are intimately intertwined. Neither, after Lakatos, should we fear to admit that the philosophy of science is always 'philosophy fabricating its own examples'. The problem, once more, is to provide (meta-)criteria according to which one fabrication is better than another. In a dynamic setting this amounts to taking the problem of justifiable changes—at the level of scientific theories as well as of scientific standards of evaluation—to be central. Moreover, since we are not prepared to accept such standards *a priori*, we need to test these principles against historical cases and general (scientific) practice. Thus, to justify a set of principles that characterise scientific practice in a given domain, one generates principles that conform to commonly accepted cases and practices. If such principles establish judgements that do not conform to general cases and practices, then principles are modified; if, however, such modifications produce principles that are intuitively unacceptable, then cases and practices can in turn be rejected or modified. Perhaps the 'reflective equilibrium' between principles and cases generated by means of such a procedure is both *unavoidable* and inherently *unstable*.[35] And so is the case for any normative theory of scientific rationality based upon it—including Lakatos'. On this view, Lakatos' sophisticated attempt at implementing Hegel's ideas within the Popperian system and 'clearing rationality in the jungle of history' may well turn out to be an illusion. Perhaps Lakatos' project is to remain incomplete. And this is why, from time to

[33] See, for example, Kitcher (1992) and Kitcher (1993), Ch. 1.

[34] Naturalism has many different versions. See Quine (1969) for the seminal advocacy of the replacment of epistemology by empirical psychology. See, for example, Giere (1988) for the extension of this project in the direction of cognitive psychology, and Churchland (1989) in the direction of neurocomputation. See Bloor (1979) and Latour (1987) for the reduction of epistemology to a social dimension.

[35] The process may be circular but virtuous according to Goodman (1983), who introduced the concept in his *Fact, Fiction and Forecast*, first published in 1965. The argument was baptized 'reflective equilibrium' by Rawls (1971, pp. 48–49), who applied the notion to moral theory, in particular, to the relation between moral principles and one's intuitive judgments. (Rawls used the term 'reflective' because 'we know to what principles our judgements conform and the premises of their derivation' and 'equilibrium' because 'our principles and judgment coincide'. See Stein (1996), Ch. 5 for a discussion on the validity and the implication of the argument for the justification of normative principles of reasoning.

time, Lakatos concedes something to historicism, anarchism and elitism. After all, Popper's clear-cut falsificationism and Feyerabend's anarchic slogans and elitist positions equating might and right immediately clear up science in its obscure aspects, but—as Lakatos himself would have put it—they are like '*Lucifer*, the chap who brings *false* light', while 'I'm shrouding you in the darkness of truth'.[36]

Acknowledgements

I should like to express my gratitude to Francesco Guala, Donald Gillies, Giulio Giorello, Brendan Larvor, John Worrall and an anonymous referee for many useful suggestions and detailed comments. The Centre for Philosophy in Natural and Social Sciences and the Imre Lakatos Memorial Fund are also gratefully acknowledged. Any mistakes fall of course under my own responsibility.

References

Bloor, D. (1979). *Knowledge and social imagery*. Chicago: University of Chicago Press.
Breck, A. D., Yourgrau, W. (Eds.). (1970). *Physics, logic and history*. New York: Plenum.
Churchland, P. M. (1989). *A neurocomputational perspective*. Cambridge, MA: MIT Press.
Duhem, P. (1954). *The aim and structure of physical theory*. English trans. of 2nd ed. Princeton: Princeton University Press. First published 1906; 2nd French ed. 1914.
Feferman, S. (1978). The logic of mathematical discovery vs. logical structure of mathematics. In P. D. Asquith, & I. Hacking (Eds.), *PSA 1978: Proceedings of the biennial meeting of the Philosophy of Science Association* (Vol. 2) (pp. 309–327). East Lansin, MI: Philosophy of Science Association.
Feyerabend, P. K. (1975). *Against method. Outline of an anarchistic theory of knowledge*. London: New Left Books.
Feyerabend, P. K. (1976). On the critique of scientific reason. In C. Howson (Ed.), *Method and appraisal in the physical sciences*. Cambridge: Cambridge University Press.
Fine, A. (1978). Conceptual change in mathematics and science: Lakatos' stretching refined. In: P. D. Asquith, & I. Hacking, (Eds.), *PSA 1978: Proceedings of the 1978 biennial meeting of the Philosophy of Science Association* (Vol. 2, pp. 328-341). East Lansing, MI: Philosophy of Science Association.
Forrai, G. (1993). From the method of proofs and refutations to the methodology of scientific research programmes. *International Studies in the Philosophy of Science, 2*, 161–175.
Giere, R. (1988). *Explaining science*. Chicago: University of Chicago Press.
Goodman, N. (1983). *Fact, fiction, and forecast*. Cambridge, MA: Harvard University Press. First published in 1965.
Hacking, I. (1979). Imre Lakatos' Philosophy of Science. *British Journal for the Philosophy of Science, 30*, 381–410.
Hegel, G. W. F. (1977). *Phenomenology of spirit*. Oxford: Clarendon Press. First published in 1807.
Hegel, G. W. F. (1991). *Elements of the philosophy of right*. Cambridge: Cambridge University Press. First published 1820.
Kadvany, J. (2001). *Imre Lakatos and the guises of reason*. Durham, NC and London: Duke University Press.
Kitcher, P. (1992). Naturalists return. *Philosophical Review, 101*, 52–114.
Kitcher, P. (1993). *The advancement of science*. Oxford: Oxford University Press.

[36] Lakatos to Feyerabend, 23 January, 1973, in Motterlini (1999), p. 313.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
Lakatos, I. (1959–1961). What does a mathematical proof prove? Published as Chapter 4 of Lakatos (1978b).
Lakatos, I. (1970). Falsificationism and the methodology of scientific research programmes. In I. Lakatos, & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–196). Cambridge University Press: Cambridge Republished as Chapter 1 of Lakatos (1978a) (pp. 8–101).
Lakatos, I. (1971a). History of science and its rational reconstructions. In R. Buck, & R. Cohen (Eds.), *P.S.A. 1970, Boston Studies in the Philosophy of Science* (8) (pp. 91–135). Dordrecht: Reidel. Republished as Chapter 2 of Lakatos (1978a) (pp. 102-138).
Lakatos, I. (1971b). Replies to critics. In R. Buck, & R. Cohen (Eds.), *P.S.A 1970, Boston Studies in the Philosophy of Science* (8) (pp. 174–182). Dordrecht: Reidel.
Lakatos, I. (1973). The problem of appraising scientific theories: Three approaches. Published as Chapter 6 of Lakatos (1978b).
Lakatos, I. (1973/1976). Understanding Toulmin. Minerva, 14, 126–134. Republished as Chapter 11 of Lakatos (1978b).
Lakatos, I. (1974). Popper on demarcation and induction. In P. A. Schilpp (Ed.), *The philosophy of Karl Popper* (pp. 241–273). La Salle, IL: Open Court. Republished as Chapter 3 of Lakatos (1978a) (pp. 139-167).
Lakatos, I. (1976). *Proofs and refutations. The logic of mathematical discovery* (J. Worrall & E. Zahar, Eds.). Cambridge: Cambridge University Press.
Lakatos, I. (1978a). *The methodology of scientific research programmes: Philosophical papers. Vol. I* (J. Worrall & G. Currie, Eds.). Cambridge: Cambridge University Press.
Lakatos, I., & Feyerabend, P. K. (1999). Correspondence 1968-1974. In M. Motterlini (Ed.), *For and against method* (pp. 119–373). Chicago: University of Chicago Press.
Larvor, B. (1998). *Lakatos: An introduction*. London: Routledge.
Larvor, B. (1999). Lakatos' mathematical Hegelianism. *The Owl of Minerva*, *31*(1), 23–44.
Latour, B. (1987). *Science in action*. Cambridge, MA: Harvard University Press.
Laudan, L. (1989). If it ain't broke, don't fix it. *British Journal for the Philosophy of Science*, *40*, 369–375.
Long, J. (1997). Lakatos in Hungary. *Journal of the Philosophy of the Social Sciences*, *28*(2), 244–311.
Motterlini, M. (1995). Has Lakatos really gone a long way towards epistemological anarchism? *Epistemologia*, *XVIII*, 215–232.
Motterlini, M. (Ed.). (1999). For and against method. Including Lakatos' lectures on method and the Lakatos–Feyerabend Correspondence. Chicago: University of Chicago Press.
Motterlini, M. (2002). Lakatos between the Hegelian devil and the Popperian blue sea. In G. Kampis, L. Kvasz, & M. Stoeltzner (Eds.), *Appraising Lakatos, mathematics, methodology, and the man*. Dordrecht: Kluwer.
Musgrave, A. (1976). Method or madness? In R.S. Cohen, P.K. Feyerabend, M.W. Wartofsky (Eds.), *Essays in memory of Imre Lakatos*. Boston Studies in the Philosophy of Science. Dordrecht: Reidel.
Newton-Smith, W. (1981). *The rationality of science*. London: Routledge.
Poincaré, H. (1982). *Science and hypothesis*. (English authorized trans. by G. B. Halsted). In H. Poincaré, *The foundation of science*. Washington, DC: University Press of America. First published 1913; original French version, La science et L'hypothèse, 1902.
Pólya, G. (1954). *Mathematics and plausible reasoning*, 2 vols. Princeton: Princeton University Press.
Pólya, G. (1973). *How to solve it*. Princeton: Princeton University Press. First published 1945.
Popkin, R. H. (1979). *The history of scepticism from Erasmus to Spinoza*. Berkeley: Berkeley University Press.
Popper, K. R. (1959). *The logic of scientific discovery*, expanded English edition. London: Hutchinson.
Quine, W. V. O. (1969). Epistemology naturalized. In *Ontological relativity and other essays* (pp. 69-90). New York: Columbia University Press.
Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.
Rediker, M. (1987). *Between the devil and the deep blue sea*. Cambridge: Cambridge University Press.
Stein, E. (1996). *Without good reason. The rationality debate in philosophy and cognitive science*. Oxford: Clarendon Press.
Szabó, A. (1958). 'Deiknymi' als mathematischer Terminus für 'Beweisen'. *Maia*, *10*, 1–26.

Szabó, A. (1960). Anfänge des Euklidischen Axiomensystems. *Archive for the History of Exact Sciences*, *I*, 27–106.
Toulmin, S. (1972). *Human understanding*. Oxford: Oxford University Press.
Truesdell, C. (1984). *An idiot's fugitive essays on science. Methods, criticism, training, circumstances*. New York, Berlin, Heidelberg, Tokyo: Springer.
Wittgenstein, L. (1958). *Philosophical investigations*. Oxford: Oxford University Press.
Worrall, J. (1985). Scientific discovery and theory-confirmation. In J. Pitt (Ed.), *Change and progress in modern science*, pp. 301-331. Dordrecht: Reidel.
Worrall, J. (1988). The value of a fixed methodology. *British Journal for the Philosophy of Science*, *39*, 263–275.
Worrall, J. (1989). Fix it and be damned: A reply to Laudan. *British Journal for the Philosophy of Science*, *40*, 376–388.
Worrall, J. (mimeo). How to deduce theories from phenomena: Newton on scientific method.
Zahar, E. (1983). Logic of invention or psychology of research? *British Journal for the Philosophy of Science*, *34*, 234–261.
Zahar, E. (1989). *Einstein's revolution*. A study in heuristics. La Salle, IL: Open Court.
Zheng, Y. (1990). From the logic of mathematical discovery to the methodology of scientific research programmes. *British Journal for the Philosophy of Science*, *41*, 377–399.

Further reading

Lakatos, I. (1960). Necessity, Kneale and Popper. Published as Chapter 7 of Lakatos (1978b).
Lakatos, I. (1963–1964a). Newton's effect on scientific standards. Published as Chapter 5 of Lakatos (1978a).
Lakatos, I. (1963–1964b). Proofs and refutations. British Journal for the Philosophy of Science, 14, 1–25, 129–139, 221–243, 296, 342. Repubished in revised form as part of Lakatos (1978a).
Lakatos, I. (1968). Changes in the problem of inductive logic. In I. Lakatos (Ed.), *The problem of inductive logic* (pp. 315–417). North Holland Publishing Co: Amsterdam. Republished as Chapter 8 of Lakatos (1978b).
Lakatos, I. (1978b). Mathematics, science and epistemology: Philosophical papers. Vol. II (J. Worrall & G. Currie, Eds.). Cambridge: Cambridge University Press.
Lakatos, I. (1999). Lectures on scientific method. In M. Motterlini (Ed.), *For and against method* (pp. 19–109). Chicago: University of Chicago Press.
Szabó, A. (1978). *The beginning of Greek mathematics*. Boston & London: Reidel.