60% Proof
Lakatos, Proof, and Paraconsistency

GRAHAM PRIEST

SCHOOL OF PHILOSOPHY
THE UNIVERSITY OF MELBOURNE
g.priest@unimelb.edu.au

NEIL THOMASON

SCHOOL OF PHILOSOPHY
THE UNIVERSITY OF MELBOURNE
neilt@unimelb.edu.au

1   INTRODUCTION

Imre Lakatos' *Proofs and Refutations*[1] is a book well known to those who work in the philosophy of mathematics, though it is perhaps not widely referred to. Its general thrust is out of tenor with the foundationalist perspective that has dominated work in the philosophy of mathematics since the early years of the 20th century. It seems to us, though, that the book contains striking insights into the nature of proof, and the purpose of this paper is to explore and apply some of these.

2   PROOF IN THE HISTORY OF MATHEMATICS

2.1   THE EUCLIDEAN PARADIGM

The contemporary conception of proof in logic and mathematics is dominated by what one might call the Euclidean paradigm, since it is modeled on Euclid's *Elements*. One starts by laying down certain axioms which are clearly true. A proof is then a logically valid deductive argument starting from the axioms, and ending in a theorem. Since valid deductive arguments preserve truth, the conclusion of the argument, the theorem, is thereby established as true.

---

[1]Lakatos (1976). Page references in what follows, unless otherwise specified, are to this. Italics in all quotes are original.

It takes little knowledge of the history of mathematics to realise that, at least until the 20th century, mathematical proof—the form of argumentation used by mathematicians to support their results—did not, generally speaking, work in this way. This is for the very simple reason that, outside of geometry, there weren't any axiom systems. No one ever specified axioms for the real numbers, negative numbers, infinitesimals. Indeed, even the natural numbers were not axiomatized until the work of Dedekind in the late 19th century.

At least in the history of mathematics, then, mathematical proof was not Euclidean proof.

2.2 LAKATOSIAN PROOF

The conception of proof that emerges from Lakatos' book, and one that appears to fit the history of mathematics more accurately, is quite different. A proof, in this sense, is indeed an argument, but the argument starts from opportunistic places, simply with claims that appear to be true, and then proceeds by argumentative steps that appear to be right, till the conclusion is established.

What is proved is proved very fallibly. Counter-examples to the theorem or other parts of the proof may turn up; and when they do, we have to revise the theorem, the proof, or the starting points, appropriately. The Euclidean notion of proof is foundationalist: proofs are built up on the solid foundation of axioms, with the theorems at the top. By contrast, Lakatosian proofs are anti-foundationalist. To use a metaphor of Karl Popper,[2] theorems are like buildings built on swamps. The arguments are like piles that are driven down into the swamp to support the building, but they never reach bedrock, and may have to be replaced or driven down further, as required. Call the places from which they start axioms if you like, but an axiom is just a good place to stop—at least for the present. It seems true; but we may have to come back and revise that judgment.

2.3 AN ILLUSTRATION: EULER'S THEOREM

Lakatos illustrates his notion of proof with Euler's Theorem concerning polyhedra. If the number of faces, vertices, and edges of a polyhedron are F, V, and E, respectively, then $V - E + F = 2$. Cauchy proved this as follows. The proof is illustrated for a cube in Fig. 1 (reproduced from Fig. 3, p. 8). Remove one face, and stretch out the rest on a flat surface. This reduces F by 1, but leaves V and E alone. For the resulting figure we therefore need to prove that $V - E + F = 1$. Now, for each polygonal face, join its corners until the only polygons left are triangles. Each joining inserts one extra edge and one extra face. $V - E + F$ therefore remains unchanged. Now, starting at the outside, remove one triangle at a time. In removing a triangle, there are two possibilities: (a) one edge and one face disappear, (b) two edges, one vertex, and one face disappear. Both

---

[2] *Logic of Scientific Discovery*, Section 30 (1934).

possibilities leave $V - E + F$ unchanged. In the end, only one triangle is left. This has three edges, three vertices and one face. Thus, $V - E + F = 1$, as required.
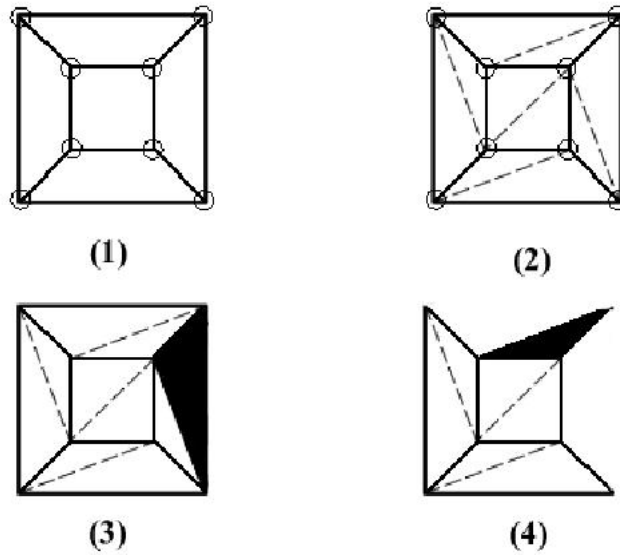

FIGURE 1

So much for the proof. But we can now note that there are some counter-examples (refutations). Consider, for example, Fig 2. In Fig 2a (the "picture frame"), $V - E + F = 2$ alright, but the proof does not work. If we remove the top, we cannot stretch the rest out on a flat surface. Fig 2b (the "crested cube") is even worse. For this, $V = 16$, $E = 24$, and $F = 11$. So $V - E + F = 3$. Lakatos calls these *local* and *global* counter-examples respectively (to just the proof, or the proof *and* the theorem).
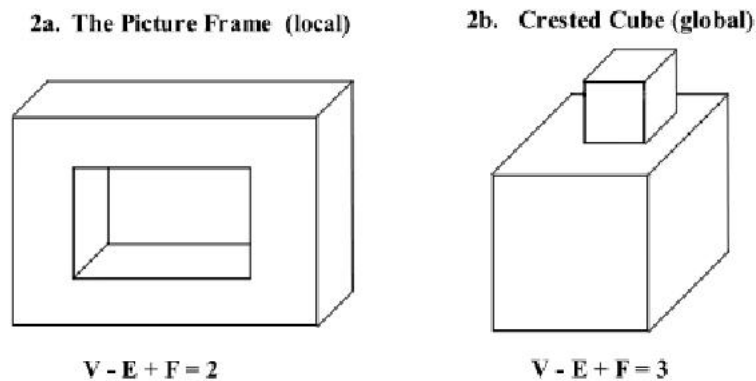

2a. The Picture Frame (local)   2b. Crested Cube (global)

$V - E + F = 2$ $V - E + F = 3$

FIGURE 2

2.4   RESPONSES TO COUNTER-EXAMPLES

Several responses to counter-examples are possible. One is 'monster barring'. We declare that the counter-examples are not really polyhedra: they are just monsters, and should be ignored. The response is obviously quite *ad hoc*: we have no principled reason to rule them out—except that they violate the theorem. Another response, Lakatos terms 'withdrawal to a safe domain'. We observe that neither of the counter-examples is convex. We therefore rephrase our theorem as 'For all convex polyhedra, $V - E + F = 2$'. This is better than monster-barring, but the move is still a crude one. We may well have over-withdrawn. Most importantly, the move ignores the information provided by the proof.

By far the best response, according to Lakatos, is to analyse the proof in the light of the counter-examples to see what has gone wrong, and how it can be revised to avoid the anomoly. Thus, the picture frame exposes a hidden assumption of our proof, namely that the removal of a face will leave something that can be stretched out. If we remove the bottom of the crested cube, it can be stretched out, but the annular face cannot be triangulated in the appropriate way. Define a polyhedron to be a *Cauchy polyhedron* if it is such that the operations of the proof can be performed on it. In the light of the analysis, what the proof shows is that for all Cauchy polyhedra, $V - E + F = 2$.

Proof analysis is clearly methodologically superior to the other responses mentioned. Not only does it leave us with an improved proof, it digs out hidden assumptions, so we learn things we did not know before, and it generates new concepts of a potentially fruitful kind.

This is just a small example of the wealth of example and methodological discussion in Lakatos' book. But it will suffice for the nonce. Lakatosian proof is fallible, but the exploitation of the fallibility gives rise to mathematical progress.

3   PROOF IN CONTEMPORARY MATHEMATICS

3.1   THE 19TH CENTURY DRIVE TO RIGOR

So much for proof in the history of mathematics. It might well be thought that the situation is now essentially different. From a modern perspective, past mathematical practice was rather sloppy. Concepts were not tied down in advance; arguments were loose. No wonder that results were so fallible. The 19th century drive for rigor, through the work of mathematicians such as Weierstrass, produced a mathematics that has put an end to all that.

Rigor requires three things:

1. First, all concepts must be defined in terms of a set of basic concepts. (The definitions cannot go on indefinitely on pain of infinite regress.)

The 19th century mathematicians worked hard to define all their concepts in more basic terms. Geometric entities, such as lines, could be defined as sets of real numbers. Real numbers could be defined as sets of rational, and ultimately natural, numbers. And natural numbers could be defined themselves as certain sets. So sets emerged as the basic mathematical concept.

2. Next, though the basic concepts cannot be defined in terms of anything more basic, their behaviour needs to be tied down. This is done by spelling out a set of axioms which characterise them. After a period of various debates, mathematicians settled for contemporary Zermelo Fraenkel set theory.

3. Finally, the logical rules to be applied in inferring from the mathematical axioms must themselves be tied down. This produced so called 'classical logic', the logic of Frege and Russell. This, too, was spelled out axiomatically (though most contemporary logicians prefer some kind of natural-deduction formulation).

Lakatos was quite well aware of this, of course. In a long note (pp. 55-6), he suggests that the method of proof-analysis itself gave rise to the increasing rigor. Be that as it may; the result is a system of mathematics where Euclideanism reigns. Not in practice, of course. Mathematicians do not give formal axiomatic proofs—certainly not proofs in the logician's sense. They give informal arguments, much as they always have done, starting with things which are generally accepted, and with lots of gaps in the argument. But in theory, anyway. The arguments could, in principle, be spelled out as fully formal axiomatic proofs; and if this is not the case, a mistake has been made.

3.2 LAKATOS AND HIS EDITORS

But, Lakatos thinks, this has not changed things essentially. Certain things may be marked out as axioms, but these are simply conventionally accepted stopping points; in due course, they themselves may be contested and revised in the light of counter-examples. As he puts it (p. 56):

> ...*different levels of rigor differ only about where they draw the line between the rigor of proof-analysis and the rigor of proof*, i.e., *about where criticism should stop and justification should start*. 'Certainty is never achieved', 'foundations' are never found—but the 'cunning of reason' turns each increase of *rigor* into an increase of *content*, in the scope of mathematics.

Lakatos never lived to see the publication of his book. It was brought to fruition by its two editors, John Worrall and Elie Zahar, whose view on the matter differs from the one Lakatos expresses here. They say (pp. 56-7):

We believe that [Lakatos' view] underplays a little the achievement of mathematical 'rigorists'. The drive towards rigor in mathematics was, it eventually transpired, a drive towards two separate goals, only one of which was attainable. These two goals are, first, rigorously correct arguments or proofs (in which truth is infallibly transmitted from premises to conclusions) and, secondly, rigorously true axioms, or first principles (which are to provide the original injection of truth into the system—truth would then be transmitted to the whole of mathematics *via* rigorous proofs). The first of these two goals turned out to be attainable (given, of course, certain assumptions), whilst the second proved unobtainable.

Frege and Russell provided systems into which mathematics could be (fallibly) translated..., and in which the rules of proof are finite in number and specified in advance. It also turns out that one can show (it is here that the assumptions just referred to come in) that any sentence that can be proved using these rules is a valid consequence of the axioms of the system (i.e., if these axioms are true, the sentence proved *must* also be true). In these systems there need be no 'gaps' in proofs, and whether a string of sentences is a proof or not can be checked in a finite number of steps.... There is no serious sense in which such proofs are fallible. (It may be that everyone who has ever checked some such proof made an inexplicable error, but that is not a serious doubt. It is true that the informal (meta-)theorem that such valid proofs transmit truth may be false—but there is no serious reason to think it is.) But the *axioms* of such systems *are* fallible in a non-trivial sense. The attempt to derive all mathematics from 'self-evident', 'logical' truths, as is well known, broke down.

According to them, then, logic, at least, is foundational.

3.3   CONTEMPORARY SET THEORY

Lakatos and his editors are in agreement about the fallibility of set theory. Set theory initially appeared at the hands of Cantor and Dedekind in an informal, non-axiomatized, form. And it quickly spawned refutations, in the shape of the contradictions that are the set-theoretic paradoxes. The theory was certainly not, then, certain. After many discussions and debates, consensus settled on an axiomatization, ZF(C)—Zermelo Fraenkel set theory with the Axiom of Choice. (We will return briefly to the process by which this happened later.) But it would indeed be foolish to suppose that these axioms are themselves beyond challenge.

The Axiom of Choice was disputed almost as soon as it was formulated explicitly. True, debates over it have now died down, but it would be an act

of hubris to maintain that they could not be reignited. More recently, we have seen debates over another of the axioms, the Axiom of Foundation. This says, essentially, that all sets are in the Cumulative Hierarchy. Modern research into non-well-founded set theory shows that an appeal to non-well-founded sets (sets outside the hierarchy) makes perfectly good sense, and has fruitful applications.[3]

More fundamentally, as an axiomatization of the truths about sets, there is something shaky at the very core of ZF. ZF avoids the paradoxes by denying the existence of a totality of all sets, and other "large" sets. But mathematicians are tempted all the time to make informal use of such sets. Category theory, for example, appears to make essential use of them.[4] The impulse for theories that add other sorts of collections, proper classes, to the sets of ZF is but one recognition of this fact. How adequate such measures are, we need not go into here. It suffices for present to note that they indicate an insecurity about the conceptual structure, and so the axioms, of ZF.[5]

The axioms of set theory, then, are soft. We might note, at this point, a formalist move that might be made here. The whole project of trying to find true axioms is misguided. What mathematicians do is simply set down any set of axioms at will; they then spend their time proving theorems in the axiomatic system specified.

This is not the place to discuss this view at great length. But it is a wildly implausible one. Mathematical reasoning is often not axiomatic. If we need a reminder of this, we have only to remember Gödel's first incompleteness theorem, which shows us that given any formal system of mathematics of the appropriate kind, we appear to be able to establish things about the entities in question by, but only by, reasoning *outside* the system. And if one attempts to circumvent the first incompleteness theorem by an appeal to second order logic, then *this* itself is not axiomatisable.

But more importantly, specifying arbitrary axiom systems is not mathematics. For a start, one can write down axiom systems for parts of physics, biology, sociology, or anything else. There is nothing intrinsically *mathematical* about axiomatic systems. Indeed, an arbitrary axiomatic system is likely to be devoid of mathematical interest. If mathematicians construct axiom systems, it is because they think that the axioms capture a notion of mathematical interest and importance. They therefore have a pre-existing and informal grasp of the notion (which is not to say that the grasp cannot be improved by suitable axiomatization, amongst other things). The intuitions in question always stand as an independent check on the axioms, and, in principle, can clash with the theorems of the system at any time.

---

[3]See Aczel (1988) and, e.g., Barwise and Moss (1996).

[4]See Priest (1987), 2.3.

[5]For further discussion, see Priest (1995), ch. 11.

3.4 CONTEMPORARY LOGIC

Let us now turn to the Editors' claim that infallibility has at least grounded out in logic—classical logic, they obviously have in mind. We find it, frankly, amazing, that the they could have made this claim. Even when *Proofs and Refutations* was published, intuitionist logic was well known, as were Brouwer's claimed counter-examples to principles such as the Law of Excluded Middle. The Editors may well not have liked such examples, but as Lakatos demonstrates, many mathematicians did not like the putative counter-examples to Euler's Theorem either. Logical axioms are just as soft as set-theoretic ones.

Maybe, in the present state of play, even softer. Since the 1970s we have seen a remarkable growth in non-classical logics.[6] Many of these take off from aspects of classical logic for which there are apparent counter-examples. Logics with truth value gaps accommodate sentences that appear to be neither true nor false. Free logics accommodate singular terms that appear not to denote. Relevant logics avoid the highly counter-intuitive "paradoxes of material implication". All such counter-examples, and the logics to which they give rise, are, of course, disputable. But that is beside the point. They show that logic has no privileged exemption from fallibility.[7]

In this context, it is illuminating to consider for a while one particular kind of refutation of classical logic: the paradoxes of self-reference. Sentences like 'this sentence is false' quickly give rise to contradictions, showing, apparently, that something can be both true and false—a possibility that certainly cannot be accommodated in classical logic.

Logicians and mathematicians have often reacted to these refutations with the same kind of stratagems that Lakatos described.[8] A lot of monster-barring goes on: such sentences are monsters of some kind: they are not to be taken seriously. They do not express propositions, or they "malfunction"—whatever that might mean.[9] They may therefore be set aside. Alternatively, some logicians have simply withdrawn to a safe domain. Tarski's hierarchy can be seen in this light. As long as we consider only sentences that can be expressed in the appropriate hierarchy, contradictions cannot arise. But as subsequent work has shown, languages with their own truth predicates can express things not expressible in the Tarski hierarchy, quite consistently. The constraints imposed by the hierarchy are overly-restrictive.[10]

---

[6]See, e.g., Priest (2001).

[7]One might suggest that it is some particular inferences, rather than systems thereof, that are immune from doubt. But in fact, there is no significant principle of inference that is not problematised in some non-classical logic or other: the Law of Excluded Middle, and Double Negation (intuitionist logic) the Law of Non-Contradiction, Contraposition (some paraconsistent logics), Distribution (quantum logics), Modus Ponens (some fuzzy logics), Conjunction Elimination (connexive logics).

[8]We do not wish to suggest that all contemporary consistent approaches to the semantic paradoxes display these methodological vices.

[9]See, e.g., Smiley (1993).

[10]See, e.g., Kripke (1975).

Such stratagems are as methodologically unsatisfactory in this case as they were in the one that Lakatos discussed. If we take the self-referential examples seriously, what they appear to show is that the Law of Non-Contradiction (LNC) is not acceptable. Some sentences of the form $A \wedge \neg A$ are indeed true. The Law might be something we are accustomed to take for granted, but the whole function of refutations is precisely to cause us to question things the truth of which we have taken for granted.

Actually, this is not quite right. One can have the LNC, at least in the form $\neg(A \wedge \neg A)$, as well as a contradiction. True, we then get a secondary contradiction, $(A \wedge \neg A) \wedge \neg(A \wedge \neg A)$, but that is no worse than the original contradiction. What one cannot have is the law of Explosion, $A \wedge \neg A \vdash B$. This is what the paradoxes of self-reference really refute—and Explosion, incidentally, is a lot softer than the Law of Non-Contradiction. Historically, its place in logic has always been less than secure.[11] Partly fuelled by the paradoxes of self-reference, we have witnessed in the lasts 40 years the development of paraconsistent logics, logics where Explosion fails.

But let us take matters a little further. Explosion is not something that is likely to be accepted simply because it recommends itself. (Aristotle, for example, rejected it, even though he subscribed to the LNC.) It is usually accepted because it follows from other things. Perhaps the most famous argument for Explosion—probably invented in about the 12th century, and certainly known to medieval logicians, is as follows:

$$
\cfrac{\cfrac{A \wedge \neg A}{A} \qquad \cfrac{\cfrac{A \wedge \neg A}{\neg A}}{\neg A \vee B}}{B}
$$

The paradoxes of self-reference, taken seriously, plus a B that is not true, provide a global counter-example to this argument. Where to point the blame? Most paraconsistent logicians finger the final step, the Disjunctive Syllogism, $A, \neg A \vee B \vdash B$. The failure of this is certainly more surprising than that of Explosion. The inference would appear to be applied in naive reasoning much more frequently, for example. Once viewed through the eyes of the paradoxes, though, it is clearly moot. The rule gets its appeal from the thought that A rules out $\neg A$ (that is, you can't accept both A and $\neg A$); so given $\neg A \vee B$, we must have B. But given the possibility of paradoxical sentences, A precisely does not rule out $\neg A$. So one would expect the principle to fail. Its tenure in thinking is the result, perhaps, of simply being unaware of, or forgetting, or not taking seriously, unusual cases. Surprising the failure may be. But the whole point of proof analysis in the light of a counter-examples is trace back blame, and so expose the illegitimate assumptions which are being taken for granted. We should expect to be surprised.

---

[11]See Priest (2007), section 2.

In any case, Lakatos got it right, and his Editors got it wrong: modern logic is no more infallible than any (other?) part of mathematics.

3.5 SET THEORY REVISITED

Before we conclude, let us return briefly to set theory. ZF set theory arose from informal set theory, partly in response to the set-theoretic paradoxes of self-reference. Its formulation was, it might well be thought, somewhat opportunistic. Zermelo and his co-workers wrote down a bunch of axioms that seemed to suffice for proving the things they wanted to prove and which did not allow the proof of contradiction. But the axiom system came later to appear more principled. The standard model of ZF is the cumulative hierarchy. This is what the axioms are most naturally taken to characterise. From this point of view, the situation appears to be a version of withdrawal to a safe domain. Set theoretic contradictions do not arise, it is true, but that is because we have withdrawn to a simple safe area. We have overwithdrawn, as the appearance of non-well-founded set theory suffices to remind us.

Indeed, in a paraconsistent context, even this withdrawal may be too severe. Cantor, who was well acquainted with the paradoxes of set theory, distinguished between consistent and inconsistent totalities.[12] Modern set theory has withdrawn entirely to the realm of the consistent totalities, leaving all the others outside. Of course, what is outside makes little sense if one subscribes to Explosion. The inconsistent collapses into the formless. But Cantor's logic is not that of Frege and Russell, and there is no reason to suppose that he subscribed to Explosion (as far as we know, anyway[13]). Certainly, once one takes paraconsistency seriously, one can recognise transconsistent set theory, just as much as transfinite.

4 CONCLUSION

Let us bring this discussion to a conclusion with one further point. As we saw in the quote from Lakatos that we gave, the method of proof analysis gives rise to the growth of mathematics. His case study in *Proofs and Refutations* shows how this may be so. The rise of paraconsistent logic demonstrates yet another twist of the "cunning of reason". As we have just observed, paraconsistent logic allows for the recognition of a whole new mathematical realm, ripe for mathematical investigation, transconsistent set theory. More generally, the consistent is a special case of the inconsistent. In the semantics of all paraconsistent logics there is a space of interpretations. And standardly, a subspace of this is constituted by the classical interpretations. These interpretations may

---

[12]Cantor, (1899). For a discussion, see Priest (1995), 8.6.

[13]What Cantor, who, like most mathematicians, reasoned informally, and who appeared to have no interest in formal logic, would have said about Explosion, or the arguments for it, one will presumably never know.

be investigated as ever. Paraconsistent logic does not abolish the investigation of consistent structures. What it does is *expand* the domain of what can be investigated. It allows us to investigate inconsistent structures and theories in a way that was unthinkable before.[14] This is, we think, a remarkable instalment in the growth of mathematics.

At any rate, and to return to the main theme of this paper. The Euclidean conception of proof cannot characterise the history of mathematics. Lakatos' conception of proof as a fallible enterprise, starting from things that appear to be true, but which are subject to revision in the light of counter-examples, appears much more plausible. Developments in the last 100 years have not changed this picture essentially. Mathematicians and logicians are undoubtedly much more self-conscious about formulating the starting points, their axioms. But the axioms are no infallible epistemological bedrock. They are merely places where proof may stop, *pro tem*; they are still liable to be challenged by appropriate counter-examples.[15] And this is just as true of the axioms of logic as those of mathematics. The development of paraconsistent logic can be seen as a clear case of this.[16]

REFERENCES

[1] Aczel, P. (1988), *Non-Well-Founded Sets* (CLSI Lecture Notes, Number 14), Stanford, CA: CSLI.

[2] Barwise, J., and Moss, L. (1996), *Vicious Circles* (CLSI Lecture Notes, Number 60), Stanford, CA: CSLI.

[3] Cantor. G. (1899), 'Letter to Dedekind', 113-17 of J. van Heijenoort (ed.), *From Frege to Gödel*, Cambridge, MA: Harvard University Press, 1967.

[4] Kripke, S. (1975), 'Outline of a Theory of Truth', *Journal of Philosophy* 72, 690-716.

[5] Lakatos, I. (1976), *Proofs and Refutations*, Cambridge: Cambridge University Press.

[6] Mortensen, C. (1995), *Inconsistent Mathematics*, Dordrecht: Kluwer Academic Publishers.

---

[14]A start in this direction is made in Mortensen (1995). The most developed investigations of inconsistent structures are presently those that concern arithmetic. For an overview of the current state of affairs in this area, see the second edition of Priest (1987), ch. 17. The recapture of classical mathematics is further discussed in 8.5 of the same book.

[15]We might think of the distinction between working within an axiom system and criticising the system in terms of Kuhn's distinction between normal science and revolutionary science.

[16]Versions of this paper were given at the Universities of South Africa, Melbourne, St Andrews, Cambridge, and Manchester. Thanks go to all those in the audiences for their thoughts and comments.

[7] Popper, K. (1934), *The Logic of Scientific Discovery*. First English edition (London: Hutchinson, 1959).

[8] Priest, G. (1987), *In Contradiction*, Dordrecht: Martinus Nijhoff; second (extended) edition, Oxford: Oxford University Press, 2006.

[9] Priest, G. (1995), *Beyond the Limits of Thought*, Cambridge: Cambridge University Press; second (extended) edition, Oxford: Oxford University Press, 2002.

[10] Priest, G. (2001), *Introduction to Non-Classical Logic*, Cambridge: Cambridge University Press.

[11] Priest, G. (2007), 'Paraconsistent Logic: Dialetheic Variations', 129-204, vol. 8 of D. Gabbay and D. Woods (eds.), *Handbook of the History of Logic*, Amsterdam: Elsevier.

[12] Smiley, T. (1993), 'Can Contradictions be True?', *Proceedings of the Aristotelian Society, Supplementary Volume* 67, 17-34.