Motivations for Paraconsistency: The Slippery Slope from Classical Logic to Dialetheism

Graham Priest

> *'Indeed, even at this stage, I predict a time when there will be mathematical investigations of calculi containing contradictions, and people will actually be proud of having emancipated themselves from consistency.'* Ludwig Wittgenstein, 1930.

Fellow Congress participants:[1] you have already been welcomed by the Vice-Rector of the University of Ghent and by the Chair of the Congress, Prof. Batens. Let me add my voice to theirs.

This is an historic occasion. Where and when the modern paraconsistent movement started is, to a certain extent, an arbitrary matter; but I think that a reasonable date is the publication of Jaśkowski's paper 'Propositional Calculus for Contradictory Deductive Systems' in 1948. By that reckoning, the movement will be coming up to its 50th birthday next year. This is the first time in those 50 years that there has been a major international conference on the subject. (There have been some national ones.) Unfortunately, some of the influential figures in the movement are no longer with us, Jaśkowski himself, Anderson, Arruda, Smirnov, Sylvan. Sylvan was an invited speaker for this conference and died suddenly last year. This conference will be the poorer for his absence. But many of those who have made significant contributions to the subject are here today: Alves, Asenjo, Batens, Brady, Dunn, Jennings, Meyer, Mortensen, Peña. May I be pardoned by those whom I have not mentioned. (Da Costa had to pull out at the last moment, unfortunately, but his paper is being read for him.) There may well be conferences on paraconsistency again—I hope there will be; but it is unlikely that so many of the founders of the subject will be gathered together again. Further, the subject is

[1]This article is the text of the opening address of the World Congress on Paraconsistency, Ghent, 1997, pretty much as it was given. I decided to leave it in this form to convey something of the spirit of the occasion. For further discussion of all the material covered in sections 1–3, see G. Priest, 'Paraconsistent Logic', in D. Gabbay and F. Guenthner (eds.) *Handbook of Philosophical Logic*, 2nd. ed. (Kluwer Academic Publishers, forthcoming). The quote from Wittgenstein is from p. 332 of *Philosophical Remarks* (Blackwell, 1964). I'd like to thank Diderik Batens for suggesting that a lecture of this kind be given, and also for helpful comments on a draft of it.

now growing at such a rate, and making connections with so many other areas, that conferences that try to cover *all* of paraconsistency, as this one does, are likely to be too unwieldy to attempt.

So whether you are a paraconsistent logician, not a paraconsistent logician, both a paraconsistent logician *and* not a paraconsistent logician, or neither, I hope that you enjoy participating in this unique occasion, and return home safely and intellectually stimulated.

1 Introduction

What I aim to do in this introductory lecture is to put paraconsistency in a certain perspective. In particular, I will try to disentangle some of the motivations that have driven people to paraconsistency, and chart the connections between them.

Let us start with a definition of paraconsistency. I will call the principle of inference that everything follows from a contradiction (α, $\neg\alpha \vdash \beta$, for all α and β), *Explosion.* (Its more usual Latin name is *ex contradictione quodlibet.*) A logic is *explosive* if it verifies explosion and *paraconsistent* otherwise.

The definition is not perfect. There are certainly logics that are paraconsistent with respect to the letter of this law, though not with respect to its spirit. For example, minimal logic is paraconsistent with respect to this definition, but in this logic contradictions entail the *negation* of everything ($\alpha, \neg\alpha \vdash \neg\beta$, for all α and β). Still, the spirit of paraconsistency is difficult to pin down exactly. (In many ways, it is what the major part of this talk is about.) And the above definition is at least simple, and now widely accepted.

Incidentally, the term 'paraconsistent' was coined by Miró Quesada at the Third Latin American Conference on Mathematical Logic in 1976. In English the prefix 'para-' has two somewhat distinct meanings: 'quasi-' or 'similar to, modelled on', as in 'para-military' or 'para-medic'; or 'beyond', as in 'paradox' (beyond belief). Newton da Costa tells me that it was the first meaning that Quesada had in mind. I have always preferred the second. I prefer to think of paraconsistent logicians, not as trying to be like consistent logicians—and not quite getting there—but as going beyond what classical logicians can do, exploring regions where they cannot venture.

Anyway, whatever the word means, I want to talk about the nature of the subject. Let me start by putting the matter into perspective, by saying a little about its history.

2 The History of Explosion

Explosion is a principle that strikes most who begin to study logic as outrageous. It is, however, a very orthodox principle of modern logic. It is sometimes thought that it is a very orthodox principle of logic, period. This is, in fact, just false: the entrenchment of explosion is a very modern phenomenon.

The first formal logic was Aristotelian Syllogistic; and this is paraconsistent. To see this, merely consider the inference:

Some men are mortal.

No mortals are men.

Hence, no men are men.

This is not a valid Syllogistic form, though its premises are certainly inconsistent. In fact, at *Prior Analytics* 64^a15, Aristotle tells us quite explicitly that some contradictions entail syllogistic conclusions and some don't. Not everything, therefore, follows from a contradiction.

The other great logical system of antiquity was Stoic logic. Unlike syllogistic, this was a propositional logic in much the modern sense, but it is notable that explosion does not appear in any of what remains to us of the writings of Stoic logicians. Moreover, it is such a striking principle that it seems unlikely that (critical) commentators such as Sextus Empiricus would not have commented on it had it been endorsed in Stoic logic. It is therefore reasonable to believe that Stoic logic was not explosive.

If explosion is not to be found in ancient logic, where does it come from? The earliest appearance in the history of logic of which I am aware, seems to be in the 12th century Paris logician William of Soissons. At any rate, William was a member of a group of logicians called the 'Parvipontinians' who made a name for themselves defending explosion. After this time, the principle is a contentious one in medieval logic, being endorsed by some logicians, such as Scotus, and rejected by others, such as the Cologne School.

The entrenchment of explosion begins only in the third of the great historical periods of logic, the contemporary period. As a result of the work of Boole, Frege and others in the second half of the 19th century, a new logical theory, now misleadingly called 'classical logic', was produced. The new logic was so much more powerful than traditional logic that it was soon entrenched in the logical community. And since classical logic was explosive, explosion was thereby entrenched.

We now know, of course, that much of the power of classical logic stems not so much from its content, as from the mathematical techniques that it deploys (proof-theory, model-theory), and that these techniques can also be deployed to produce very different accounts of logical consequence. Indeed, in the first half of this century, other such accounts were constructed, the most famous of these being intuitionistic logic. But all the systems constructed were explosive, until the rise of the modern paraconsistent movement, to which we can now turn.

3 The Modern Paraconsistent Movement

The modern movement certainly had its precursors before 1948. Paraconsistent logics (of very different kinds) were proposed by the Russian logicians Vasil'ev (*c.*1910) and Orlov (1929), but these were voices crying in the wilderness. And the modern movement was strikingly foreseen as early as 1930 by Wittgenstein, as shown by the quotation at the start. However, it was to begin in earnest only after the Second World War.

Paraconsistency was then an idea whose time had come. In the space of about 20 years it seems to have occurred independently to a number of different logicians in very different countries—and reinvented by many others since then—Jaśkowski in Poland (1948), Asenjo in Argentina (1954), da Costa in Brazil (1963), Anderson and Belnap in the USA (early 1960s). Since this time, the ideas have developed rapidly in many different places, but I think it fair to say that the epicentres of the movement have been Brazil and Australia.

The paraconsistent logicians I have just mentioned were often driven by very different considerations. More of this in a moment. They also came up with rather different *kinds* of paraconsistent logic. Let me illustrate just three of these. (They are by no means exhaustive.) If one defines a paraconsistent logic model-theoretically, to invalidate explosion one needs a notion of propositional interpretation according to which it is possible for α and $\neg\alpha$ to hold simultaneously. One way to achieve this (in effect, Jaśkowski's) is to take an interpretation to be a Kripke model, M, for a modal logic, say $S5$, and to define 'α holds in M' to mean that α is true in *some* world in M. According to this approach, although α and $\neg\alpha$ may hold in M, $\alpha \wedge \neg\alpha$ never does. Hence the rule of adjunction ($\alpha, \beta \vdash \alpha \wedge \beta$) fails.

According to another approach (one of those invented by da Costa, with Alves), an interpretation is the same as in the classical propositional calculus, except that negation is not truth functional. Thus, the truth value of α does not determine the truth value of $\neg\alpha$. In particular, both may take the value 1. A feature of this approach is that all of the positive (negation-free) part of the classical propositional calculus is preserved.

A third approach is many-valued (this was the approach of Asenjo). Take three truth values, *true*, *both*, and *false*, whose logical relationships are depicted in the following Hasse lattice diagram:

$$\begin{array}{c} true \\ \uparrow \\ both \\ \uparrow \\ false \end{array}$$

Conjunction is the minimum, disjunction is the maximum, and negation inverts the order. In particular, *both* is a fixed point for negation. If we take the designated values to be $\{true, both\}$, then we obtain a paraconsistent logic. Throw in a fourth truth value, *neither*, so that the Hasse diagram now looks like this:

$$\begin{tikzcd} & true & \\ both \arrow[ur] & & neither \arrow[ul] \\ & false \arrow[ul] \arrow[ur] & \end{tikzcd}$$

where *neither* is also a fixed point for negation, but where the designated values are the same, and one has essentially Dunn's semantics for the relevant logic First Degree Entailment.

4 Three Levels of Paraconsistency

Let us now return to the topic that I raised earlier, of the motivations for paraconsistency; this will occupy us for the rest of the talk.

Paraconsistent logicians (those I have already mentioned and others) have been driven by a number of different motivations. Amongst these motivations, there are at least three quite distinguishable ones—though doubtlessly they shade into each other at the edges. I will call these three *levels of paraconsistency*. I should say straight away that the levels do not match up in any neat way with the various approaches just outlined. To a large extent, one might try to proceed at any level, by applying any of the approaches mentioned—though whether one can always do this successfully is another matter. I will describe the levels in increasing order of "strength".

The first level of paraconsistency (gentle-strength paraconsistency) is a simple dissatisfaction with explosion as a valid principle of inference (maybe for certain purposes). As I have already observed, explosion is historically moot, and quite counter-intuitive to those studying logic for the first time. It is therefore unsurprising that people should have been moved to construct systems of logic in which it fails. This was the motivation of a number of the pioneers of paraconsistent logic, notably American relevance logicians, such as Anderson and Belnap.

Though the interest of the first level of paraconsistency is with the inference relation itself, those involved at this level have not been blind to the applications of paraconsistency. The one that has received most attention here concerns information processing. There is no way that one can prevent data of any degree of sophistication, as represented in a computational database, from being inconsistent; and if inferences are to be drawn from it, a paraconsistent logic is clearly desirable.

The second level of paraconsistency (full-strength paraconsistency) is the idea that there are interesting and/or important inconsistent but non-trivial theories, which should therefore be investigated. It seems clear that such theories occur naturally in the history of science and mathematics: the original theory of the calculus, Bohr's theory of the atom, naive set theory. In each case we have a challenge—and often a very difficult one: to formalise the theory with an underlying paraconsistent logic in such a way as to do justice to its integrity. Friends of consistency can hardly deny the existence of inconsistent theories in the history of science. What they are likely to claim is that the deductions of arbitrary conclusions from the inconsistencies were blocked by some, maybe *ad hoc,* mechanisms. But this is just to say that the theories did have a *de facto* underlying paraconsistent inference mechanism. Why, then, not take this to be *de jure* if it can be suitably analysed?

Alternatively, we may want to construct new theories that enshrine natural but inconsistency-producing principles. Thus, for example, we may construct a theory of truth which incorporates the T-schema, giving rise to the Liar, and other semantic paradoxes. Alternatively again, in the semantics of paraconsistent logics, inconsistent theories have inconsistent models—such as the intriguing inconsistent models of arithmetic. These may have mathematical structure that cries out to be investigated.

The second level of paraconsistency has motivated many paraconsistent logicians; the first, arguably, being da Costa. And much of his and other people's work has been on the investigation of inconsistent mathematical theories and their models.

The third level of paraconsistency (industrial-strength paraconsistency) is the thought that some of these inconsistent but non-trivial theories may be *true*. This is a view now called 'dialetheism' (a phrase coined by Sylvan and myself in 1982). This third level is particularly characteristic of the Australian paraconsistent movement. The most natural candidates for the status of true inconsistent theory are certain theories of sethood and truth, which enshrine the paradoxes of self-reference. There are, however, other candidates. For example, theories of change—and even good old fashioned dialectics.

These, then, are the three levels of paraconsistency. It could be argued that it is the second level that marks the real beginning of paraconsistency, and that the first level is really a different enterprise. There is some plausibility to this, but I think it better to draw the paraconsistent net wider. The reason for this is simply that so many of the issues that arise in the second and third levels of paraconsistency are also integral to the first. As Sylvan often emphasised, what theories a consequence relation makes possible is an important piece of the evidence concerning the adequacy of that relation. And certainly, what sorts of logical forms may be true is highly relevant to the question of a correct inference relation.

5 ...and the Relationship Between Them

Let us now turn to the relationship between the three levels of paraconsistency. For a start, they are quite distinct; one can subscribe to any weaker form without subscribing to a stronger form—though not conversely.

Thus, one may take it that the correct inference relation (maybe with respect to a given purpose) is paraconsistent (first level) without holding that any of the inconsistent theories this makes possible are of any interest (second level). One might hold, for some reason, that any inconsistent theory is simply beyond the pale. Of course, the set of claims that follows from an inconsistent database is logically closed, and so is a theory in the logicians' sense. But it is not a theory in the more usual sense of that word, as in physics or mathematics (Newton's theory of gravity, category theory). Such structures are a lot more than deductively closed sets of sentences.

But if one thinks that there are interesting and important inconsistent non-trivial theories (second level) then one obviously has to take their consequence relation to be paraconsistent (first level)—though one might feel more freedom in letting the theory constrain, and so partially determine, the inference relation if one approaches things from level two; approaching from level one, it is more natural to get the inference relation right in the first place, and then see what theories fall out of it.

In a similar way, one may subscribe to the second level of paraconsistency without subscribing to the third. One may take certain inconsistent theories to be intrinsically interesting, or to have important mathematical structure, without supposing that those theories are candidates for the truth. Even if the theories are applied ones, one may take it that they are important because they are useful instrumentally, or good approximations to the truth; one still does not have to hold that they may be true. Or one may take the models of inconsistent theories to represent impossible situations, of the kind described in some fictions, or by the antecedents of counterfactuals of certain kinds (such as: 'If you were to square the circle, I would give you all my money'). There is no commitment here to the possibility of truth in these cases.

The third level does require the second, however. If some inconsistent theories (about, e.g., sets or change) are true, then undeniably there are interesting and important such theories. Naturally, one does not have to suppose that all interesting and important inconsistent theories may be true.

6 The Slippery Slope

Despite the differences between the three levels of paraconsistency, there is a very natural progression down them, forced by the weight of (rational) gravity. Suppose that one subscribes to the first. Then, like it or not, there are going to be inconsistent and non-trivial theories (in the logicians' sense). Now it could just be that none of these has any intrinsic interest, applicability or whatever; but this seems rather unlikely. Indeed, we know that it is not the case. The inconsistent theories of arithmetic and their models have an elegant mathematical structure, interesting in its own right. Moreover, inconsistent theories may obviously be useful; witness the Bohr theory of the atom. Hence the first level of paraconsistency gives rise, almost inevitably, to the second (though, naturally, any given individual may not be interested in pursuing things further than the first level).

The slide from the second level to the third is also a very natural one. If there are structurally rich, and even empirically applicable, theories that are inconsistent, why *shouldn't* some of these be true? In the semantics of paraconsistent logics, contradictions can be true in an interpretation, so why not true *simpliciter*? Let me elaborate on the second question. A model may naturally be thought of as representing a possible situation—at least, a situation possible in as far as the meanings of the logical constants goes. And if this is right, then the fact that there are inconsistent models means that it is logically possible for contradictions to be true. Now, one might argue that some models, and especially the inconsistent ones, are not candidates for representing *real* situations—whatever that may mean; but the onus is now on someone who rejects the third level of paraconsistency to say why.

What can one say? There is really only one move here: to invoke the Law of Non-Contradiction (LNC), in the form: no contradiction may be true. It is absolutely crucial to distinguish between this and paraconsistency. Aristotle, for example, subscribed to the LNC, though his logic was paraconsistent, as we have seen. Moreover, unlike explosion, the LNC is deeply entrenched in Western philosophy (though, arguably, not in Eastern philosophy). It is true that some thinkers have challenged the law (some Presocratics—such as Heraclitus; some Neoplatonists—such as Cusanus and, maybe, Plotinus; and some dialecticians—such as Hegel and Engels). But the weight of historical consensus has sided firmly with it.

It is important to ask, though, what *reasons* are there for accepting the LNC. Do we have any reason to suppose that the adherence to the Law is anything more than a sociological fact? After all, nearly everything else that Aristotle endorsed has been overthrown! The major historical defence of the Law is to be found in Aristotle himself (*Metaphysics*, Γ4). How good are Aristotle's arguments? Terrible. The main long argument, which takes up half the chapter, is convoluted and opaque. It is not even clear *how* it is supposed to work, let alone *that* it works. The other six or seven (it depends how you cut the cake) are usually little more than throw-away arguments. It is not clear that any of them carry much weight. Worse: most of them argue, not for the LNC, but for the much weaker claim that it is impossible for *all* contradictions to be true (or even that it is impossible to *believe* all contradictions to be true). This is, naturally, quite compatible with the third level of paraconsistency. All this was pointed out long ago in 1910 by Łukasiewicz (who almost certainly influenced Jaśkowski, since he was his teacher). Aristotle is therefore of no help here.

What others arguments are there? It is a rather amazing fact that there has never been a sustained defence of the LNC since Aristotle's. Western philosophy seems to have taken the LNC for granted—rather unwisely, given its slim basis. This does not show that there are no other arguments, of course. By far the most powerful argument appeals to explosion. If this principle is valid, and given that not all contradictions are true, it follows that none are. But this argument is precisely not available to someone who is at the second (or even first) level of paraconsistency. There are a few other arguments for the LNC that I am aware of, but none withstands much scrutiny. This is not the place to go into details, so let me just say, for the record, that such arguments have a tendency simply to beg the question at some crucial place.

To summarise: unless there are good arguments for the LNC that are acceptable to someone at the second level of paraconsistent involvement—and I know of none—then the movement from the second level to the third beckons ineluctably. One might wait to be convinced by an appropriate inconsistent candidate for truth (though for my money there are already such around), but at least in principle, the correctness of the third degree of paraconsistent involvement must be accepted.

7 The Classical Backlash

I have now said all I wish to say about the modern paraconsistent movement here. Let me finish by saying a little about the orthodoxy against which it has developed. It is not unusual for new developments in science to be carried out in the face of opposition from orthodoxy. The paraconsistent movement is certainly a case of this. Though the situation is now changing—in some places—for a long time, paraconsistent logics were widely regarded as too outrageous to be given serious attention. They were usually ignored in public, and sometimes even ridiculed in private. I think that much of this must be put down to rank conservatism.

The crucial question is, again, whether there are any *principled* reasons for rejecting paraconsistency—of any level. In virtue of the slippery slope down the levels, my advice to anyone who wishes to remain true to the classical faith would be: *fight even the first level of involvement.* Yet this is a hard counsel. The reason why is simple. The first level of paraconsistent involvement is usually depicted as a loss: there are forms of inference that are classically valid, but that can no longer be used, such as the Disjunctive Syllogism ($\alpha, \neg\alpha \vee \beta \vdash \beta$). In fact, this is exactly the opposite of the truth. Going paraconsistent is a *gain.* In all the mechanisms for paraconsistency that I mentioned, classical interpretations are a special case (where, in a Jaśkowski interpretation, there is only one world; where, in a da Costa interpretation, negation is *de facto* truth functional; where, in a many-valued logic, nothing takes a non-classical value). Thus, a paraconsistent logician may analyse every situation that a classical logician may analyse in exactly the same way. But they can also analyse *more.* Paraconsistent logic therefore has excess power. This can all be made quite precise with the non-monotonic (adaptive) logics pioneered by Batens. These are paraconsistent logics that give precisely classical logic when reasoning from consistent information (or about consistent situations).

Paraconsistent logic is therefore preferable to classical logic for exactly the same reason that classical logic was preferable to traditional logic. It is a more powerful and flexible inference engine. Why prevent yourself from exploring what is beyond the consistent, when you lose nothing by this? Why shackle yourself with the *necessity* of consistency, when you don't have to? (With apologies to Marx and Engels:—) Logicians of the world unite! You have nothing to lose but your chains.

University of Queensland, Australia.