Introduction

This Element is an introduction to some uses of paraconsistent logic for mathematics. It is for beginners – or at least, readers who know enough about what the words in the title mean to have picked this up, but not much more. In writing it, I have mostly tried to take a neutral "tour guide" approach, both in the selection of material and in avoiding trying to "sell" the reader anything, though my biases have inevitably shown through, especially by the end. The views expressed in this Element are the author's and do not necessarily reflect the position of Paraconsistency Inc or its affiliates.

Each of the first four little sections exposits some key ideas (including, inevitably, using some formal symbolism); the last section of each discusses a more general philosophical issue that arises. The final section is a brief philosophical and critical appraisal, looking to the future of this little field.

1 Invitation to Paraconsistency in Mathematics: Why and How?

Is mathematics *consistent*? Must it be? Or could new mathematical discoveries be found where previously no one had thought possible to look – in the *inconsistent*?

Paraconsistency in mathematics allows the development of mathematics that either is or could be inconsistent but without absurdity. On a mainstream approach in logic, any inconsistent theory is as good (or bad) as another, because all contradictions are equivalent. Standard mathematics uses *classical logic*, according to which there is only one inconsistent theory – the one that contains every single sentence, the absurd or *trivial* theory. With classical logic, if there is even a single contradiction, everything collapses. In a slogan, the conventional view is that "*the inconsistent has no structure*" (Mortensen, 2010, p. 3).

A nonclassical, paraconsistent conception of logic takes a more fine-grained approach to contradictions. Paraconsistency distinguishes between a theory being inconsistent (it includes at least one contradiction) from the notion of being incoherent, absurd, or trivial (it includes every sentence); from a paraconsistent point of view, a theory does not have to be consistent to be coherent. Paraconsistency in mathematics thus provides a rigorous framework for both a *cautious* approach to contradiction – for all we know, a theory might be inconsistent, so we use a logic that can handle it – and a much *less* cautious project: of investigating and describing contradictory abstract objects and structures.

Until recently, even the possibility of a contradictory, yet coherent, mathematical theory would have been taken as facially absurd. (In some quarters, of course, it still is.) Developments in formal logic since the 1950s, though, have established that such theories can indeed exist. "Paraconsistent logic" now has an official mathematics subject classification code.[1]

One aim of this emerging, diverse field of work is to widen the horizons of mathematics by discovering and studying new objects – much in the way that historically mathematics has advanced by admitting the existence of "bad" entities like zero, negative numbers, irrational numbers, imaginary numbers, transfinite sets, geometries where parallel lines can meet, and so forth. Newton da Costa suggests that "it would be as interesting to study the inconsistent systems as, for instance, the non-euclidean geometries (da Costa, 1974, p. 498)." For paraconsistency in mathematics, as Robert Meyer puts it, "what is to be hoped for most of all are not new routes to old truths, but an expansion of the pragmatic imagination (Meyer, 2021a, p. 158)."

The aim of this Element is to give the interested reader a critical sense of some of the work in this area to date, its strengths and weaknesses, and to indicate what might be next.

1.1 Motivations

Let us begin with an example of possible inconsistency in mathematics that motivates using paraconsistent logic. The discussion proceeds informally. Then we will get into a few details of how the logic itself works.

Mathematics is, by standard accounts developed in the twentieth century, based on set theory – itself a mathematical theory of collections that provides a foundation for all other areas of mathematics. Even very cautious philosophers like Quine have grudgingly accepted that sets are indispensable for mathematical (and so scientific) practice. Mathematics takes place in the universe of sets.[2]

Consider, then, the universe of sets – the collection of all sets, $\mathcal{U}$. This is, one imagines, a very big collection, the most inclusive collection of sets there could be, one containing *every* set. This collection is, intuitively, the domain of discourse for statements that set theorists are interested in, such as "every set can be wellordered" or "there are no self-membered sets"; and whether or not those statements are true, they do clearly seem to be – have been taken to be – meaningful. The universal collection is the prima facie basis of the meaning

[1] Under the Mathematics Subject Classification 2010, database of the American Mathematical Society, 03B53: "Logics admitting inconsistency (paraconsistent logics, discussive logics, etc)."

[2] For a good account, see Potter (2004).

of these statements: when you study set theory, $\mathcal{U}$ is "where" you work. It has been traditional at least since Whitehead and Russell's *Principia Mathematica* in 1910 to define the universal class with a "property posessed by everything," for example, the collection of all sets x such that $x = x$, or

$$\mathcal{U} = \{x : x = x\},$$

which is universal because everything is, after all, self-identical.[3] Every set x is a member of this collection, $x \in \mathcal{U}$.

Within $\mathcal{U}$ are subcollections of all sorts – the continuous functions, the commutative groups, the set of all collections with exactly three members, and so forth. One such collection is all the *singletons*. Every x has a singleton, $\{x\}$, and the collection of all singletons comprises the sets with just *one* member. Now, that means every x in the universe, every $x \in \mathcal{U}$, can be paired up *exactly*, into a one-to-one correspondence, with its singleton, in pairs

$$\langle x, \{x\} \rangle.$$

This shows a natural sense in which the universe and its subcollection of all singletons have exactly the same "size": they can be paired off perfectly. But of course, there are *more* objects in the universe than just the singletons, since most things are not singletons. So there is also a natural sense in which $\mathcal{U}$ is *not* the same size as the subcollection of all singletons.

This is a little puzzling, but we are talking about the entire universe after all, so we should be prepared for some surprises – and indeed, the outstanding mathematician Richard Dedekind used exactly this fact in 1888 to *define* what it means for a set to be *infinite*; namely, having a proper part of the same size. He then used a variant on the aforementioned argument to prove that infinite sets exist.[4]

Now consider *all* the subcollections of the universe collected together, $\mathcal{P}(\mathcal{U})$ (called the *powerset* of $\mathcal{U}$). Since $\mathcal{U}$ is maximally inclusive, both $\mathcal{P}(\mathcal{U})$ and all its members are *inside* of $\mathcal{U}$. Writing subsethood as "⊆," then

$$\mathcal{P}(\mathcal{U}) \subseteq \mathcal{U}. \tag{1}$$

---

3 See Whitehead and Russell (1910, p. 216). Here, we are focusing on the universe of sets, so "everything is self-identical" is short for "every *set* is self-identical."

4 In the infamous Theorem 66 of his *Was sind und was sollen die Zahlen?* of 1888 (reprinted in (Dedekind, 1901, p. 64)), he argues that the set of his thoughts is infinite because for each thought x there is also the thought of that thought $\{x\}$, the thought of the thought of that thought $\{\{x\}\}$, and so forth. See Priest (2006, p. 33, sec. 10.1). For alternative ways to think about the sizes of infinite sets, see Mancosu (2009).

But similarly, each member of $\mathcal{U}$ is a set, which in turn has only sets as members; so for any $x \in \mathcal{U}$, if $z \in x$ then $z \in \mathcal{U}$, which is to say that x is also a *subset* of $\mathcal{U}$ and hence a member of the powerset of the universe, showing

$$\mathcal{U} \subseteq \mathcal{P}(\mathcal{U}) \tag{2}$$

But, if all the members of $\mathcal{U}$ are members of $\mathcal{P}(\mathcal{U})$ and vice versa, then these collections are exactly the *same* collections; by (1) and (2),

$$\mathcal{U} = \mathcal{P}(\mathcal{U}) \tag{3}$$

because sets with all and only the same members are the same set, by the principle of *extensionality*.

A powerset contains all possible recombinations of elements of a set, so this equality is a little odd – but also natural enough, maybe, since the set in question is the universe. Note that, since $\mathcal{U}$ is a subset of itself, that is, $\mathcal{U} \in \mathcal{P}(\mathcal{U})$, then this means $\mathcal{U} \in \mathcal{U}$. The universe is contained in itself; the universe is *everything*, after all.

But now we have a real problem. By (3), $\mathcal{U}$ and $\mathcal{P}(\mathcal{U})$ must be the same size (since they are the same set). If they are the same size, there is a way to pair off their members in a one-to-one correspondence. Call such a pairing f, that matches members $x \in \mathcal{U}$ with members $y \in \mathcal{P}(\mathcal{U})$, as in $f(x) = y$. For members x of $\mathcal{U}$ paired up with subsets y of $\mathcal{U}$, sometimes x will be in that subset, and sometimes not.[5] So consider

$$r = \{x : x \notin f(x)\}$$

This is a subset of the universe comprising all the things that are not in the set they are paired with. But then, since f pairs off everything, and $r \in \mathcal{P}(\mathcal{U})$, there must be some $x \in \mathcal{U}$ such that $f(x) = r$. Now we just have to ask: is $x \in r$, or not? If it is, then $x \notin f(x)$ by definition, so $x \notin r$ after all; yet if $x \notin r$, then $x \in f(x)$ again by definition, so $x \in r$ after all. Since x is either in r or it is not, it is both: contradiction.

Classically, this general argument has been taken as a reductio – requiring the *rejection* of some assumption, usually the existence of a pairing off between a set and its powerset, or a universal set, or both (see §2.3). But the existence of a set theoretic universe is extremely hard to shake; and since we independently established that $\mathcal{U} = \mathcal{P}(\mathcal{U})$, it looks like there must be at least one such pairing

[5] For instance, if $a = \{1, 2, 3\}$ and $b = \{2, 4, 6\}$ then if $f(1) = a$, we would have $1 \in f(1)$, and if $f(1) = b$ then $1 \notin f(1)$.

off, namely, *identity* $f(x) = x$. Under this identity mapping, where $f(r) = r$, then writing "iff" for "if and only if,"

$$\begin{aligned} r \in r \text{ iff } & r \in f(r) \\ \text{iff } & r \notin f(r) \\ \text{iff } & r \notin r. \end{aligned}$$

So we seem to have shown that

$$r \in r \text{ and } r \notin r$$

as an apparently natural consequence of some apparently true premises. There are no assumptions to reject, or at least, none the rejection of which would be less puzzling than the conclusion that some "diagonal" subset of the universe has contradictory properties. In brief, "naive" set theory is inconsistent.

This fact was discovered in the late nineteenth century,[6] but was labeled a paradox: an apparently sound argument for an apparently impossible result. In more than a century since then, no fully satisfactory solution to this paradox has been found, as witnessed at least by a continuing stream of dissatisfied research on the subject, as we will see. A paraconsistent approach offers a unique take on this situation: there is no problem here to solve. The paradox is simply a *proof*, and its conclusion is a *theorem*: some sets are inconsistent. Put cosmically, the universe is larger than itself.

Now, none of this *forces* us into "inconsistent mathematics": historically, this and other examples all have consistent accounts. In set theory, the official solution is that there is no universal set; the collection $\mathcal{U}$, the domain of set theory (and mathematics) itself is not a set. If that seems good enough to you, I do not plan to try to talk you out of it.[7] Crucially, to see a place for paraconsistency in mathematics, it seems like I do not *need* to talk you out of it. Indeed, as we will be canvassing, the majority of people working in paraconsistency are happy to accept standard mathematics. To move ahead with paraconsistent investigations, all we need is the suggestion that there are interesting things that seem to be inconsistent when we think about them. If there could be more to learn than is allowed by the assumption of consistency, there is a place for

6 The fact that there cannot be a one-to-one correspondence between a set and its powerset was proven by Cantor and is called *Cantor's theorem*. The fact that Cantor's theorem becomes inconsistent at the universe was known to Cantor by 1895 or so but was made especially public by Russell in 1902 and is called *Cantor's Paradox*. See van Heijenoort (1967, p. 124) and Section 2.

7 If it does not seem good enough, though, I think you might be right and suggest you see Priest (2006, ch. 2) and Weber (2021, ch. 1).

paraconsistency. Whether or not we should think of paraconsistent theories as *true* or their investigations as a *challenge* to the official consistent approaches in mathematics is not obvious and is a topic we will return to.

And foundational disciplines like set theory are not the only place some have found more to learn. We will look at further topics in arithmetic, calculus, and topology. Chris Mortensen, a key founder of the field of inconsistent mathematics, has focused on impossible *pictures*. These have philosophical motivations, but from a purely mathematical point of view, perhaps it is enough just to glimpse something new there to study: the inconsistent has some structure after all, and we can pursue structure for its own sake. We may benefit from dropping what Meyer, Routley/Sylvan, and Priest call the *universal consistency hypothesis* – the limiting supposition that there is nothing of rational interest outside of the bounds consistency[8] – and widen our horizon of investigation to points beyond.

Let us look at some precise ways one might do this.

1.2 Methods

1.2.1 Paraconsistent Logic Tutorial

A paraconsistent logic allows inconsistency without absurdity. In a non-paraconsistent logic, any inconsistent premises p, $\neg p$ will have any arbitrary q as a valid conclusion (*ex contradictione quodlibet*, or *explosion*); a logic is paraconsistent if and only if explosion is invalid. Denying explosion is *all* that is required for a logic to be paraconsistent.

More precisely, let us say that a *logic* is determined by a consequence relation $\vdash$ that relates some sentences (premises) to another (conclusion).[9] When the consequence relation

$$p_0, \ldots, p_n \vdash q$$

holds then the argument from $p_0, \ldots, p_n$ to q is *valid*; and if not, not. And then let us say that a *theory* is a set of sentences closed under logical consequence: the "starting" sentences, and all the ones that validly follow under $\vdash$. An inconsistent theory contains both some sentence p and its negation $\neg p$. And so an inconsistent theory under a non-paraconsistent logic will include every sentence, which makes the theory trivial. Thus, if an inconsistent theory is to be

8 "...that all that can be spoken of or described (non-trivially) is consistent" (Priest, Routley, & Norman, 1989, p. 4; cf. Routley and Meyer, 1976).

9 There is a lot of good (paraconsistent) work on *multiple conclusion* consequence relations; see Beall and Ripley (2018, p. 744). The focus on single conclusion here is to keep it simple.

nontrivial, it must, it seems, be embedded in a paraconsistent logic, where the principle of explosion is invalid:

$$p, \neg p \nvdash q.$$

Paraconsistent logics are the basis for the study of paraconsistent theories.

There are many strategies for making a logic paraconsistent, and within these strategies there are many – infinitely many – paraconsistent logics. For concreteness, let us look at one simple approach due to Asenjo (1966) (cf. Asenjo and Tamburino, 1975) and then to Priest (1979). This is to generalize the standard truth conditions on logical evaluations and the definition of semantic validity, opening some extra space that leaves classical conditions as a special case.

Just as with classical logic, we are given a formal language with connectives $\neg$ (negation, "not"), $\wedge$ (conjunction, "and"), and $\vee$ (disjunction, "or"), with propositional atoms $p, q, \ldots$ connected by these connectives into complex expressions $A, B, \ldots$. The material conditional $A \supset B$ is defined $\neg A \vee B$ and a biconditional $A \equiv B$ is defined as $(\neg A \vee B) \wedge (\neg B \vee A)$. The new twist is in an assignment v taking sentences of the language to (two) truth values, t and f. Sentences may be true, or false, or – now diverging from classical logic – *both*.

On this arrangement, no sentence *must* be "both" but some *can*. This is possible because v is a *relation*, rather than a *function*.[10] Relations can be multiple, as in "y is a place x has lived" can take one x to multiple values for y. While an evaluation function would treat a "true" contradiction as having a value *equal* to both true and false, $\mathsf{t} = v(p) = \mathsf{f}$, and hence[11] $\mathsf{t} = \mathsf{f}$ (which not even paraconsistentists will approve of), a relation will let t and f be *among* the values of $v(p)$, one not always ruling the other out.

Relational truth conditions for negation, conjunction, and disjunction may be spelled out in a standard-looking homophonic way:

$$v(\neg A) \text{ is at least } \mathsf{t} \text{ iff } v(A) \text{ is at least } \mathsf{f}$$
$$v(\neg A) \text{ is at least } \mathsf{f} \text{ iff } v(A) \text{ is at least } \mathsf{t}$$
$$v(A \wedge B) \text{ is at least } \mathsf{t} \text{ iff } v(A) \text{ is at least } \mathsf{t} \text{ and } v(B) \text{ is at least } \mathsf{t}$$
$$v(A \wedge B) \text{ is at least } \mathsf{f} \text{ iff } v(A) \text{ is at least } \mathsf{f} \text{ or } v(B) \text{ is at least } \mathsf{f}$$

[10] Following an idea from J. Michael Dunn in the 1960s (see Omori and Wansing, 2019b), published as Dunn (1976); cf. Priest (2008, p. 161). Both Asenjo and Graham Priest present this as a three-valued logic, where the "both" value is a third distinct status along with truth and falsity. They treat the relation v as a function, assigning each sentence exactly one of the three possible statuses. There are strong reasons why it is philosophically preferable to take the relational approach, as presented here; see Weber (2021, ch. 3).

[11] At least, assuming the transitivity of identity: $a = b, b = c \vdash a = c$. This is disputed in Priest (2014).

$$v(A \vee B) \text{ is at least } \mathtt{t} \text{ iff } v(A) \text{ is at least } \mathtt{t} \text{ or } v(B) \text{ is at least } \mathtt{t}$$

$$v(A \vee B) \text{ is at least } \mathtt{f} \text{ iff } v(A) \text{ is at least } \mathtt{f} \text{ and } v(B) \text{ is at least } \mathtt{f}$$

On this account of negation, a sentence is evaluated as true iff its negation is evaluated as false; a sentence is evaluated as false iff its negation is evaluated as true; so if a sentence is evaluated as *both* true and false, then its negation is also evaluated as both true and false.

An argument is *valid* if, whenever all the premises are *at least* true, then so is the conclusion; an argument is invalid if there is some way for the premises to be at least true but the conclusion to be *not at all* true. Heuristically, validity is *un-truth preservation backward*, from conclusion to at least some premises. If we also assume the relational valuation is exhaustive,

$$v(A) \text{ is at least } \mathtt{t} \text{ or } v(A) \text{ is at least } \mathtt{f}$$

then this is the idea behind the logic now known as LP (for "logic of paradox"). It is paraconsistent because if we just consider the sentences p and q, then it is possible for p to be assigned both true and false, while q is not assigned true. Then p and $\neg p$ are both assigned at least true, but not at all so for q, and so ex contradictione quodlibet is invalid – it has a counterexample.

This is a way of *generalizing* classical logic. If the relation v were tightened up to be a function, so that, for example, $v(\neg A) = \mathtt{t}$ iff $v(A) = \mathtt{f}$, then these conditions simply are those of classical logic. Without assuming functionality, the logic allows for *gluts*. We do assume that the relation v assigns every sentence at least value t or f, making negation exhaustive:

$$A \vee \neg A$$

is always assigned at least value t. If one wanted to, this condition could be dropped, and then the logic would allow *gaps* too; that would deliver the logic FDE, which can then serve as a base to extend to several other sorts of paraconsistent logics; see Belnap (1977). In LP there are gluts but no gaps.[12]

Nothing about the general idea of paraconsistency would seem to commit to *actual* gluts (or true contradictions), only the hypothetical that *even if* there were some true contradiction, *still* not everything would be true.[13] This opens

[12] In this Element, we will not be considering "gappy" approaches, because intuitionistic and constructive mathematics have already told us a lot about the "incomplete" (but consistent) side of things, and it seems worthwhile to try to understand the dual. Also, without the law of excluded middle (LEM), many of the standard paradoxical proofs of contradictions – those that motivate inconsistent mathematics to begin with – fall apart; one can derive theorems of the form "A iff $\neg A$" but no further. The question of whether one can still derive contradictions without the LEM has been a point of contention between Priest and Brady; see Brady and Rush (2008); Priest (2019).

[13] See Barrio and Da Re (2018).

a space, then, for different approaches to paraconsistency in mathematics, which we will look at. Moderate approaches will be based on a paraconsistent consequence relation, without any commitment to gluts, whereas so-called strong paraconsistency endorses some true contradictions, sometimes called *dialetheism*.

1.2.2 A Very Brief History of Paraconsistency in Mathematics

There are many approaches to paraconsistency one could take; all that is required for paraconsistency is that the argument from $p, \neg p$ to arbitrary q be invalid. There are almost as many approaches to paraconsistency as there are logicians who have taken them. For present purposes, here is a potted history (restricting attention only to mathematical applications, not the development of paraconsistent logic). This also serves as a forecast for the sections ahead.

Paraconsistent logics were developed in various forms in the first half of the twentieth century, most notably by Jaśkowski in 1948. The first paraconsistent approaches connected with mathematical practice were developed independently by Florencio Asenjo in 1954 and then very prominently by Newton C. A. da Costa in 1963. Asenjo presented what he called antinomic number theory. Da Costa introduced paraconsistent set theory; this was investigated with Ayda Ignez Arruda in the 1960s and 1970s. In 1986, da Costa proved that his system is nontrivial. This flourishing line of research continues in work by many, including Walter Carnielli, Marcelo Coniglio, Itala D'Ottaviano, João Marcos, and many others, and is sometimes called *paraconsistent mathematics*.[14]

Diderik Batens, founder of the adaptive approach to paraconsistency, worked with Arruda on set theory when she visited Europe in the late 1970s. Batens had proposed what he then called dynamic dialectical logic, which is intended to model reasoning in science and mathematics. He has most recently applied adaptive logics to set theory, in light of some significant developments from Peter Verdée.

In the 1970s in Australia, Richard Routley/Sylvan and Robert Meyer started investigating what they called "dialectical" theories, including set theory and arithmetic. Routley visited the State University of Campinas in 1976, and da Costa visited the Australian National University in 1977.[15] Unpublished manuscripts by Routley and Meyer were widely distributed, not only to colleagues in Australia but also da Costa and others, containing the outlines

[14] Interested readers may consult Gomes and D'Ottaviano (forthcoming).

[15] According to Arruda (1989, p. 107). Lengthy handwritten correspondence from the late 1970s between Routley and da Costa ("Dear Brother Richard/Dear Brother Newton") are preserved in the archive at the University of Queensland.

of what is now sometimes called *inconsistent mathematics*. These included Routley (1977) and Meyer (2021a, 2021b).

Circa 1976, Meyer produced a model of inconsistent arithmetic; such models were further developed by Meyer and Mortensen through the 1980s, and then Priest in the 1990s, using techniques due to Dunn. Important limiting results for Meyer's arithmetic were found by Harvey Friedman and Meyer in 1992. Circa 1978, Ross Brady produced a model of inconsistent set theory, which he developed further with Routley and reached culmination in a 2006 book; that set theory was studied further by Weber (me!). In 1995, Mortensen and collaborators published a book outlining what he termed inconsistent mathematics across several new areas, including calculus, topology, and category theory, and in 2010 followed this with further investigations in inconsistent geometry.[16]

Today, with the widespread acceptance of the legitimacy (if not correctness) of nonclassical logics, paraconsistent logics are being developed, sometimes with applications to mathematics, in many directions across the globe. The alternating terms "paraconsistent mathematics" or "inconsistent mathematics" have been used to some extent as shorthand to demarcate regional traditions and variations. Marcos has urged that logic does not carry a passport, and I similarly think there is little to be gained by insisting on geographic labels. What matters is that the formal development of paraconsistent logic(s) creates an opportunity in mathematics: the rigorous study of theories that are possibly or actually inconsistent.

1.2.3 Goals: Recapture, Expansion, Revision

Paraconsistency in mathematics deploys a nonclassical logic for mathematical work. How sweeping a change, if any, is this to ordinary practice?

We just saw that there is no one project or program called "paraconsistency in mathematics" or even "inconsistent mathematics." The story of paraconsistent mathematics told here is of two countervailing forces. On the one hand, there is a *moderate* or conservative line of research, seeking mainly to fit a paraconsistent logic into the wider standard mathematical picture, perhaps as a kind of buttress or insurance policy against any future inconsistency.[17] On the other hand, there is a *radical*, revisionary line that not only seeks to extend knowledge in novel nonclassical directions, but perhaps also to overturn previously accepted wisdom, rewriting the rules of the game. The conservative aim is to prove that $1 + 1 = 2$, even if there is some chance of inconsistency;

[16] Interested readers may consult Brady and Mortensen (2014).
[17] There is guarantee that mathematics is consistent, due to Gödel's theorems; see §3.