Tatton-Brown, O. M. W. (2020). Rigour and Proof. *Review of Symbolic Logic*. https://doi.org/10.1017/S1755020320000398

Rigour and proof

Abstract

This paper puts forward a new account of rigorous mathematical proof and its epistemology. One novel feature is a focus on how the skill of reading and writing valid proofs is learnt, as a way of understanding what validity itself amounts to. The account is used to address two current questions in the literature: that of how mathematicians are so good at resolving disputes about validity, and that of whether rigorous proofs are necessarily formalizable.

1 Introduction

What's going on in mathematical proofs? How do they establish the truth of their conclusions? By *proof* here I mean the kind of proof mathematicians actually write and exchange with each other and accept as valid.

One way to attempt to understand proof is via derivations, the formal objects that logicians use to model deduction. Gentzen intended his system of natural deduction to be

> a formalism that reflects as accurately as possible the actual logical reasoning involved in mathematical proofs (Gentzen 1969, p. 74)

and he described its derivations as having

> [a] close affinity to actual reasoning (ibid., p. 80).

One can read this as meaning that all valid mathematical inferences should be instances of the logical rules of natural deduction, or closely related to them.[1] If mathematicians did actually explicitly work according to certain fixed formal rules, then proof would be unmysterious: understanding proof would just come down to understanding the relevant formal rules. However as many authors point out, in reality mathematical proofs do not proceed according to any list of rules that one could specify in advance: there is far

[1]This is apparently the reading of Goethe and Friend (2010, pp. 274–275). It is also possible that Gentzen merely intended that the logical substructure of proofs should be expressible in natural deduction – aspects such as introducing and eliminating premises, moving from a statement about an arbitrary x to a statement about all x, and so on.

too much variety of inferences for that, and new proofs will often contain inferences that are somewhat (or completely) novel.[2]

More plausible than a naive rules based view of proof is one on which the correctness of a proof consists in its being *formalizable* – translatable into a derivation according to some given system of formal rules. What exactly this description amounts to will depend on what notion of translation is employed (and on the underlying system of formal rules). Though historically often implicitly assumed by philosophers of mathematics to be correct, numerous authors have recently objected to this view. There have thus been many calls to develop a new, more plausible account of mathematical proof and its epistemology.[3]

This paper aims to provide such an account, or at least the beginnings of it – found largely in sections 3 and 4. The account is not actually of proof in general, but only of *rigorous* proof, rigour being the standard of acceptable proof in much of modern mathematics. I do not view this as a significant limitation. Indeed I argue in section 2 that questions like "what is mathematical proof?", asked in full generality, are unlikely to receive a satisfying answer: there is no univocal notion of proof in mathematics, or at least not one we can expect to obtain a substantial philosophical analysis of. The account of rigorous proof given here is not the last word on the subject, and I note places where it could be expanded on.

The account put forward here takes a somewhat novel approach: instead of focusing exclusively on proofs, considering also the ability of mathematicians to produce and recognize valid inferences. Where does this ability to recognize validity come from? What can be said about it? Answering these questions is one way to gain insight into what is going on in proofs themselves - how they justify their conclusions.

The resulting account is used to address two questions about proof that have been raised in the literature. The first is that of how mathematicians are so able to generate agreement about the validity of proofs. This phenomenon – that if a mathematician thinks a proof is valid, they can generally convince others, or be convinced themselves of a flaw in it – has been noted by various authors, including Azzouni (2004, pp. 83–84) and Antonutti Marfori (2010, pp. 267, 270–271). Explaining it is one of the major motivations for the "derivation–indicator" view of proof that Azzouni (2004) puts forward. Azzouni's analysis of proof has met with controversy, with for

[2]These points are argued for instance by Tragesser (1992), Celluci (2009), Leitgeb (2009), Goethe and Friend (2010), and Larvor (2012). A nuanced version of the rule based approach to proof is given by Hamami (2019). Due to the arguments of the previously mentioned authors, I do not think an account like Hamami's can ultimately be successful, though I do not have space to discuss it in detail here.

[3]As illustrations of the objections, and the calls for improvement, see for instance Rav (1999, 2007), Celluci (2009), Detlefsen (2009), Leitgeb (2009), Pelc (2009), Goethe and Friend (2010), Antonutti Marfori (2010), Larvor (2012), Weir (2016), De Toffoli and Giardino (2016), and Larvor (2019).

instance Tanswell (2015) raising what appear to be valid criticisms. Using the account of rigour put forward here, in section 5 I give an explanation of this agreement about validity that aims to be simpler and more plausible than Azzouni's.

The second question is that of whether rigorous proofs are necessarily formalizable. In section 7 it is argued not that formalizability is directly required of valid proofs (as is sometimes thought), but that it is a consequence of the norm of rigour in mathematics – as spelled out in sections 3 and 4 – that valid proofs are necessarily formalizable. The worries of Rav (1999) and Weir (2016) about whether the process of filling in intermediate steps in an argument will ever terminate are addressed, and a possible response to the related worry of Pelc (2009) is sketched.

One cannot properly discuss formalizability without discussing what formal system one is targeting, and first section 6 considers the argument that mathematicians should really be regarded as working in naive set theory, with its unrestricted comprehension scheme. This has been claimed for instance by Jones (1998) and Leitgeb (2009). If correct this would render the formalizability claim empty of interest, since set theory with unrestricted comprehension is inconsistent and so any argument is trivially formalizable in it (with every inference justified by your favourite set theoretic paradox). This position is evaluated, and dismissed.

These are not all the worries about formalizability that have been or could be raised, and consideration of further qualms – such as the concern expressed by for instance Larvor (2012) that informal arguments may undergo some sort of violence, or essential loss, when formalized – will have to await consideration in further work.

2 Initial remarks on rigour

The main account of rigour is found in sections 3 and 4 and, but first there are some preliminary remarks worth making. To begin with, we will see some examples of non rigorous mathematics; this helps illustrate the distinctive nature of mathematical rigour, and is also used to argue against the idea that there is a unified notion of "proof" in mathematics that is worth conceptually analyzing. Indeed one of the intended lessons of this paper is that it is rigorous proof, not proof in general, which is the philosophically interesting concept. The latter part of this section summarizes the discussion of rigour given by Burgess (2015), which makes a number of valid, significant points, but which does not address the questions discussed in section 1 that this paper attempts to answer.

First, the examples of non rigorous mathematics. One good example consists of manipulations involving infinitesimals in the 17^{th} and 18^{th} centuries, which – before the introduction of limits into analysis – were not generally rigorous. For a toy example of how they often worked, we can determine the derivative of the function $x \mapsto x^2$. If we let dx be small, then we have

$$\frac{(x+dx)^2 - x^2}{dx} = \frac{x^2 + 2xdx + dx^2 - x^2}{dx} = \frac{2xdx + dx^2}{dx} = 2x + dx$$

and then since dx is small we discard it, obtaining $2x$ as the derivative of $x \mapsto x^2$ at x. Arguments along these lines (and more complicated versions) were carried out by various authors, with Fermat perhaps being the first to give this particular kind of calculation (Kline 1990b, pp. 344–345). These methods met with criticism however, as it was not clear what the status of "small" quantities such as dx was, or what was allowed when manipulating them. Indeed if dx is small but non zero then the result is only approximate; for the result to be exact, we require that $2x + dx = 2x$, but then we obtain by subtraction that $dx = 0$ and so we cannot divide by dx to begin with. Rolle pointed this out (Mancosu 1989, pp. 230–231), followed more famously by Berkeley who complained that infinitesimals appeared to be the "ghosts of departed quantities" (Berkeley 1999, pp. 80–81).

Other common methods in the 17$^{\text{th}}$ and 18$^{\text{th}}$ centuries also lacked rigour. Often arguments proceeded by assuming that what held for the finite also held for the infinite, with infinite series being manipulated as though they were finite sums, without worrying about issues of convergence. For instance Jacob Bernoulli argued (essentially) that

$$\begin{aligned} 1 &= \left(1 + \frac{1}{2} + \frac{1}{3} + \dots\right) - \left(\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots\right) \\ &= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \dots = \frac{1}{2} + \frac{1}{6} + \frac{1}{12} + \dots \end{aligned}$$

despite knowing that the sum $1 + \frac{1}{2} + \frac{1}{3} + \dots$ is infinite (Kline 1990b, p. 443).[4] As illustrated by many further examples in Kline (ibid., Chapter 20), infinite series were freely manipulated in this period without worrying about convergence, despite these methods sometimes leading to false or contradictory conclusions.

Arguments like these gradually came to be seen as unacceptable. In the 19$^{\text{th}}$ century both the calculus and the study of infinite series were rephrased in terms of the concept of limit, putting them on a firm footing (Kline 1990a, Chapter 40). Infinitesimals were then largely eschewed in analysis until Robinson demonstrated how one could in fact reason rigorously about them,

[4] One can adjust this argument to make it rigorous by telescoping the partial sums:

$$\frac{1}{2} + \frac{1}{6} + \dots + \frac{1}{n(n+1)} = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \dots + \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \frac{1}{n+1}$$

via the logical concept of a non standard model (Robinson 1996). Recent work has shown how the inconsistency in arguments like the above involving infinitesimals can be embraced and utilized, in a suitable paraconsistent logic (see for instance Brown and Priest 2004, and Sweeney 2014).

Other ways of reasoning formerly regarded as valid also came to be shunned, such as appeals to intuition. Proofs were demanded of even very intuitive statements, with the Jordan curve theorem being a famous example. This states that if $\phi : S^1 \to \mathbb{R}^2$ is continuous and injective then $\mathbb{R} \setminus \phi(S^1)$ consists of exactly two connected components, one of which is bounded and the other unbounded, and $\phi(S^1)$ is the boundary of each component. If one explained all the relevant terms – "continuous", "injective", "bounded" and so on – to a layman, they would likely state that this was simply immediate, and one might well get the same reaction from some undergraduates (particularly those with a preference for applied mathematics). It takes some effort to imagine a curve ϕ in such a way that the conclusion seems anything but obvious. Nonetheless mathematicians were not satisfied with this, and providing a rigorous proof turned out to be very difficult. Bolzano had already noted that this fact required proof (Coulston 1970, p.274) and this proof was only provided by Jordan (published in Jordan 1887) more than 30 years after Bolzano's death.[5]

There are a number of lessons to be drawn from these examples. Firstly, they can be used to assess the adequacy of an account of rigour, such as that given in sections 3 and 4 – it needs to be able to explain why they are not considered rigorous. Secondly, such examples illustrate that an argument does not have to be rigorous to be reliable, or explanatory, or valuable. Arguments involving infinitesimals in the 18^{th} century could plausibly be all three (such as the differentiation example above), as could manipulations of infinite series. Much of modern non rigorous physics and engineering is also presumably reliable, explanatory and valuable. Nonetheless rigour does bring benefits, some of which will be discussed in this paper.

These examples also illustrate the broadness of the notion of mathematical proof in times past. This, I believe, tells against the desire to seek a philosophical account of the general notion of proof – to discover what proof in general "really is", or where its boundaries are drawn. Indeed in the above examples, reasoning of various kinds all distinct from the usual mathematical paradigm of deduction is seen. To make arguments involving infinitesimals one postulates a new manner of calculation, in which a quantity is treated at one stage as non zero, and later as small enough that it can be neglected. This can otherwise be viewed as the postulation of a new kind of entity with these apparently odd properties. Either way, it is essentially

[5]Some controversy followed Jordan's proof, with Veblen (1905) claiming it was flawed, and claiming to give the first rigorous proof. However Hales (2007) argues that Jordan's proof was basically valid, though perhaps not as polished as it could be.

a form of abductive reasoning: reasoning of a kind which is not justified by anything that has gone before, but instead by its immense success in solving all manner of differential problems.[6] It is not so different to the postulation of new principles or entities in physics, except that it is confirmed by mathematical applications and deductions, rather than by experiments. Then the manipulation of infinite series as though they were finite is essentially a case of argument by analogy, again backed up by its apparent success in solving problems. Finally we have appeals to intuition, delivering conclusions that one finds very hard to doubt because of one's intuitive grasp of the concepts involved – not so different from the intuition that philosophical zombies could exist, or that nothing can cause itself.

I do not think there is much to be gained by seeking to discover what these disparate forms of reasoning "have in common". They are all taken to justify high credence in their conclusions, and they all concern abstract, mathematical subject matter, but beyond that I am not sure there is much to be said. Certainly one could conduct a fruitful investigation into abductive reasoning, or intuition as a form of evidence, but there is unlikely to be much distinctive to say about either in the context of mathematical proof that does not apply to more general contexts. There is also not much I think to say about why such methods were accepted, beyond that they were felt to be reliable. The best assessment of proof in general may just be that there are different kinds of permissible actions that one may carry out, as Larvor (2012) puts it. It is *rigorous* proof which I think is more deserving of philosophical attention: this is where the ideal of flawless deduction that Euclid aspired to takes its purest form (Burgess 2015, pp. 36–38), and where mathematical reasoning is found at its most distinctive, and epistemologically robust.

The variety of kinds of inference allowed in proofs historically makes it clear that it is only for rigorous mathematics that one could defend the formalizability of arguments. For instance what would a formal system for arguments like Bernoulli's, involving manipulation of divergent series, look like? Would it formalize analogies between the finite and the infinite? Similarly there is no reason to think that irreducibly intuitive reasoning would be formalizable.

Given the major differences between the historical standard of proof and what modern rigour permits, an obvious question is how and why the shift to rigour came about. Considering that would take us too far afield, but it is one topic which Burgess discusses (ibid., Chapter 1) in his account of rigour, to which we now turn. Burgess is mainly concerned with implications of the norm of rigour for the debate over structuralism, and does not give explicit

[6] As an anonymous referee remarks, part of the confidence in the use of infinitesimals stemmed from how in many cases one could use them to solve a problem – like the volume of a sphere – whose solution was already known, and thus check that the infinitesimal methods gave the correct results.

arguments concerning the kinds of questions raised in section 1 – the epistemology of proof, the ability of mathematicians to reach agreement about the validity of proofs, the issue of formalizability and so on. Nonetheless some observations he makes are worth highlighting.

Burgess emphasises that any piece of rigorous mathematics takes place in a context of existing results and definitions which can be appealed to (Burgess 2015, pp. 149–158). One can then extend the boundaries of knowledge with a new argument, stringing together definitions and proofs of propositions, appealing to existing results and using existing concepts where needed. Burgess also emphasises that often it does not matter whether a fact one is appealing to is a basic principle or a consequence thereof, or how concepts used were actually defined as long as the properties one needs of them do hold. This is the basis of his critique of structuralism as a metaphysical position.[7]

Both aspects of the rigorous process – definitions, and proofs of propositions – merit some attention. Burgess notes that when introducing a new concept, one is required a clear definition in terms of existing ones (ibid., p. 7). This definition does not have to be completely formal – for instance one can state that a vector space is a set equipped with an abelian group structure and a scalar multiplication operation, without specifying exactly how this is coded set theoretically: as a triple $(V, +, \cdot)$, or $((V, +), \cdot)$, or as $(+, \cdot)$, or in some other way. It just needs to be clear that the definition could be made completely precise in such a way that all uses made of the concept would be valid.

Burgess also discusses what the standard of rigour requires for proofs of propositions. He considers various possibilities, and ultimately comes to the (tentative) conclusion that:

> What rigor requires is that each new result should be obtained from earlier results by presenting enough deductive steps to produce conviction that a full breakdown into obvious deductive steps would in principle be possible (ibid., p. 97)

This I think is basically right. However there is more to say before this can be brought to bear on the issues discussed in section 1. If we are interested in the epistemology of proof, then this is only a sketch rather than a full account. How is this conviction generated? How is it reliable? If mathematicians are judging formalizability in principle (which is roughly what "full breakdown into obvious deductive steps" might amount to), how

[7]He argues that exactly how the concepts one uses were defined is often irrelevant, as all one will need are certain derived properties. Thus one need not care about how things were defined when doing mathematics. It is this irrelevance of definitions that Burgess argues has been mistaken by structuralists for a metaphysical truth about the nature of mathematical structures, with structuralists hoping to infer for instance that mathematical objects have only general structural properties (Burgess 2015, Chapter 3).

are they able to judge this? As discussed in section 1, most mathematicians have no experience of or interest in formalization, after all. There is also the question of how mathematicians are so good at resolving disputes, discussed in section 1, which we could hope to answer. Addressing these issues is the purpose of the remainder of the paper.

3 A rigorous eduction

As mentioned in section 1, the account of rigorous proof put forward here uses a somewhat novel approach: to try to understand the skill of mathematicians in judging and producing rigorous proofs by thinking about how this skill is acquired. For this we will start at the beginning. There are many different universities around the world that teach rigorous mathematics, and they may teach it in somewhat different ways, but there are some common features that can be pointed to. Students are generally taught the basics of rigorous proof by seeing and working through examples, paired with descriptions of how and why the reasoning involved works. An example of a basic early result students might see is displayed in fig. 1.

2.1.8 Theorem **(a)** *If $a \in \mathbb{R}$ and $a \neq 0$, then $a^2 > 0$.*
(b) $1 > 0$.
(c) *If $n \in \mathbb{N}$, then $n > 0$.*

Proof. (a) By the Trichotomy Property, if $a \neq 0$, then either $a \in \mathbb{P}$ or $-a \in \mathbb{P}$. If $a \in \mathbb{P}$, then by 2.1.5(ii), $a^2 = a \cdot a \in \mathbb{P}$. Also, if $-a \in \mathbb{P}$, then $a^2 = (-a)(-a) \in \mathbb{P}$. We conclude that if $a \neq 0$, then $a^2 > 0$.
(b) Since $1 = 1^2$, it follows from (a) that $1 > 0$.

Figure 1: "Introduction to Real Analysis", Bartle and Sherbert, p.26 (proof of (c) is omitted)

This demonstrates a fact commonly assumed without question: that the square of a non zero real is positive. Probably the only part of the argument that requires explanation is the symbol $\mathbb{P}$, which denotes the set of (strictly) positive real numbers. Axioms concerning this set have been stated a few pages previously. The relevant axioms are that:

(i) For any $a \in \mathbb{R}$, either $a \in \mathbb{P}$ or $(-a) \in \mathbb{P}$ or $a = 0$, with exactly one of these holding.

(ii) If $a, b \in \mathbb{P}$ then $a \cdot b \in \mathbb{P}$.

From these the above proof proceeds straightforwardly, arguing by cases.

The main thing to note about this proof is just how incredibly detailed it is. Virtually all of the logical structure of the argument is right there on the page. There are places where one could be even more explicit, in particular in the assertion that $a^2 = (-a)(-a)$, and indeed this follows immediately from the facts that $1 = (-1)(-1)$ and that $(-a) = (-1)a$, which are both given as exercises (Bartle and Sherbert 2000, p. 29). Nonetheless the proof is very close to the formal level and would be no challenge to formalize.

We can call this level of very great detail that proofs can be carried out at the "week 2 level of detail". Of course students may not see this particular argument in week 2, or at all; it is just a convenient name. We are not defining the "week 2 level of detail" here in terms of what is comprehensible to certain students– instead we give examples of basic arguments at this level of incredible detail, such as the above and also for instance basic number theoretic results (Taylor and Garnier 2014, Theorem 6.2; Silverman 2012, Lemma 7.1) or basic results from linear algebra (Axler 1997, Propositions 1.1–1.6).

As students learn the subject they won't just be passively reading proofs like this. They will also typically (and importantly) be proving these kinds of basic facts themselves, demonstrating them with arguments written out at this very explicit level of detail. The hope is that by doing this they will gain what we can call "proficiency at week 2 detail", the ability to prove simple facts like this one by chaining together these kinds of very basic steps.

A bit more will be said about how this basic level of proficiency is gained later in this section. For now we proceed onwards through the curriculum. As time passes the arguments the students are presented with will gradually get faster, and have fewer of the details filled in. Some time later – perhaps a few months, or a term – they may encounter an argument like that in fig. 2.

6.2.1 Interior Extremum Theorem *Let c be an interior point of the interval I at which $f: I \to \mathbb{R}$ has a relative extremum. If the derivative of f at c exists, then $f'(c) = 0$.*

Proof. We will prove the result only for the case that f has a relative maximum at c; the proof for the case of a relative minimum is similar.

If $f'(c) > 0$, then by Theorem 4.2.9 there exists a neighborhood $V \subseteq I$ of c such that

$$\frac{f(x) - f(c)}{x - c} > 0 \qquad \text{for} \quad x \in V, \ x \neq c.$$

If $x \in V$ and $x > c$, then we have

$$f(x) - f(c) = (x - c) \cdot \frac{f(x) - f(c)}{x - c} > 0.$$

But this contradicts the hypothesis that f has a relative maximum at c. Thus we cannot have $f'(c) > 0$. Similarly (how?), we cannot have $f'(c) < 0$. Therefore we must have $f'(c) = 0$. Q.E.D.

Figure 2: "Introduction to Real Analysis", Bartle and Sherbert, p.168

This theorem presents another fundamental fact: that if a function on an interval has a "relative extremum" – a local maximum or a local minimum – at an interior point, and is differentiable there, then the derivative must be zero. This is clear by visualizing the situation, but we are doing rigorous mathematics so are not satisfied with that, and we demand a proof.

The argument given is again fairly detailed, but is slightly less explicit than the previous example: not all the details are there. It only actually covers the case where $f'(c) > 0$, showing that this cannot happen, and the task of showing that $f'(c) < 0$ cannot occur is left to the reader. If the reader has understood the argument they should have no problem seeing how this would go, or writing it out. This aspect of the proof is fundamental to the way we learn rigorous mathematics. Students will hopefully not be treating proofs like the deliverances of some oracle: lecturers will ideally encourage them to engage with the proofs, to see if they could have proved the results themselves, to see if they can prove similar results by similar methods, and to see if they can fill in any parts where the proof is sketchy, and check any parts of the proof they are not sure about in more detail.

The course did not start by teaching students the week 2 level of detail material just to pad the schedule. The hope is that now when they meet an argument like this which is a little bit faster, strung together out of inferences that are simple but not necessarily completely basic, they can check any inference they are not sure of by proving it in more detail, using their "proficiency at week 2 detail" that they have hopefully already attained. Thus they can sharpen their judgement of which simple (but not completely basic) inferences are valid, checking such inference whenever necessary by seeing if they can be proved.

Students won't just be seeing theorems like this however. They will also be proving these kinds of slightly higher level statements themselves, by stringing together inferences that are simple (but not necessarily completely basic). By doing so they will hopefully gain what we can call "proficiency at term 2 detail", the ability to prove these slightly higher level inferences by stringing together simple inferences, and to reliably judge the validity of simple inferences (checking whenever necessary by proving them at the week 2 level of detail).[8] Again we define the "term 2 level of detail" by giving examples, such as the above and also for instance from Bartle and Sherbert (2000, Theorem 5.2.1), Axler (1997, Proposition 2.9, Proposition 2.13) and Silverman (2012, Lemma 9.2).

As the terms go by the students are gradually exposed to more and more condensed arguments. After another year or so they might meet an argument like that in fig. 3.

Here the students see a proof that power series can be differentiated term

[8]A note on terminology. I find it natural to speak of *proving inferences*, as in proving them in greater detail, though strictly speaking this may be a category error: inferences are things that we draw, assess, or justify, and we normally only speak of proving statements and propositions. Nonetheless I think it is clear what is meant – replacing a given inference by a chain of intermediate inferences, which collectively constitute a proof of the conclusion from the premises – and it is a convenient and expressive idiom.

9.4.12 Differentiation Theorem *A power series can be differentiated term-by-term within the interval of convergence. In fact, if*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \qquad \textit{then} \qquad f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} \qquad \textit{for} \quad |x| < R.$$

Both series have the same radius of convergence.

Proof. Since $\lim(n^{1/n}) = 1$, the sequence $(|na_n|^{1/n})$ is bounded if and only if the sequence $(|a_n|^{1/n})$ is bounded. Moreover, it is easily seen that

$$\lim\sup\left(|na_n|^{1/n}\right) = \lim\sup\left(|a_n|^{1/n}\right).$$

Therefore, the radius of convergence of the two series is the same, so the formally differentiated series is uniformly convergent on each closed and bounded interval contained in the interval of convergence. We can then apply Theorem 9.4.4 to conclude that the formally differentiated series converges to the derivative of the given series. Q.E.D.

Figure 3: "Introduction to Real Analysis", Bartle and Sherbert, p.270

by term. The proof is another step up in terms of compression, in terms of relying on the intelligence of the reader. This can be seen in the first line, where the reader is expected to see that a certain sequence is bounded if and only if another sequence is. Also in the second sentence, the reader is expected to "easily see" that a certain equation holds. These statements are extremely plausible; and if a student has any doubt, they can check them by proving them in more detail, using the "proficiency at term 2 detail" ability they have hopefully gained. They do not need to take these statements on trust, and they do not need to guess.

Again we can talk roughly about this "year 2 level of detail", giving further examples of arguments at about this level of detail to help explicate it, again for instance from analysis (Bartle and Sherbert 2000, Theorems 9.3.2 and 10.1.3), from number theory (Silverman 2012, Theorem 42.1), and also from ring theory (Aluffi 2009, Proposition III.3.11, III.4.5) and complex analysis (Bak and Newman 2010, Proposition 3.1).

This process continues in the obvious way. As the years progress a student is exposed to gradually faster and faster arguments, arguments where gradually more and more of the details are left out and more is left to the reader's intelligence. We can pick out further levels of detail a student will encounter, in the same way as above. First, we define a "year 3 level of detail" by giving examples, now with a wider variety: from functional analysis (Rudin 1987, Theorems 4.6–4.12), complex analysis (Conway 1978, §IV.2), measure theory (Fremlin 2010, Chapter 12), general topology (Munkres 2000, §33), algebraic topology (J. M. Lee 2000, Chapter 13), differential geometry (J. Lee 2012, Chapter 3), commutative algebra (Aluffi 2009, §V.1), representation theory (James and Liebeck 2001, Chapter 6), number theory (I. M. Niven, A. Niven, and Zuckerman 1991, §1.2), combinatorics (Szemerédi 1975, Facts 1 & 2), logic (Cori and Lascar 2000, §1.1), and category theory (Awodey 2010, Proposition 2.10). Again these may not be arguments a given student actually sees in their third year, but the level of detail is intended to be one that competent third year students will be gaining proficiency at, for both reading and writing proofs.

Detail here is not the same thing as accessibility. An argument can be very detailed but still difficult, for instance because it involves difficult concepts, or relies on difficult results, or because the result is poorly motivated and the proof strategy unexplained – or just because the argument is too long. Detail here means explicitness, and proximity to definitions, and how much the proof says of what could be said. It is the antonym of "how much is left out".

Naturally these predicates "week 2 level of detail", "year 2 level of detail" and so on will be vague: we may not be able to always determine precisely whether an argument is at the year 2 level of detail or not, just as we may not be able to decide whether a jumper is red, or perhaps orange instead. That does not undermine these predicates' validity or usefulness. Although there will be borderline cases, there will also be cases where we can in fact state with confidence that an argument is at around the year 2 or year 3 level of detail, rather than the week 2 or graduate level of detail (defined shortly). Of course an argument may not all take place at the same level of detail, so that describing the different levels of detail its parts take place at may be more appropriate than trying to shoehorn the whole argument into one category – as with a multicoloured jumper. One issue with these levels of detail that does not have such a comparison with jumpers is that it can potentially be quite difficult to compare the detail of pieces of mathematics from very different areas, where the reasoning is of a very different style. We can mitigate that as here by giving examples from a wide range of areas when characterizing levels of detail.

We now continue in the same way, defining further levels of detail a student will encounter. It should perhaps be emphasised that this terminology of levels of detail is new terminology I am introducing, and not a standard part of mathematical discourse. There are times one might see something like it used – for instance if a mathematician presented an unconvincing argument to a colleague, and after some questioning the colleague asked them to explain it more slowly, like they were talking to a grad student. Also, concepts like these can perhaps be seen as implicitly underlying some mathematicians' talk of detail in mathematics, an example of which will be seen in section 5 when discussing how these levels of detail can help mathematicians resolve disputes about the validity of proofs.

We will actually now pick out two different graduate levels of detail. First, we define the "graduate level of detail (explicit)" by giving examples, from functional analysis (Banach 1987, Theorem II.1), complex analysis (Conway 1978, §IV.6), measure theory (Schwartz 1954), general topology

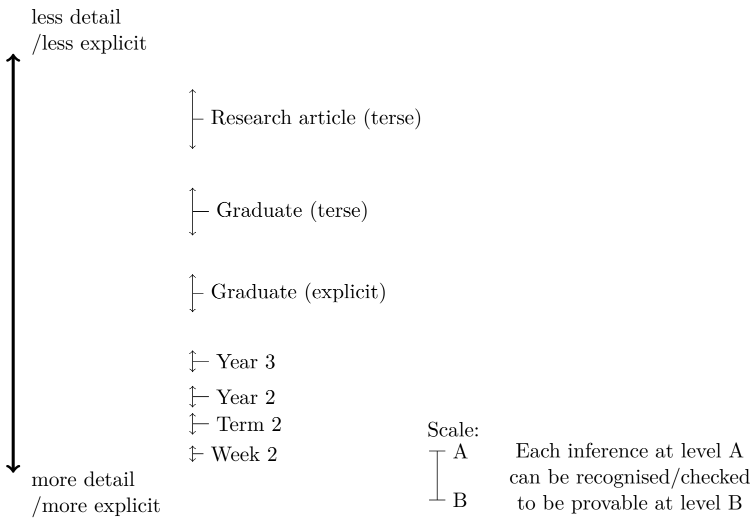

Figure 4: Levels of detail

(Walker 1974, §1.1–1.16), algebraic topology (Switzer 2002, Chapter 4), differential geometry (Hirsch 1976, §1.3), algebraic geometry (Eisenbud and Harris 2000, §I.1.4), commutative algebra (Eisenbud 1995, Chapter I.2), representation theory (Fulton and Harris 1991, Lecture 4), number theory (I. M. Niven, A. Niven, and Zuckerman 1991, §5.7), combinatorics (Erdös 1947, Theorem 1), logic (Prawitz 1965, Chapters I–III), and category theory (MacLane 1998, §II.3–II.6).

The basic idea is hopefully now clear, but we can keep going and pick out a "graduate level of detail (terse)" by giving some examples: from algebraic topology (Hatcher; 2001, §2.2), differential geometry (Thurston 1997, §2.7), algebraic geometry (Hartshorne 1977, §II.3) and category theory (MacLane 1998, Chapter IX). As seen in the examples above, some research mathematics is written out at a level of detail already covered: for instance Szemerédi (1975, Facts 1 & 2) at the year 3 level of detail, and Banach (1987, Theorem II.1), Schwartz (1954), Erdös (1947, Theorem 1) and Prawitz (1965, Chapters I–III) at the graduate level of detail (explicit). However plenty of research mathematics does take place at a level of greater compression, and one could keep going and pick out levels of this, named perhaps "research article level of detail (terse)", "research article level of detail (very terse)", and maybe one or two more. A range of levels of detail we can obtain in this way is seen in fig. 4.

To be clear, there is nothing privileged about the levels of detail listed in fig. 4. There is a continuum of levels of detail that we could potentially pick from (perhaps idealizing somewhat, given that statements in proofs are finite objects), and nothing to mark out those in fig. 4 as special; they are just useful examples for the purposes of this paper.

A little more should be said about the upper reaches of mathematics, at the top of fig. 4 and beyond. In fact it is clear that nothing too different happens as one approaches the research frontier – the gradual ascent to levels of greater and greater compression continues. That nothing radically different is happening can be seen in the ability of professors to take the most important, beautiful or useful results in a field – once it has matured – and collate them into a textbook with proofs accessible to graduate students. In the process proofs may be simplified or altered, but there is never any great obstacle to writing out what was once research level mathematics at a level of detail that graduate students can follow. One can see this in for instance the titles in the Springer Graduate Texts in Mathematics series.

A particularly extreme example of this is given by the case of Perelman's proof of the Poincaré conjecture. This was one of the most major conjectures in mathematics, and the subject of a Clay Millennium Prize. In 2002 and 2003 Perelman uploaded three papers containing a claimed proof (Perelman 2002, 2003b,a). The papers were written at a very high level, containing mathematics sketchy enough that despite only totally 70 pages, it took teams of mathematicians 3 years – working in correspondence with Perelman – to verify the argument as correct. Perelman was then offered a Field's medal, and subsequently a Millennium Prize, both of which he declined. Since 2006 more detailed expositions of his argument have been produced, such as Morgan and Tian (2007), which is a textbook intended to be accessible to graduate students. Indeed it looks to be at about the graduate level of detail (explicit) or graduate level of detail (terse). It comes in at 521 pages – about 8 times longer than Perelman's original papers. This is quite an increase, but even so it shows that there is not too dramatic a leap in terms of compression from maths at a level graduate students can understand to some of the most concise mathematics acceptable as a proof.

Now to say a little more about the lower end of fig. 4. First, a potential issue is that some areas of maths are generally only studied at a high level, because they have substantial mathematical prerequisites or typically involve subtle or complex arguments. This is the case with harmonic analysis, the theory of functions of several complex variables, and modern algebraic geometry, amongst other areas. This presents a potential problem with regards to the lower levels of detail in fig. 4, since for instance no-one does algebraic geometry in week 2 of their degree. In some cases this is unproblematic since expanding an argument in great detail will lead to basic inferences like those found in other areas – for instance combinatorics or analysis, in the case of harmonic analysis. In other cases reasoning used is more *sui generis*. Nonetheless in these cases I think it still makes sense to talk about what an inference carried out at say the term 2 level of detail would look like. In fact we sometimes see this happen: when a new manner of argument is introduced, even to advanced students, a few very explicit examples are often given of how it works. This is so that students have a sense of what is underlying more complex arguments, and know what to fall back on if they ever find more complex arguments hard to follow or produce. An instance of this from differential topology – a fairly advanced subject, done in full generality – is seen in J. Lee (2012, Proposition 2.4). Here Lee is giving an example of how to use the smooth charts on a manifold to prove local facts about them, and it is written up in great detail to make clear to students how this works, at around the week 2 or term 2 level of detail. One could do the same for other advanced subjects, for instance writing out the arguments of Eisenbud and Harris (2000, §I.1.4) from algebraic geometry mentioned above at the term 2 level of detail.

A second issue about levels of greater detail is one that was rather passed over in the discussion above: what takes place in the initial stages of learning rigorous mathematics, before the ascent up the levels of compression can start. To begin with, students will be taught how the basics of proof work, and what to do with the logical vocabulary $\wedge, \exists, \neg, \forall$ and so on, by a combination of examples and informal descriptions of what is going on and why. For instance they will (hopefully) learn by seeing and working through examples that to prove a statement $\forall n\, \phi(n)$ about all numbers, one can take an unknown number n, prove that $\phi(n)$ holds without assuming anything special about n, and then deduce that indeed $\forall n\, \phi(n)$ holds – essentially the $\forall$-introduction rule from natural deduction. The logical workings of proof can be stated in simple, clear, precise form – in terms of the hypotheses active at each stage, and how to introduce and exploit them – and when students can grasp how this works from examples it is not such a great surprise. Anyway not all students do manage to learn the rules from examples, and how to best teach the logic of proof is much discussed in the mathematics education literature (see for instance S. S. Epp 2003; S. Epp 2009). As well as the logical vocabulary students will learn the basics of set theory, including how to determine if two sets are equal via extensionality (whether this is an axiom or an inference rule is not important to the students, and the distinction may not be clear to them), and they will be shown various acceptable ways of forming new sets – power sets, cartesian products and so on.

There is a further basic aspect of proof that students are expected to infer from examples, and this is the ability to prove results by describing algorithms or procedures to achieve some desired mathematical goal, with Euclid's algorithm being an early example students often encounter (Silverman 2012, pp. 33–34), and further examples coming in linear algebra (Axler 1997, Proposition 2.6, Proposition 2.7; Artin 1991, pp. 14–15) and other areas. Though some such proofs can be rephrased as arguments by induction, sometimes they may implicitly require the definition of functions by recursion, and this is not usually something early undergraduates will be in a position to justify formally – the set theoretic treatment of recursion is typically taught later on in a more general form that applies to recursions on all ordinals (such as in Jech 2006, Theorem 2.15). In fact this does not present any sort of problem, and is not so different to the cases above where students grasp principles from examples; in this case the general principle implicitly underlying these kinds of recursive arguments is the axiom of dependent choice. This states that if X is a set and R is a binary relation on X such that for all $x \in X$ there is $y \in X$ with xRy, then for all $x \in X$ there is a sequence $(x_0, x_1, \ldots)$ of elements of X where $x_0 = x$ and for all i, $x_i R x_{i+1}$. That this is a statement rather than an inference rule, and not a basic axiom of set theory (it is deduced from the axiom of choice) is not important here. It is intuitive, and can be stated clearly, simply and precisely, and it is not a surprise that students can grasp from examples what kinds of arguments are in line with it.

The basic axioms governing sets are exempted from the general requirement that inferences be justifiable by proofs in greater detail. One can rightly assert this without actually deciding on which the basic axioms are; for instance it does not matter whether one regards the statement that function sets B^A exist as a basic axiom, or as justified by an argument that appeals to more basic axioms (union, separation, pair set and power set perhaps). Whatever the basic axioms are, they need not be justified by a proof, and other basic properties of sets are justified in terms of them (perhaps out of sight of students). There has been some discussion of exactly what means of set formation are allowed in mathematics – in particular, whether the unrestricted comprehension principle of naive set theory is used – and this issue will be addressed in section 7.

With these points addressed, we have the essentials of how rigour is learnt. We can pick out different levels of detail maths is done at by giving examples, and students proceed upwards through these levels of detail as described above: once they have gained proficiency at a certain level of detail they are in a position to engage with more concise arguments, with a tutored sense of what less explicit are valid – tutored by their experience at proving such inferences. If they are ever unsure, they can use their existing proficiency to check a larger inference and see if it can in fact be justified by a proof; and if so, they can ask why they were suspicious about it, and consider how to adjust their instincts to recognize such inferences as valid in future. It is essential to the normal process of learning rigorous mathematics that students are in a position to check inferences for themselves in this way, rather than just being presented with high level arguments they are intended to imitate. This is the most significant difference between mathematics as taught rigorously, and mathematics as taught in a physics degree (for instance). As discussed above there are also cases early on where students are expected to infer general principles from examples. There are only a few such cases though, and the reasoning each general principle encompasses can be characterized simply and precisely.

As mentioned in section 1, there are various places where this account could be expanded on. This could be a task for further investigations on the subject. For instance, one might seek a better understanding of exactly how students "adjust their instincts" to recognize a wider variety of inferences as valid, having seen particular cases to be provable in greater detail; or how, and with what degree of success, they "infer general principles from examples" in the early stages. There is not space to properly consider such questions in this paper, however.

4 The concept of rigour

Implicit in the process of learning rigorous maths described above is that each time a student is trying to master a new level of greater compression, it is constitutive of inferences being at that level of compression that they be provable at a previous level of detail, a level the student is already comfortable at – so that there are no leaps in the process of learning rigorous mathematics where a student is unable to check inferences for themselves (apart from when grasping certain basic principles). Indeed if inferences at the level of greater compression didn't need to be provable in more detail, then "checking" them by seeing which inferences can be justified with a proof would be a mistake. As discussed above, the basic axioms of set theory are exempted (whatever exactly they are taken to be), and are intended to be accepted by students without argument, though possibly with the assurance that they are "obvious".

Similarly, it is constitutive of rigour that nontrivial inferences can potentially be proved at a level of greater detail, so that if there is ever any unclarity – or disagreement – about the correctness of an inference it can be resolved by seeking a proof, or requesting one. Again proving in more detail here means an appreciable step up in detail, going up a notch in terms of levels of detail that we can pick out. This is one useful mechanism in mathematics for resolving disputes about the validity of proofs, as discussed in section 5. For the purpose of resolving disputes, this requirement is of less practical importance at levels of very great detail, as mathematicians may agree immediately about sufficiently basic inferences; but equally, for these simpler inferences it is generally more obvious that they can be proved in greater detail, and how such a proof would go, so the requirement is no unnecessary burden. Also, the ability to gain greater clarity about the correctness of inferences at all levels by seeking a more detailed proof is very important for the purposes of learning mathematics, as mentioned above. Again, the basic axioms of set theory are exempted from the requirement of being justifiable in greater detail: their correctness is not up for debate (within mathematics) and is supposed as a precondition for the mathematical enterprise to get going.

This is, I think, the key feature of rigorous mathematics: that there is a range of levels of detail that it can take place at, where inferences at a more compressed level can necessarily be proved at an appreciably greater level of detail. This requirement that inferences be provable in more detail is why the examples discussed in section 2 were not rigorous: the manipulations of infinitesimals and infinite series discussed could not be justified in greater detail, and could not be regarded as basic rigorous rules in themselves since it was not clear how to demarcate what reasoning was acceptable.[9] Similarly brute intuitions such as for the truth of the Jordan curve theorem are not in themselves suitable in proofs unless they can be backed up with more detailed arguments.

The necessity that inferences be provable in greater detail applies to levels of detail that students (and mathematicians) reach after progressing onwards from the basic level at which the subject is first taught – assumed above to be the week 2 level of detail. However not all students do manage to directly grasp how proof works at this level of detail, and some need more explicit demonstration of the rules that proof is implicitly following. There are courses and textbooks which provide this, such as Velleman (2006), which teaches how proof works essentially by teaching how to prove statements using the natural deduction rules: here the premises being used are explicitly tracked and calculated with, according to the rules governing the various bits of logical vocabulary. We can call this most basic, most explicit level of detail the "intro to proof level of detail". Students can use this as a stepping stone to gain comprehension of how basic arguments at the week 2 level of detail work, and implicit in this (as above with later levels of detail) is that inferences at the week 2 level of detail be provable at the intro to proof level of detail – otherwise gaining a grasp of how arguments work at the intro to proof level would be misleading as to what is going on at the week 2 level. Thus we can extend the above argument that inferences need to be provable in greater detail all the way down until we reach the intro to proof level of detail, where arguments explicitly use the natural deduction rules.

A minor caveat to this ability to prove in greater detail is that some arguments may require rephrasing when proving in more explicit formal terms – for instance one would often justify the Euclidean algorithm as an informal recursive process of repeated division with remainder, described for integers $a, b > 0$ perhaps by saying:

$$\text{Write } a = q_0 b + r_0 \text{ with } 0 \leqslant r_0 < b,$$
$$\text{write } b = q_1 r_0 + r_1 \text{ with } 0 \leqslant r_1 < r_0,$$

[9] Reasoning with infinitesimals was put on a firm footing by Robinson (1996) in terms of the logical concept of non standard model, and Nelson (1977) showed how to give a rigorous axiomatization for the approach. One can also use a paraconsistent approach to embrace the contradictory nature of infinitesimals, as seen in Brown and Priest (2004) and Sweeney (2014).

$$\text{write } r_0 = q_2 r_1 + r_2 \text{ with } 0 \leqslant r_2 < r_1,$$
$$\text{and so on, until we reach } r_n = 0$$

which if proved more explicitly would be transformed into some sort of formal recursive definition (of the form justified by the axiom of dependent choice – though there is no choice here – as discussed in section 3). In these kinds of cases one is replacing a part of a proof with a similar more detailed version, rather than literally filling in inferences in greater detail. However these kinds of cases are the exception rather than the rule, and do not make any essential difference to any of the discussion below, so we will generally include them under "proving in greater detail" (speaking a little loosely).

Now although the account above has focused on the ability to prove inferences in greater detail, it should be emphasised that one does not generally have to see how to prove an inference in greater detail to understand it, or accept it. For instance, to someone familiar with the notion of homeomorphism it probably feels obvious that the sphere $\{x \in \mathbb{R}^3 : \|x\| = 1\}$ is homeomorphic to the cube $\{x \in \mathbb{R}^3 : \max_i |x_i| = 1\}$, but sitting down and trying to write out a proof of this could well take a while. Such examples are not limited to topology. In logic – a subject where perhaps one would expect "intuition" would play less of a role – it might well feel obvious that substituting term t for variable x in a formula ϕ, when x is not free in ϕ, will just return ϕ, but again proving this in detail would take a bit of work (though probably less insight than the previous example). These kinds of higher level judgements about inferences – without a proof in mind – are an essential part of mathematics. Nonetheless it is important in rigorous mathematics that the option of proving inferences in more detail is there, to aid in gaining a firm grasp of any new concepts, and in guiding and sharpening one's judgement in any difficult cases – not all homeomorphic spaces are as obviously homeomorphic as the two above.

Also, when checking a proof one does not always necessarily actually check that every inference in is valid. Indeed if an unsurprising claim in a proof is supported by reasoning that looks like the right kind of thing, and the right amount of effort, then an experienced mathematician may pass on without checking every single detail. This is seen in interviews conducted by Andersen (2017) with mathematics referees, and also mentioned by Thurston (1997, p. 32). It appears this can be a fairly good guide to the overall correctness of results, though numerous commentators have remarked on the unreliability of the mathematical literature and the pervasiveness of errors in proofs, to which this manner of refereeing may be a contributing factor (Jaffe and Quinn 1993; Thurston 1994, p. 33; Nathanson 2008; Grcar 2013, pp. 421–422).

5 Disagreements about validity

It is traditional when studying deduction to think in terms of a single sharply defined notion of validity, that every inference either has or lacks. I think the above analysis of rigour rather tells against this conception.

Indeed having picked out various levels of detail that mathematical inferences can take place at, we can introduce a cumulative hierarchy of validity predicates – "valid at the week 2 level of detail", "valid at the term 2 level of detail", and so on, where for instance "valid at the term 2 level of detail" means an inference either at the term 2 level of detail, or at a level of greater detail. One can keep ascending in this way, defining validity predicates which allow larger and larger, more and more compressed inferences, inferences that are increasingly challenging for even an experienced mathematician to follow. At some point one will reach inferences compressed enough that they are well beyond the bounds of what mathematicians consider to be valid. However there appears to be no natural place on the continuum of levels of detail to draw a line, and say *this* is the limit of validity: that inferences at least that detailed are valid, while those less detailed are not. Asserting that there is a precise such limit seems to just be philosophical dogma, unsupported by the facts of the practice. Suppose for instance that a preprint of an article is uploaded to the arXiv, and read by two mathematicians experienced in the field – one of whom concludes that it is perfectly rigorous, the other that the proof of a certain lemma is too sketchy and incomplete. Who are we to say as philosophers that one is definitely right, and the other definitely wrong? I think it is more plausible to say that judgements of validity are – to some extent – both vague and subjective.

Some philosophers are aghast at this suggestion, it is worth explaining why, given the account of rigour above, it is not as damaging a claim as might be thought. In as much as there is disagreement about the validity of a certain inference, the framework of rigour provides a mechanism for resolving it.

That mathematicians are good at resolving disagreements about validity – that if a mathematician believes a proof is valid, they are generally able to convince others of this, or become convinced themselves of a flaw in it – has been noted by various authors as a fact that deserves explanation, and is one of the main motivations of Azzouni's derivation indicator view of proof is to attempt to provide this (Azzouni 2004, pp. 83–84; Antonutti Marfori 2010, pp. 267, 270–271). The idea behind Azzouni's account is that the informal proofs mathematicians write serve to indicate formal derivations. This has met with criticism, with Tanswell (2015) pointing out that attempted proofs may have many different attempted formalizations, which poses a problem since Azzouni wants to characterize validity of the informal proof in terms of success of the indicated formal derivation. Additionally as seen in section 3, mathematics generally proceeds at a much greater level of compression than is found in formal proofs, and (as mentioned in section 1) most mathematicians have no experience of or interest in the activity of formalization, so Azzouni's account is rather far removed from how mathematicians typically engage with proofs in practice.

Instead of hoping for an explanation in terms of completely formal proofs, it is more promising to look to the process of proving in more detail itself. Indeed if a mathematician ever puts forward a purported proof in which an inference is not convincing, then a more detailed justification for that inference can be requested – an appreciable step up to a level of greater detail, perhaps from the research article level of detail (terse) to the graduate level of detail (terse), as discussed in section 3.[10] At this level of greater detail inferences are more transparent and judgements of validity are more reliable, and this may already serve to resolve the controversy – with the new argument being acceptable, or an obvious error in it discovered. If not, and an inference in this more detailed argument is still controversial, a more detailed justification can be asked of it in turn, taking us to a still greater level of detail at which errors will be even more obvious – perhaps we now reach the third year level of detail. In principle this process will terminate when one reaches the level of complete formalization, though in practice if both sides are of sound mind and proceeding in good faith then the controversy will be resolved well before that.[11]

Thus if one believes a proof to be rigorous, in as much as this belief is correct one can always (in principle, and usually in practice) fill in the details of any inferences that are felt to be sketchy, to increase the level of detail to one which is found acceptable. In the imagined case considered initially of two mathematicians disagreeing about an arXiv preprint – with one finding that a particular lemma was argued for too briefly – the dispute would normally be resolvable in this way, bringing the proof into a form acceptable to everyone.

Thus though in my view there may be some subjectivity and vagueness to where exactly the limits of rigour are drawn, much more significant is the strong form of objectivity afforded by rigour, in which there are always robust reasons available to win over an objector – provided one's assessment of a proof as valid is justified.

In fact this dispute resolution mechanism is essentially that described in the Princeton Companion to Mathematics:

> [T]he fact that arguments can in principle be formalized provides a very valuable underpinning for the edifice of mathematics,

[10] Though this terminology of levels of detail is new, the process is not.

[11] The resolution of disagreement in this manner is an example of the idea from argumentation theory that a debate goes down to the level of detail that will satisfy both parties (using the apt words of an anonymous referee). For more on argumentation theory and mathematics, see Aberdein and Dove (2013).

> because it gives a way of resolving disputes. If a mathematician produces an argument that is strangely unconvincing, then the best way to see whether it is correct is to ask him or her to explain it more formally and in greater detail. This will usually either expose a mistake or make it clearer why the argument works. (Gowers, Barrow-Green, and Leader 2008, p.74)

Sections 3 and 4 can be seen as a clarification and elaboration of the process outlined in this quote. As an explanation of how disputes in mathematics can be resolved, this seems to be both more straightforward and more plausible than Azzouni's account, and better grounded in mathematical practice. However there may well be more to be said about the reality of disputes over validity in mathematics, and how well this simple account fits it.

6 Naive set theory?

The final topic of the paper is how this account of rigour impacts on the question of formalizability. Before the discussion of this in section 7, we take a brief detour from the main course of the paper, to address a sceptical view about the basic principles used in mathematics: the idea that mathematicians should be viewed not as working in a system like ZFC(U), but in naive set theory, with its axiom scheme of unrestricted comprehension. This has been suggested by Leitgeb (2009), and some mathematicians have made similar claims about their own understanding (Jones 1998, p. 205; Aluffi 2009, p. 1). Indeed "Naive set theory" is actually the title of a set theory textbook by Halmos (Halmos 2011). If mathematicians are best regarded as using unrestricted comprehension, this would make the claim that mathematical proofs can be formalized trivial, since set theory with unrestricted comprehension is inconsistent and so any argument can be immediately formalized in it (with every inference justified by a set theoretic paradox).

The key step when considering this possibility is to distinguish different senses of the term "naive set theory". Certainly most mathematicians do not know what the axioms of ZFC are, but they do have a solid grasp of how to legitimately form new sets: by taking unions, subsets, power sets, Cartesian products, function sets, equivalence classes, and so on. This understanding may be "naive" in the sense that it is not accompanied by explicit awareness of how these operations are justified in terms of the basic axioms – but that is totally different to "naive" set theory in the logicians' sense, in which the central principle is that of unrestricted comprehension, the scheme

$$\exists y \; (x \in y \Leftrightarrow \phi(x))$$

for all formulae ϕ in which y does not occur free.[12] Indeed the above set

[12] The set theory in Halmos's textbook is naive in an even weaker sense. Halmos does

forming operations are not generally justified in any more direct terms by unrestricted comprehension than by ZFC; to form Cartesian products for instance, one still needs to go through the rigmarole of defining what ordered pairs are and what a family of objects is, and the availability of unrestricted comprehension does not significantly simplify this. Moreover, there is no evidence of mathematicians making essential appeals to unrestricted comprehension, and this being accepted as valid. In the normal course of mathematics, all classes one would like to be sets are easily seen to be set sized using the standard set forming operations. In category theory, where size issues are encountered, the axiom of universes was introduced specifically so that they could be dealt with in a rigorous way. There are occasional instances where classes are manipulated as though they were sets, for instance in the definition of the Grothendieck group as a quotient of the set of isomorphism classes of finitely generated modules over a ring R; but in this case there is nothing genuinely troubling going on since one can easily define a set of representatives of the isomorphism classes instead (the quotient modules $\frac{R^n}{M}$ of powers of R, with isomorphic quotients identified), or one can appeal to the axiom of universes.

Moreover there are genuine mathematical cases where the distinction between sets and classes is crucial, and the use of unrestricted comprehension would be disastrous. For instance the general adjoint functor states that if $G : D \to C$ is a functor with D complete and locally small, then G has a left adjoint iff it preserves all limits and satisfies the solution set condition (MacLane 1998, p. 121). The solution set condition states that a set of morphisms with a certain property exists, and in this case the fact that this be a set rather than a class is key, as there is always a class of morphisms with the required property. In the presence of unrestricted comprehension, the general adjoint functor theorem would become the claim that any limit preserving functor whose domain is complete and locally small has a left adjoint, which is false in general. The issue of which functors have adjoints is not some category theoretic curiosity – it arises naturally in various areas of mathematics including algebra and topology.

Thus the way mathematicians form sets may be naive in the sense that it need not be founded in explicit knowledge of the basic principles, but there is no indication that it is naive in the sense of relying on unrestricted comprehension. If it did, signs of this ought not to be too hard to find.

---

in fact state the usual basic axioms of set theory (with no mention of unrestricted comprehension), and he uses them to derive various set theoretic operations and facts, but he says the approach is naive in that “the language and notation are those of ordinary informal (but formalizable) mathematics”.

7 Formalizability

Now to the question of formalizability itself. As advertised previously, it will be argued here that valid rigorous proofs are formalizable, in principle, though what this means requires clarification. The prospects for feasible formalization will also be touched on.

For the purposes of this section we will introduce the concept of "deductive grounding" between levels of detail. If L, L' are levels of detail, then we say that L' is *deductively grounded* in L if every inference valid at level of detail L' is provable at level of detail L.[13] This concept of grounding is justificatory rather than metaphysical, and does not share all the properties of standard notions of metaphysical grounding (for instance it is reflexive). This concept of grounding uses a notion of provability, and the kind of modality employed here needs to be spelt out before we know what deductive grounding amounts to. The notion of provability we will use is one of "in principle" provability, abstracting from limitations in terms of time or other resources (and thus perhaps abstracting away from the limitations of our own physical universe). One could otherwise use a notion of what is actually feasibly provable, given the physical and biological constraints on us, and obtain thereby a notion of "feasible deductive grounding".

We can relate this notion of deductive grounding to a notion of "in principle formalizability". Indeed, we can fix some standard system of first order natural deduction (for instance that of Prawitz 1965, Chapter I), and we take a formal proof to be a complete derivation in this system in the language of set theory with all premises amongst the axioms of ZFCU. We then define the "formal level of detail" to consist just of these formal proofs. We characterize a proof as being *formalizable in principle* if every inference in it is provable at the formal level of detail. Thus the claim that every proof at a level of detail L is formalizable in principle is just the claim that L is deductively grounded in the formal level. This is a weak notion of formalizability in principle, and for instance if the Riemann hypothesis is a theorem of ZFC then the one line proof of the Riemann hypothesis from no premises is formalizable in principle by this definition. Nonetheless it is one available sense of formalizability in principle; the question of what we do and should mean by formalizability will be returned to later in this section.

The key property of this notion of in principle provability is that it satisfies a version of the converse Buridan formula. In general the converse Buridan formula is (the scheme of formulae) of the form

$$\forall x \, \Diamond(\phi(x)) \Rightarrow \Diamond(\forall x \, \phi(x)).$$

[13]As was noted in section 4, some inferences at level L' may be part of an informal section of a proof which as a whole needs to be replaced by a more detailed and formal version at level L; this was illustrated with the example of the Euclidean algorithm. This caveat makes no real difference to what follows.

It is of the same general form as the converse Barcan formula, though with an existential instead of universal quantifier (Konyndyk 1986, p. 94). This converse Buridan formula is not typically valid. For instance if I have a well stocked fridge, it may be the case that for every item in the fridge I can have that item as part of my dinner, but that it is impossible for me to have every item in the fridge as part of my dinner. Nonetheless in cases where one abstracts from resource constraints it can be valid. In particular if we have a finite set S of inferences, and write Prov to indicate that we have obtained a proof of $s \in S$, then we do have that

$$\forall s \in S\ \Diamond(\mathrm{Prov}(s)) \Rightarrow \Diamond(\forall s \in S\ \mathrm{Prov}(s))$$

since if each element of S is provable, then – given sufficient time – it will be possible to obtain coeval proofs of every element of S.

It follows that if level of detail L' is deductively grounded in level of detail L, and we have a proof p of result s valid at level of detail L', then one can in principle obtain a proof of the s at level of detail L. Indeed every inference in p is provable at level of detail L, so that (as just discussed) it is possible to obtain a simultaneous proof of every inference in p; concatenating these then gives a proof of s at level of detail L, as claimed.

Thus we can obtain that the notion of deductive grounding is transitive. Indeed suppose we have levels of detail L, L', L'' with L' deductively grounded in L and L'' deductively grounded in L'. Let s be an inference valid at level of detail L''. Then by the definition of deductive grounding, we can in principle obtain a proof of s at level of detail L'; but then as just discussed, given such a proof one can in principle obtain a proof of s at level of detail L. Thus s is indeed provable at level of detail L.[14]

Now as discussed in section 4, it is crucial for rigour that valid mathematical inferences be provable in greater detail (unless they are already basic). This is an essential part of how rigorous mathematics is learnt, and of how validity is reliably judged – since for inferences that are not immediately convincing, one can always clarify the situation by seeking a proof in more detail. In both cases proving in more detail means an appreciable step up in detail, going up a notch in terms of levels of detail that we can pick out. This is an important mechanism for resolving disputes in mathematics, as seen in section 5.

We will now argue that for the levels of detail that a student moves through on their way to mastering research level mathematics, each more compressed level is deductively grounded in its more detailed predecessors. We will consider a student gradually moving through an education in mathematics, where at each time t there is the level of detail m_t of mathematics that they have mastered so far, and the level of detail l_t that they are learning at that point. We will assume that the collection of times t making up

[14] This argument could be carried out more formally, and implicitly uses the rule $\Diamond\Diamond\phi \Rightarrow \Diamond\phi$ of S4, which holds for the kind of metaphysical possibility being employed.

this period of education forms a complete ordered set, which we denote by I. I may or may not have endpoints – an initial point, and/or a final point. If L and L' are levels of detail we will write $L \leqslant L'$ to denote that L' consists of mathematical inferences at least as compressed as those at level L. We will assume that the function $t \mapsto m_t$ is monotonic, i.e. that if $t' \geqslant t$ then $m_{t'} \geqslant m_t$. For there to be no magical jumps in this process of learning, it needs to be the case that for all times t not initial in I, there is some $t' < t$ such that $m_t \leqslant l_{t'}$, so that the level of detail mastered at time t was actually learnt at some previous point in time. We also need that for all times t not final in I, there is some $t' > t$ such that $m_{t'} \leqslant l_t$, so that there is no magical jump after time t where at all subsequent times, a level of detail has been mastered that is greater than that which was being learnt at time t. Finally, as discussed above, it is constitutive of learning rigorous mathematics that at each stage, the level of detail one is learning is deductively grounded in a level of detail one has already mastered, i.e. that l_t is deductively grounded in m_t.

Now for the argument. For each t there is some $t' < t$ such that $m_t \leqslant l_{t'}$, but $l_{t'}$ is deductively grounded in $m_{t'}$, so that by transitivity m_t is also deductively grounded in $m_{t'}$ (call this backwards grounding). Also, for each t, there is $t' > t$ such that $m_{t'} \leqslant l_t$, so that $m_{t'}$ is deductively grounded in m_t (call this forwards grounding). Now suppose for contradiction that there is $s > t$ such that m_s is not deductively grounded in m_t. We let r be the infimum of

$$\{s \mid m_s \text{ not deductively grounded in } m_t\}$$

(a set which is bounded below by t). By forwards grounding, there is $t' > t$ such that $m_{t'}$ is deductively grounded in m_t, so that by monotonicity of m, we have that $r \geqslant t' > t$. Thus r is not initial in I, and so by backwards grounding for r, there is $t' < r$ such that m_r is deductively grounded in $m_{t'}$. But by the definition of r, we must have that $m_{t'}$ is deductively grounded in m_t, and thus by transitivity that m_r is deductively grounded in m_t. But then forwards grounding for r gives that there is $t' > r$ such that $m_{t'}$ is grounded in m_r, and thus in m_t; then by monotonicity of m, if $r \leqslant t'' \leqslant t'$ then $m_{t''}$ is deductively grounded in m_t, so that t' is actually a lower bound for

$$\{s \mid m_s \text{ not deductively grounded in } m_t\},$$

so that r is not the infimum of this set, giving the required contradiction. Thus from the assumptions we have made, it follows that if $s > t$, then m_s is deductively grounded in m_t.

Thus we obtain that all levels of detail that can be mastered rigorously, from a process starting at the week 2 level of detail, are deductively grounded in the week 2 level of detail. Then it is clear by inspection that arguments at the week 2 level of detail can be written out as formal derivations, as was noted in section 3, and their premises are all amongst the basic axioms of set theory, so that the week 2 level of detail is deductively grounded in the formal level. Otherwise, as discussed in section 4, one can define an "intro to proof level of detail", in which the natural deduction rules are explicitly used, and which some students use as a stepping stone to master the week 2 level of detail; it follows (as above) that the week 2 level of detail is deductively grounded in the intro to proof level of detail, whose proofs are trivially formalizable – they are already natural deductions of a slightly nonstandard kind. Either way we obtain that the week 2 level of detail is deductively grounded in the formal level, and thus that all the levels of detail that can be mastered rigorously are deductively grounded in the formal level. In other words, all rigorous mathematics is, in principle, formalizable. This is not a mere empirical fact obtained by looking at examples of mathematical arguments (except perhaps for the step from the formal level to the intro to proof or week 2 level), but a direct consequence of the norm of rigour in mathematics as enunciated here in terms of levels of detail.

This account answers a potential worry about formalizability: if rigour requires proofs to be formalizable, how are mathematicians so good at judging this? The answer is that mathematicians are not directly judging formalizability, but are instead judging the rigour of inferences (as discussed in sections 3 and 4), and that formalizability is obtained as a consequence of rigour.

In the literature the main worry about this kind of in principle formalizability is raised by Rav (1999), reiterated by Weir (2016). Rav considers a situation where one has an inference $A \to B$ in a proof, and after some thought fills this in with intermediate inferences to obtain $A \to A_1, A_1 \to A_2, \ldots, A_n \to B$. Perhaps one is then questioned by a student or non specialist as to why A_1 follows from A, and comes up with a new interpolation $A \to A', A' \to A_1$. Rav claims we can give no "theoretical" reason why this process of adding interpolations ought to ever terminate (Rav 1999, pp. 14–15).

The basic problem with this description is the lack of any attempt to characterize the form the interpolating inferences must take. One cannot just write in any intermediate inferences that suit one's fancy, whether justifying an inference to oneself or to a sceptic. For instance suppose we are trying to argue that the fact there are infinitely many primes (IP) is a consequence of the fundamental theorem of arithmetic (FTA), and we write RH for the Riemann hypothesis and ST for Szemerédi's theorem. It is clearly nonsense to try to justify FTA → IP by filling in intermediate inferences of the form

$$\text{FTA} \to \text{RH},\ \text{RH} \to \text{ST},\ \text{ST} \to \text{RH},\ \text{RH} \to \text{IP}.$$

These may be valid as material implications (in the presence of the axioms of ZFC), but they are totally useless as intermediate inferences for us; and to explain why they are the wrong kind of intermediate inferences, we have to start putting inferences on some sort of scale of plausibility, or simplicity, or evidentness, and require that adding intermediate inferences take us in the increasing direction on this scale. This is the first step towards thinking of mathematics in terms of levels of detail, as in sections 3 and 4, and to the requirement – implicit in Rav's description – that nontrivial inferences be provable in greater detail. Then the straightforward observations that 1) this requires an appreciable step up in detail, in terms of levels of detail that we can pick out, and that 2) there are only finitely many such steps up in terms of detail that we can pick out, and that a student can pass through on their way from week 2 mathematics to the research frontier, give the dissolution of Rav's worry (I do not know whether these count as a "theoretical" reason, in Rav's terms).

The argument given above concerned in principle formalizability, but it is also possible to say a little from this perspective about how practical formalization might be. Indeed we can use the above framework to address a weaker version of Rav's worry put forward by Pelc (2009), who accepts that the process of filling in with intermediate inferences will necessarily terminate, but questions whether we have any reason to believe this process will result in a formal proof of feasible length (given some initial proof of reasonable size). Pelc defines a vast number M in terms of various fundamental constants, large enough so that there is no possibility of us ever (in practice) constructing a formal proof of this length. His number M is at least the number of particles in the universe divided by the Planck time (in seconds), and thus at least 10^{120} on standard estimates; so a special case of Pelc's worry is whether when formalizing a proof of reasonable length, we have any reason to believe the resulting formal proof will be less than 10^{120} symbols.

The framework of levels of detail developed above can also be useful when addressing this kind of worry. Indeed one can get from the week 2 level of detail to the research article level of detail (terse) in a small number of steps up: via the year 2 level of detail, the graduate level of detail (explicit), and the research article level of detail (explicit). For each of these steps one can estimate by considering examples what kind of factor increase in length one generally obtains, when writing out an inference from the more compressed level at the more detailed level. Though it would have to be confirmed by more careful investigation, my belief from considering examples is that this factor is not generally too large, with a factor of less than 5 being common, a factor of 10 being fairly rare and a factor of 20 being very rare (as a proportion of inferences). This is supported by the example of Perelman's proof of the Poincaré conjecture discussed in section 3, where an extremely high level argument was written out at around the graduate level of detail with a factor of increase in length of 8. Thus though one may obtain exponential growth in proof length as one fills in details in a proof, to bring it down to the week 2 level of detail, the base and exponent are both fairly small: the former being the factor of increase with each step to greater detail, the latter being the number of such steps (4 in the case above). Even with a factor of 20 at each stage, and another factor of 20 (again a crude upper bound) to reach the formal level from the week 2 level, this gives us an overall factor of increased length of at most $20^5 = 3,200,000$ to formalize an argument at the research article level of detail (terse). This is vastly less than the factor Pelc is concerned might be necessary. Though this could undoubtedly result in very unwieldy proofs (if this crude upper bound was attained), they would still be within the bounds of feasibility, as usually conceived – requiring perhaps a few gigabytes or tens of gigabytes of space to store on a hard drive.

Even if one can fill out the above sketched argument to make a convincing case that all rigorous proofs of reasonable length will be feasibly formalizable, this does not meet one major kind of objection to formalizability: namely, that the process of formalization so dramatically changes a proof that the formal proof that results cannot rightly be regarded as the "same proof" as the original, and thus should not be regarded as a formalization of it. Larvor writes of the "violence or essential loss" that can result from formalization (Larvor 2012, p. 717). A related question concerns what practical relevance formalization has, or could have, to mathematics; it played only an indirect role in the account of rigour from sections 3 and 4, and turned out to be inessential to the process of resolving disputes in mathematics described in section 5. These questions will have to await possible consideration in future work.

8 Final thoughts

To conclude, there are a few loose ends concerning rigorous proof to discuss.

First, to say a bit about a topic that has been neglected here: understanding proofs. Indeed instead of just checking each line of a proof, one generally also wants to understand the proof "as a whole". This can be phenomenologically quite distinct from mere confidence that each line follows from the previous ones, as Tieszen (1992, pp. 58–59) emphasises. Though it is tempting to think about understanding in terms of the distinctive subjective subjective experience of grasping a proof, this characteristic sensation may not always be attainable – for instance it may be hard to gain the feeling of an immediate grasp of a proof taking longer than a page or so, or of a proof which relies on substantial previous results. An attractive alternative is to think of understanding as an ability, as advocated by Avigad (2011): for instance understanding a proof may mean that one can recreate it oneself, convey its ideas informally to another mathematician, use its ideas or techniques to prove similar results, generalize it, suggest how it could have been discovered, and so on. In many of these capacities that displays understanding of a proof, the ability to create or recognize valid proofs – proofs in which every inference is valid (as discussed above) – is key: recreating the proof means writing out a similar valid proof of the same result, generalizing it means writing out a similar valid proof of a more general result, and so on. Arguably, to convey the ideas behind the proof to another mathematician means to use words, gestures, diagrams and so on to equip the recipient with the means to recreate for themselves a similar valid proof of the result in question. As Thurston (1994, pp. 31–32) emphasises, it can often be much easier to convey mathematical ideas by informal communication than by embedding them in and then excavating them from rigorous proofs.

A second point is that in practice in mathematics it is not demanded that every inference in a proof actually be valid. Total validity is intended when many philosophers speak of proof (as has been the attitude in this paper), but in reality if an argument published in a mathematics journal contains a number of typos and minor logical errors, it may still be regarded as a perfectly acceptable proof. We could call the mathematicians' use of the word proof "proof in the weak sense", to distinguish it from the philosophical use of the term. The key feature I think for proof in this weak sense is I think that the proportion of valid inferences is high, or very high, and that those inferences which are invalid are each fixable relatively easily. Thus for instance the classification of finite simple groups may well be a proof in this weak sense, even though it is so enormously long that it is practically certain that it contains errors. This is more or less the view taken by Aschbacher (2005). A proof of a result in this weak sense still establishes that its conclusion is a logical consequence of the basic principles used – since in principle one could convert it into a proof in the strict sense, in which case the result would be a logical consequence of the relevant basic principles (as seen in section 7), and whether or not the conclusion is a logical consequence of the basic principles is independent of whether any such strict proof is actually written out.

Finally, a note on the epistemology of mathematics. In sections 3 and 4, an attempt was made to understand how rigour is judged in mathematics by thinking about how rigorous proof is learnt – by a gradual ascent up levels of greater and greater compression, at each stage being able to tutor one's judgement of which inferences are valid by checking if they can be proved in greater detail. There is undoubtedly more that could be said about this process, but at any rate it is only part of a full epistemology of mathematics. Indeed it was argued in section 7 that a valid proof shows its conclusion to be a logical consequence of the basic principles used; but this in itself does not imply that the conclusion is *true*. That would require further arguments, for instance arguments that the basic principles themselves are true. Thus the epistemology sketched here is one component of a full epistemology of proof, which would need to be supplemented by an epistemology of the basic principles themselves. Whether the basic principles generally used in mathematics – those of set theory – are true, and how we could know this, are exactly the kinds of epistemological questions that philosophy of mathematics has long wrestled with, and which some advocates of the shift to focusing on actual mathematics like to disparage, or describe as irrelevant to mathematics itself (for instance Rav 1999, Goethe and Friend 2010 and De Toffoli and Giardino 2016). It is of note that thinking about mathematical practice and the epistemology of proof leads us back naturally to exactly this standard epistemological question.

References

Aberdein, Andrew and Ian J. Dove, eds. (2013). *The Argument of Mathematics.* Logic, Epistemology, and the Unity of Science. Springer Netherlands.

Aluffi, Paolo (2009). *Algebra: Chapter 0.* Providence, R.I: American Mathematical Society. 713 pp.

Andersen, Line Edslev (2017). "On the Nature and Role of Peer Review in Mathematics". *Accountability in Research* 24.3, pp. 177–192.

Antonutti Marfori, Marianna (2010). "Informal Proofs and Mathematical Rigour". *Studia Logica: An International Journal for Symbolic Logic* 96.2, pp. 261–272.

Artin, Michael (1991). *Algebra.* Englewood Cliffs, N.J: Pearson. 618 pp.

Aschbacher, Michael (2005). "Highly complex proofs and implications of such proofs". *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 363.1835, pp. 2401–2406.

Avigad, Jeremy (2011). "Understanding Proofs". In: *The Philosophy of Mathematical Practice.* Ed. by Paolo Mancosu. Oxford, New York: Oxford University Press, pp. 317–353.

Awodey, Steve (2010). *Category Theory.* Oxford ; New York: OUP Oxford. 328 pp.

Axler, Sheldon (1997). *Linear Algebra Done Right.* New York: Springer. 251 pp.

Azzouni, Jody (2004). "The Derivation-Indicator View of Mathematical Practice". *Philosophia Mathematica* 12.2, pp. 81–106.

Bak, Joseph and Donald J. Newman (2010). *Complex Analysis.* Undergraduate Texts in Mathematics. New York: Springer-Verlag.

Banach, S. (1987). *Theory of Linear Operations.* Elsevier. 249 pp.

Bartle, Robert Gardner and Donald R. Sherbert (2000). *Introduction to real analysis.* John Wiley & Sons Canada, Limited. 416 pp.

Berkeley, George (1999). “The analyst”. In: *From Kant to Hilbert Volume 1: A Source Book in the Foundations of Mathematics*. Ed. by William Bragg Ewald. OUP Oxford, pp. 60–92.

Brown, Bryson and Graham Priest (2004). “Chunk and Permeate, a Paraconsistent Inference Strategy. Part I: The Infinitesimal Calculus”. *Journal of Philosophical Logic* 33.4, pp. 379–388.

Burgess, John P. (2015). *Rigor and Structure*. Oxford, United Kingdom: OUP Oxford. 240 pp.

Celluci, Carlo (2009). “Why Proof? What is a Proof?” In: *Deduction, Computation, Experiment*. Ed. by Rossella Lupacchini and Giovanna Corsi. Springer Verlag Gmbh, pp. 1–27.

Conway, John B. (1978). *Functions of One Complex Variable I*. Graduate Texts in Mathematics, Functions of One Complex Variable. New York: Springer-Verlag.

Cori, René and Daniel Lascar (2000). *Mathematical Logic Part 1: Propositional Calculus, Boolean Algebras & Predicate Calculus: A Course with Exercises*. Trans. by Donald H. Pelletier. Oxford ; New York: Oxford University Press, USA. 360 pp.

Coulston, Charles, ed. (1970). *Dictionary of Scientific Biography Volume II: Hans Berger - Christoph Buys Ballot*. Charles Scribner’s Sons.

De Toffoli, Silvia and Valeria Giardino (2016). “Envisioning Transformations - The Practice of Topology”. In: *Mathematical Cultures: The London Meetings 2012-2014*. Ed. by Brendan Larvor. New York, NY: Birkhäuser, pp. 25–50.

Detlefsen, Michael (2009). “Proof: Its Nature and Significance”. In: *Proof and Other Dilemmas: Mathematics and Philosophy*. Ed. by Bonnie Gold and Roger Simons. Washington, D.C.: Cambridge University Press, pp. 3–32.

Eisenbud, David (1995). *Commutative Algebra: With a View Toward Algebraic Geometry*. Springer Science & Business Media. 822 pp.

Eisenbud, David and Joe Harris (2000). *The Geometry of Schemes*. Graduate Texts in Mathematics. New York: Springer-Verlag.

Epp, Susanna (2009). “Proof Issues with Existential Quantification”. In: *Proof and Proving in Mathematics Education: The 19th ICMI Study*. Ed. by Fou-Lai Lin et al. Vol. 1, pp. 154–159.

Epp, Susanna S. (2003). “The Role of Logic in Teaching Proof”. *The American Mathematical Monthly* 110.10, pp. 886–899.

Erdös, P. (1947). “Some remarks on the theory of graphs”. *Bulletin of the American Mathematical Society* 53.4, pp. 292–294.

Fremlin, David (2010). *Measure Theory Volume 1*. Vol. 1. 5 vols. Torres Fremlin. 102 pp.

Fulton, William and Joe Harris (1991). *Representation theory: a first course*. Springer-Verlag. 576 pp.

Gentzen, Gerhard (1969). *The Collected Papers of Gerhard Gentzen.* Amsterdam: North-Holland Pub. Co.

Goethe, Norma B. and Michèle Friend (2010). "Confronting Ideals of Proof with the Ways of Proving of the Research Mathematician". *Studia Logica* 96.2, pp. 273–288.

Gowers, Timothy, June Barrow-Green, and Imre Leader, eds. (2008). *The Princeton Companion to Mathematics.* Princeton: Princeton University Press. 1056 pp.

Grcar, Joseph (2013). "Errors and Corrections in Mathematics Literature". *Notices of the American Mathematical Society* 60.4, pp. 418–425.

Hales, Thomas C. (2007). "Jordan's Proof of the Jordan Curve Theorem". *Studies in Logic, Grammar and Rhetoric* 10.23.

Halmos, Paul R. (2011). *Naive Set Theory.* Mansfield (CT): Martino Fine Books. 114 pp.

Hamami, Yacin (2019). "Mathematical Rigour and Proof". *The Review of Symbolic Logic.* (online only First View), pp. 1–41.

Hartshorne, Robin (1977). *Algebraic Geometry.* Graduate Texts in Mathematics. New York: Springer-Verlag.

Hatcher; Allen (2001). *Algebraic Topology by Allen Hatcher.* Cambridge University Press.

Hirsch, Morris W. (1976). *Differential Topology.* Graduate Texts in Mathematics. New York: Springer-Verlag.

Jaffe, Arthur and Frank Quinn (1993). "Theoretical mathematics: toward a cultural synthesis of mathematics and theoretical physics". *Bulletin of the American Mathematical Society* 29.1, pp. 1–13.

James, Gordon and Martin Liebeck (2001). *Representations and Characters of Groups.* Cambridge, UK ; New York, NY: Cambridge University Press. 468 pp.

Jech, Thomas (2006). *Set Theory: The Third Millennium Edition, revised and expanded.* Berlin ; New York: Springer. 796 pp.

Jones, Vaughan (1998). "A credo of sorts". In: *Truth in Mathematics.* Ed. by Harold G. Dales and Gianluigi Oliveri. Clarendon Press, pp. 203–214.

Jordan, Camille (1887). *Cours d'analyse de l'ecole polytechnique.* 3 vols.

Kline, Morris (1990a). *Mathematical Thought from Ancient to Modern Times, Vol. 1.* Vol. 1. 3 vols. New York: Oxford University Press. 432 pp.

— (1990b). *Mathematical Thought from Ancient to Modern Times, Vol. 3.* Vol. 3. 3 vols. New York: Oxford University Press. 444 pp.

Konyndyk, Kenneth (1986). *Introductory Modal Logic.* Notre Dame, Ind: University of Notre Dame Press. 144 pp.

Larvor, Brendan (2012). "How to think about informal proofs". *Synthese* 187.2, pp. 715–730.

— (2019). "From Euclidean geometry to knots and nets". *Synthese* 196.7, pp. 2715–2736.

Lee, John (2012). *Introduction to Smooth Manifolds.* New York ; London: Springer. 726 pp.
Lee, John M. (2000). *Introduction to Topological Manifolds.* Graduate Texts in Mathematics. New York: Springer-Verlag.
Leitgeb, Hannes (2009). "On Formal and Informal Provability". In: *New Waves in Philosophy of Mathematics.* Ed. by O. Bueno and Ø Linnebo. Basingstoke ; New York: Palgrave Macmillan, pp. 263–299.
MacLane, Saunders (1998). *Categories for the Working Mathematician.* New York: Springer. 332 pp.
Mancosu, Paolo (1989). "The metaphysics of the calculus: A foundational debate in the Paris Academy of Sciences, 17001706". *Historia Mathematica* 16.3, pp. 224–248.
Morgan, John and Gang Tian (2007). *Ricci Flow and the Poincare Conjecture.* Providence, RI: American Mathematical Society. 521 pp.
Munkres, James (2000). *Topology.* Upper Saddle River, NJ: Pearson. 537 pp.
Nathanson, Melvyn (2008). "Desperately seeking mathematical truth". *Notices of the American Mathematical Society* 55.7, p. 773.
Nelson, Edward (1977). "Internal set theory: A new approach to nonstandard analysis". *Bulletin of the American Mathematical Society* 83.6, pp. 1165–1198.
Niven, Ivan Morton, Avan Niven, and Herbert S. Zuckerman (1991). *An Introduction to the Theory of Numbers.* New York: John Wiley & Sons. 544 pp.
Pelc, Andrzej (2009). "Why Do We Believe Theorems?" *Philosophia Mathematica* 17.1, pp. 84–94.
Perelman, Grisha (2002). "The entropy formula for the Ricci flow and its geometric applications". arXiv: `math/0211159`.
— (2003a). "Finite extinction time for the solutions to the Ricci flow on certain three-manifolds". arXiv: `math/0307245`.
— (2003b). "Ricci flow with surgery on three-manifolds". arXiv: `math/0303109`.
Prawitz, Dag (1965). *Natural Deduction: A Proof-Theoretical Study.* Dover Publications.
Rav, Yehuda (1999). "Why Do We Prove Theorems?" *Philosophia Mathematica* 7.1, pp. 5–41.
— (2007). "A Critique of a Formalist-Mechanist Version of the Justification of Arguments in Mathematicians' Proof Practices". *Philosophia Mathematica* 15.3, pp. 291–320.
Robinson, Abraham (1996). *Non-standard Analysis.* Princeton, N.J: Princeton University Press. 320 pp.
Rudin, Walter (1987). *Real and complex analysis.* McGraw-Hill. 438 pp.
Schwartz, J. (1954). "The Formula for Change in Variables in a Multiple Integral". *The American Mathematical Monthly* 61.2, pp. 81–85.

Silverman, Joseph H. (2012). *A Friendly Introduction to Number Theory.* Boston: Pearson. 432 pp.

Sweeney, David John (2014). “Chunk and Permeate: The Infinitesimals of Isaac Newton”. *History and Philosophy of Logic* 35.1, pp. 1–23.

Switzer, Robert M. (2002). *Algebraic Topology - Homotopy and Homology.* Classics in Mathematics. Berlin Heidelberg: Springer-Verlag.

Szemerédi, E. (1975). “On sets of integers containing no k elements in arithmetic progression”. *Polska Akademia Nauk. Instytut Matematyczny. Acta Arithmetica* 27, pp. 199–245.

Tanswell, Fenner (2015). “A Problem with the Dependence of Informal Proofs on Formal Proofs”. *Philosophia Mathematica* 23.3, pp. 295–310.

Taylor, John and Rowan Garnier (2014). *Understanding Mathematical Proof.* Boca Raton: Routledge. 416 pp.

Thurston, William P. (1994). “On proof and progress in mathematics”. *Bulletin of the American Mathematical Society* 30.2, pp. 161–177.

— (1997). *Three-Dimensional Geometry and Topology, Vol. 1.* Ed. by Silvio Levy. Princeton, N.J: Princeton University Press. 328 pp.

Tieszen, Richard (1992). “What is a proof?” In: *Proof, Logic and Formalization.* Ed. by Michael Detlefsen. London ; New York: Routledge, pp. 57–76.

Tragesser, Robert (1992). “Three Insufficiently Attended to Aspects of Most Mathematical Proofs: Phenomenological Studies”. In: *Proof, Logic and Formalization.* Ed. by Michael Detlefsen. London ; New York: Routledge, pp. 162–198.

Veblen, Oswald (1905). “Theory on Plane Curves in Non-Metrical Analysis Situs”. *Transactions of the American Mathematical Society* 6.1, pp. 83–98.

Velleman, Daniel J. (2006). *How to Prove It: A Structured Approach.* Cambridge ; New York: Cambridge University Press. 398 pp.

Walker, Russell C. (1974). *The Stone-ech Compactification.* Springer-Verlag. 350 pp.

Weir, Alan (2016). “Informal Proof, Formal Proof, Formalism”. *The Review of Symbolic Logic* 9.1, pp. 23–43.