Global Philosophy (2023) 33:38
https://doi.org/10.1007/s10516-023-09691-1

ORIGINAL PAPER

Recalcitrant Disagreement in Mathematics: An "Endless and Depressing Controversy" in the History of Italian Algebraic Geometry

Silvia De Toffoli[1] · **Claudio Fontanari**[2]

Received: 20 October 2022 / Accepted: 27 July 2023

1 Introduction

The monograph *Lectures on Curves on an Algebraic Surface*, published by Mumford (1966), occupies a special place among works devoted to algebraic geometry. It is indeed a small masterpiece (only 200 pages long), containing both an enlightening presentation of the recent theory of schemes, developed in the 1960s by Alexander Grothendieck, as well as an astonishing application of that theory to the solution of one of the deepest and most obscure classical problems in the field. In Mumford's words,

> The goal of these lectures is a complete clarification of one "theorem" on Algebraic surfaces: the so-called completeness of the characteristic linear system of a good complete algebraic system of curves, on a surface F. If the characteristic is 0, this theorem was first proven by Poincaré in 1910 by analytic methods. Until 1960, no algebraic proof of this purely algebraic theorem was known. (*Although an endless and depressing controversy obscured this fact.*) (vii, *emphasis added*)

The protagonists of this "endless and depressing controversy" were Italian mathematicians Federigo Enriques and Francesco Severi. The disagreement, which unfolded in a series of articles published between 1921 and 1949, was over the correctness of

✉ Silvia De Toffoli
silvia.detoffoli@iusspavia.it

Claudio Fontanari
claudio.fontanari@unitn.it

1 Scuola Universitaria Superiore IUSS Pavia, Linguistics and Philosophy IUSS Center, Palazzo del Broletto, Piazza della Vittoria, 15, 27100 Pavia, PV, Italy

2 Department of Mathematics, University of Trento, Via Sommarive 14, 38123 Povo, TN, Italy

Published online: 07 August 2023

a series of putative proofs of the theorem discussed by Mumford, the so-called *Fundamental Theorem of Algebraic Surfaces*.

In this article, we aim to discuss the very possibility of disagreement in mathematics. To do so, we first develop a theoretical framework in which we identify several different types of disagreement in mathematics. We then focus on one such type that we find particularly interesting: recalcitrant disagreement about the correctness of multiple putative proofs of the same result. Our case study is a prime example of this type of disagreement. We then reconstruct the protracted controversy between Enriques and Severi and use it, in the discussion, to shed light on (i) the links between formalization, rigor, and correctness, (ii) the difference between a general criterion of rigor and specific working criteria of acceptability for rigorous proofs – which we put to use to characterize what *well-functioning* mathematical practices are, and (iii) the nature and very possibility of disagreement about the correctness of multiple putative proofs within the Italian school of algebraic geometry. We then discuss Enriques's and Severi's different attitudes to examine the role and status of speculative (non-rigorous) mathematics.

We proceed as follows. In Sect. 2, we put forward our theoretical framework by providing a working taxonomy of different types of disagreement in mathematics. We isolate a specific class on which to focus, *recalcitrant disagreement about the correctness of putative proofs*. In Sect. 3, we dive into our case study, the "endless and depressing controversy" between Enriques and Severi, which is a peculiar example of this type of disagreement. In Sect. 4, we discuss the controversy focusing on points (i), (ii), and (iii) above. In Sect. 5, we assess the proper role of speculative results within mathematical practice. In the Conclusion, we sum up our results.

2 Different Types of Disagreement in Mathematics

The phenomenon of disagreement has recently attracted the attention of epistemologists (Kelly 2010; Elga 2007; Christensen 2009). How should we respond to disagreement? Should we be *conciliationists* and lower our confidence in our beliefs or outright abandon them when others disagree with us or hold steadfast to them? It depends. The rational thing to do in the face of disagreement about a specific issue with an expert is certainly different compared to the rational thing to do in the face of disagreement about the same issue with a novice. It is to side-step this problem that the literature has mostly focused on *peer disagreement*, that is, on disagreement among *epistemic peers*, subjects who have comparable cognitive abilities and share the same evidence. But still, this is not enough to articulate a general reply. Different classes of beliefs need to be treated differently. For instance, it is reasonable to adopt divergent attitudes in the face of disagreement about core beliefs involving one's religious or political identity compared to beliefs about the result of a mental calculation.

Disagreement is widespread. We disagree about superficial as well as deep propositions. For example, we might disagree about the very existence of human-caused climate change or about the specific type of information that we can extract from climate-change models. Scientists might disagree about certain hypotheses or about the methodology of an empirical experiment.

However, in mathematics, disagreement seems rarer. Certainly, we do not disagree about basic mathematical propositions such as 2+2=4. The degree of consensus that the mathematical community often manages to achieve (at least in modern times) is striking and sets mathematics apart from other domains of discourse. Azzouni (2006, 208) called it the "benign fixation of mathematics." It is common to explain this "fixation" by appealing to the method used to establish mathematical propositions: mathematical proofs.

> What is called a "proof" of a given mathematical proposition, p, is really just a deduction (or deduction-sketch) of p from the relevant axioms.[1] In other words, a mathematical proof of p shows that if the relevant axioms are true, then so too is p. (Clarke-Doane 2020, 239)

Though mathematical results have been discovered in the greatest variety of ways, and never more so than today, with an abundance of computer numerical simulations and visualizations, it has long been accepted that one is not justified in claiming to have established a result as a theorem, and not just a conjecture, until one has produced a proof. It is true that many of what passed as "proofs" in the mid-nineteenth century or even later would not get past referees and editors of mathematical journals today. Still, already more than a hundred and thirty years ago, a leading mathematician as Henri Poincaré could claim that perfect rigor had been obtained:

> If we read a book written fifty years ago, the greater part of the reasoning we find will strike us as devoid of rigor. […] So we see that we have advanced towards rigor; and I would add that we have attained it and our reasonings will not appear ridiculous to our descendants […] But how have we attained rigor? It is by restraining the part of intuition in science and increasing the part of formal logic. […] Today only one [intuition] remains, that of whole number; all the others are only combinations, and at this price *we have attained perfect rigor*.[2] (Poincaré 1889)

If proofs are the only legitimate method for establishing a mathematical claim, then there is no room for disagreement in mathematics. But there is a caveat. A proof with proposition p as conclusion does not by itself justify proposition p; it justifies the conditional statement "if the axioms are true, then p is true." Thus, disagreement can arise with respect to the choice of the starting points. How are the axioms

[1] It is true that, as an anonymous referee pointed out, very few mathematicians would be able to state the "relevant axioms." This points to a divergence between this characterization of proof with actual mathematical proofs, whose starting points are subject-specific acceptable starting points, rather than (foundational) axioms. Still, mathematicians do generally acknowledge that such "relevant axioms" could be *in principle* made explicit – that is why this characterization will do for this context. The more so because, as it will be clear shortly, we won't focus on disagreement concerning axioms.

[2] It is not entirely clear what "perfect rigor" is, and even less clear whether that has been attained (Burgess and De Toffoli 2022; Paseau 2016). Remember that, of course, many of the proofs that were deemed to be rigorous at the time of Poincaré's writing, would not be acceptable nowadays. Nevertheless, it is important that still at the time, an ideal of rigorous proof was shared by many mathematical communities. We will return on this issue later.

themselves justified? There are multiple options.[3] Neither of these options, however, forces us to believe in the axioms in the manner that a proof forces us to believe in its (conditional) conclusion.[4] It is for this reason that the (relatively scant) literature on disagreement in mathematics has mainly focused on foundational disagreement over the choice of axioms (Clarke-Doane 2020). Well-known examples of AXIOM-DISAGREEMENT concern large cardinal axioms or, more mundanely, the legitimacy of the law of excluded middle.[5] As John L. Bell and Geoffrey Hellman (2006, 64) put it:

> Contrary to the popular (mis)conception of mathematics as a cut-and-dried body of universally agreed-on truths and methods, as soon as one examines the foundations of mathematics, one encounters divergences of viewpoint and failures of communication that can easily remind one of religious, schismatic controversy.

According to Bell and Hellman, disagreement in mathematics can be persistent (and even impossible to resolve) but is mostly circumscribed to foundational and philosophical issues.

Aside from issues about axioms and logical principles, Bell and Hellman consider questions such as: "'What is mathematics about?' 'What makes mathematical truths true?'" (*ibid.*, 65). These questions give rise to PHILOSOPHICAL-DISAGREEMENT. A prototypical disagreement about the philosophy of mathematics is over the nature of mathematical objects. According to mathematical platonists, they exist – and they are abstract, mind-independent, etc. Anti-platonists disagree.

In addition to these types of disagreement, there can be, of course, disputes over the practice itself. These disagreements target, for example, the hierarchical structure, the biases, and the power dynamics of a given mathematical practice. We call these SOCIAL PRACTICE-DISAGREEMENTS. By way of example, consider the case of category theorist Olivia Caramello. In (Rittberg et al. 2020) the authors classify her failed attempts to publish proofs of so-called *ghost theorems* (i.e., theorems that are largely accepted among the members of a mathematical community but lack a written published proof) as instances of epistemic injustice. These failed attempts arise due to the disagreement among the members of Caramello's mathematical community over how merit should be attributed and what results should be classified as novel. Another case of SOCIAL PRACTICE-DISAGREEMENT is concerned with the regulations of the attribution of Fields medals. According to Barany (2018), the fact that only young mathematicians are eligible to receive such a high honor has a negative effect both in terms of the image of mathematics as well as in terms of its actual practice, namely by perpetuating the misguided idea that a successful mathematician has to be a precocious genius (and quite likely male).

Is this all? Bell and Hellman claim that "there is indeed universal agreement on a substantial body of mathematical results[.]" (*ibid.*, 64). If that is so, then it might

3 See, for example, (Maddy 2011).

4 Arguably, however, some of the axioms of set theory "force themselves upon us as true" (Gödel 1964, 271).

5 Indeed, foundational disagreement can also concern the choice of a logic.

seem that disagreement over mathematical results is rare. Of course, there can be instances of mathematicians who are unreasonable or stubborn and disagree. But this is hardly of any interest. According to this straightforward picture, there cannot be *rational* disagreement over mathematical results conditional on axioms. But this picture is too good to be true, or, more precisely, it is too good to capture actual mathematical practice faithfully. There are at least three respects in which it departs from real mathematical practice: (a) mathematicians use methods other than proofs, (b) mathematicians disagree on how proofs should be conceived and individuated, and (c) mathematicians are not infallible in discerning proofs.

Regarding (a), mathematicians at times make use of so-called *probabilistic proofs* or other non-deductive methods[6] – we can call disagreement about what is the proper place of heuristics in mathematical practice, HEURISTIC-DISAGREEMENT. For example, according to Alexander Paseau (2015, 791), non-deductive arguments can generate mathematical knowledge: "we can derive knowledge of mathematical propositions from knowledge of related physical ones." To exemplify this claim, he appeals to simple geometric facts that, according to him, can be known through "paper and pencil experiments." But this position is up to dispute.

Regarding (b), mathematicians at times disagree on how proofs should be conceived and individuated. For example, disagreement may arise about what technological tools should be allowed in proofs. Should computer-assisted proofs be listed among genuine proofs? The first such proof, proposed in 1976 by Appel and Hacken, checked an enormous number of cases in order to establish the 4-color Conjecture. The proof, however, was not universally accepted because of the ineliminable role of computer computations it involved. Another related reason is the size of proofs. According to some, proofs should be the kind of things that a single subject with the relevant capacities and training should be able to grasp.[7] But this norm is violated not only by computer-assisted proofs but also by those proofs that involve large-scale collaborations – such as the proof of the classification of finite simple groups.

Moreover, mathematicians do disagree about how proofs should be individuated. When we formalize a proof, for example, how can we establish whether the resulting proof is the same proof we started with? Similar issues arise when we start with a proof containing diagrams and convert it into a diagram-free proof. [8] We can call these types of disagreement: CONCEPTION OF PROOF-DISAGREEMENT.

In this article, we will zoom in on (c). These are cases in which reasonable mathematicians disagree on whether a given putative proof (from now on, *p-proofs*) amounts to a genuine proof or not. We call this type of disagreement PUTATIVE PROOF-DISAGREEMENT or PP-DISAGREEMENT for short.[9] To be sure, there

[6] See (Paseau 2015). Note that, notwithstanding the (missleading) name, *probabilistic proofs* are not genuine (deductive) proofs at all.

[7] This is the "surveyability" requirement for proofs (Tymoczko 1979).

[8] For instance, De Toffoli (2023) argues that in some cases diagrams are essential to the proof in which they figure. That is, that there are plausible criteria of identity for proofs such that when eliminating the diagrams from certain diagrammatic proofs we would inevitably transform such proofs into different ones.

[9] This type of disagreement can also be seen to be linked to CONCEPTION OF PROOF-DISAGREEMENT. According to Wagner (2022, 5) the "problem of consensus [over the correctness of putative

are *hybrid* disagreements as well. An example will be useful. When Perelman proposed his proof of Poincaré's Conjecture, some mathematicians (students of Fields medalist Shing-Tung Yau) claimed that it was too gappy to be correct. They did not solely disagree with the substance of the proof, but also with the attribution of merit to Perelman (Nasar and Gruber 2006).[10] This case thus counts as a case of SOCIAL PRACTICE-DISAGREEMENT as well as PP-DISAGREEMENT.

Although PP-DISAGREEMENT has been mostly disregarded by philosophers of mathematics focusing on an idealized account of mathematics (and working mainly within a foundationalist framework), philosophers of mathematics more interested in the history and practice of mathematics have been made aware of it. Topics pertaining to the *Philosophy of Mathematical Practice* are gaining increasingly more attention in the recent literature. A precursor of such trend in the philosophy of mathematics can be found in Lakatos's work on the nature of mathematical proof. In Lakatos's quasi-empiricist account of mathematics, PP-DISAGREEMENT is a natural phenomenon since it arises from the elastic nature of mathematical concepts. However, it is hardly generalizable to the mathematics that developed after the process of rigorization at the end of the nineteenth century. In *Proofs and Refutations*, Lakatos (1976) offers a rational reconstruction of how an initial conjecture turns into a theorem. The process, he explains, is not linear but is marked by a series of proofs and refutations. For instance, counterexamples of different types are used to refine the original conjecture and the concepts involved – counterexamples naturally generate instances of PP-DISAGREEMENT because they show that what was put forward as genuine proof actually failed to amount to one. Lakatos' discussion focuses on the analysis of Euler's formula for polyhedra: V-E+F=2, where V are the vertices, E the edges, and F the faces of a given polyhedron. The text is in the form of a dialogue among (quite brilliant) students, and the ample footnotes supply historical context.

Let's take stock of the types of disagreements noted so far. This does not pretend to be an exhaustive taxonomy, but it nonetheless captures the main types of disagreement that arise in the context of mathematics:

1) AXIOMS-DISAGREEMENT
2) PHILOSOPHICAL-DISAGREEMENT
3) SOCIAL PRACTICE-DISAGREEMENT
4) HEURISTIC-DISAGREEMENT
5) CONCEPTION OF PROOF-DISAGREEMENT
6) PP-DISAGREEMENT

The first five types of disagreement are often persistent. Intuitionists reject the law of the excluded middle, anti-platonists reject the existence of mathematical objects, etc. As with many philosophical issues, there is no agreement in sight. This is true, albeit perhaps to a lesser extent, for SOCIAL PRACTICE-DISAGREEMENT, HEU-

proofs] and individuation of proof are intertwined." However, for the scope of this argument, we prefer to keep them separate.

[10] Thanks to Fenner Tanswell for encouraging us to think of this case as a hybrid type of disagreement.

RISTIC DISAGREEMENT, and CONCEPTION OF PROOF-DISAGREEMENT as well.

But what about disagreements over p-proofs? Unlike the other types of disagreement, it tends to last for a shorter time. That is, PP-DISAGREEMENT is often *evanescent*, at least if we consider mathematics that has, in Poincaré's words, achieved "perfect rigor," that is mathematics from the end of the nineteenth century onwards.

Consider, for example, the famous case of the 4-Color Theorem, the first argument for which was published by Alfred B. Kempe in 1879.[11] It was a convincing argument that was accepted by the mathematical community. Eleven years later, however, Percy Heawood found a substantial gap in Kempe's supposed proof. Contra Kempe and others, Heawood believed that the Conjecture had not been established. But once the mistake was spotted, it didn't take long to convince the other party. The gap was indeed there: there were some configurations that had not been considered by Kempe's p-proof. After attention was drawn to it, nobody objected.

This is not an isolated example. More common cases involve new results: if you think you found a proof, after carefully checking it by yourself, it is a good idea to submit it to external scrutiny and be ready to work on problems you failed to notice. In case of disagreement arising from experts scrutinizing your proof, it is widely accepted that the rational thing to do is to lower your confidence or even suspend judgement. After all, you are (or, at least, you should be) aware of your own fallibility. And this holds even in those cases in which you are right (that is, you do have a genuine proof), and they are wrong.[12]

PP-DISAGREEMENT is often *evanescent* because it is generally generated by specific mistakes or gaps. Even when a disagreement arises due to other reasons, say, the presence of a counterexample, a mistake is often spotted. This is possible because of the presence of what Easwaran (2015) calls the *convertibility norm*. Taking inspiration from Lakatos' (1976) discussion of counterexamples in mathematics, Easwaran argues that there is a norm governing mathematical practice operating on defeaters of p-proofs. When a *rebutting* defeater is found, that is, evidence that a result is false, such as a counterexample to a claim, then it must be possible to find an *undercutting* defeater, that is, evidence that the original evidence was misleading, such as a specific incorrect passage in the p-proof:

> Convertibility is a condition on potentially *incorrect* proofs – they should be such that any counterexample can reveal the incorrect step, which can allow us to replace the claimed theorem with a related theorem[.] (*ibid*., 156)

This norm contributes to the success of mathematics by underwriting the possibility of correcting local mistakes present in p-proofs that are seen to be problematic for global reasons. Convertibility can be satisfied in practice because the arguments that

[11] See (De Toffoli 2022, 256) for an analysis of this case from an epistemological perspective.

[12] Note that the correctness of this reaction is accepted even from those epistemologists who advocate that in general one should stick to one's own original position in face of disagreement (Kelly 2010, 199).

are accepted as proofs generally contain enough details to allow for step-by-step checking, as we will discuss later, this is connected to the rigor of p-proofs.[13]

> Although published proofs are not required to be so detailed that they are indefeasible (as the solutions to simple assignments in geometry and logic might be), they are required to have enough detail to impose some strong conditions on any potential defeater. (*ibid*., 148)

The convertibility norm assures us that, in general, all problems with our (rigorous) p-proofs can be *localized* and consequently swiftly resolved – this means that disagreement over p-proofs tends to be evanescent. Precisely because of their short-lived nature, evanescent PP-DISAGREEMENTS are not particularly problematic.

At least in recent times, after the rigorization of mathematics, PP-DISAGREEMENTS tend to be rare and to arise within mathematical practices that have the resources to resolve controversies. These are stable practices in which when problems arise, mistakes are spotted and progress is made – this is a crucial form of self-correction that the community exercises on itself.[14] These practices are reliable (albeit not infallible) in producing correct mathematical proofs – that is, not too many incorrect p-proofs are accepted by the practice. Because of their internal stability and the satisfaction of the convertibility norm, we call these mathematical practices *well-functioning*. Disagreement within well-functioning practices is often promptly resolved by conciliating and re-evaluating.

PP-DISAGREEMENT is more problematic when it is *recalcitrant*. As the name suggests, this type of disagreement tends to last for extended periods of time. Recalcitrant disagreement can be over a single p-proof or over multiple p-proofs of the same results. It is characteristic of mathematical communities that are not well-functioning, that is, that lack the resources for resolving disagreements over the results they produce (Fig. 1).

The type of recalcitrant disagreement that is at issue here is not linked to general conceptions of what proofs should be (CONCEPTION OF PROOF-DISAGREE-

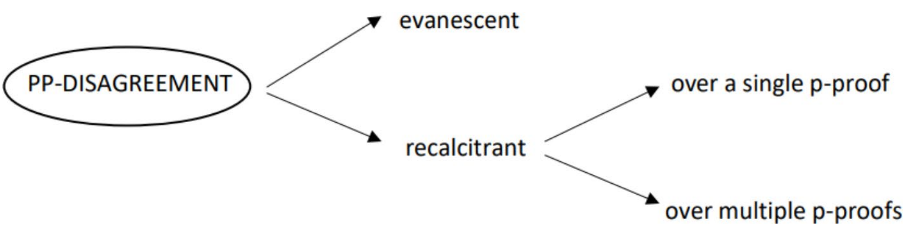

Fig. 1 PP-DISAGREEMENT: evanescent and recalcitrant

[13] This is, however, not always the case. Vladimir Voevodsky, long before finding a mistake in his own results that granted him the Fields medal, was made aware of a counterexample. He wrongly believed that his results were in good standing and that the putative counterexample came from a fallacious argument (Voevodsky 2014).

[14] In order for this process of self-correction to be possible, mathematical arguments must be *shareable* among practitioners, see (De Toffoli 2021a).

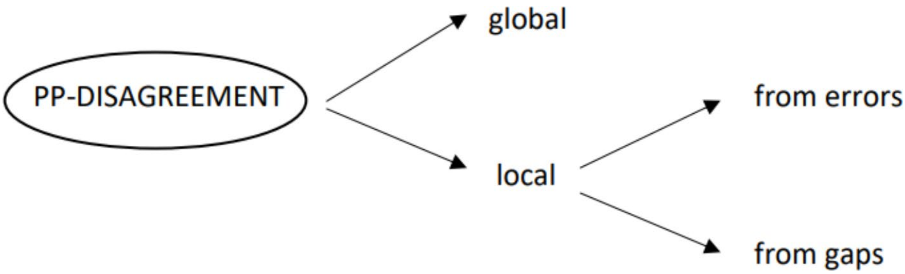

Fig. 2 PP-DISAGREEMENT: global and local

MENT), but to specific technical points, issues that have to do with counterexamples or the correctness of specific mathematical techniques. PP-DISAGREEMENT generated by these sorts of reasons can be recalcitrant because it defies quick resolution – as when the party disagree over the correctness of a specific step without managing to convince each other for an extended period of time – or, as in the case we will consider, because it is constituted by a series of evanescent PP-DISAGREEMENT over the correctness of different p-proofs of the same result.

The very existence of recalcitrant PP-DISAGREEMENT in recent mathematics seems to be controversial – for one, Bell and Hellman do not consider it at all (but we suspect that they would be happy to accept the existence of disagreement over general conceptions of proofs). We will argue in favor of its existence by examining a prominent instance of it from the history of modern mathematics. In our case, recalcitrant PP-DISAGREEMENT takes the form of a long controversy over the correctness of several p-proofs of a fundamental result about algebraic surfaces. A key difference between our case study and the one analyzed by Lakatos (which is also constituted by a series of evanescent PP-DISAGREEMENTS) is that we consider recalcitrant PP-DISAGREEMENT occurring after the turn of the twentieth century, that is, after the rigorization of mathematics. Indeed, recalcitrant PP-DISAGREEMENT is much more common and certainly less controversial if we look at pre-nineteenth century mathematics.[15] As Wagner explains: “disagreement concerning the validity of [putative] proofs was more deeply entrenched in that mathematical culture than we would expect from our acquaintance with contemporary mathematics” (Wagner 2022, 11).

Our case study is particularly interesting because it involves a number of different p-proofs of the same results and unfolds in multiple decades. For this reason, it can be set apart from other, even more recent cases of recalcitrant PP-DISAGREEMENT. Examples of such other cases are the one involving fundamental results in symplectic geometry (Hartnett 2017) and the one, which lately received much discussion, about the correctness of Mochizuki’s p-proof of the *abc* Conjecture, an important number theoretic conjecture – which, among other things, would imply Fermat’s Last Theorem (more on this later).

It is now time to turn to our case study.

[15] See (Goldstein 2013) discussing the ubiquity of PP-DISAGREEMENT in 17th century mathematics.

3 An Endless and Depressing Controversy

3.1 The Italian School of Algebraic Geometry

What is now commonly referred to as the *Italian school of algebraic geometry* lasted for about a century. We can trace its origin to the day when Luigi Cremona obtained the first Italian chair in *Geometria Superiore* at the University of Bologna: on July 10, 1860. Its end can be plausibly said to coincide with the death of Francesco Severi: on December 8, 1961. The apex of its trajectory was reached before the First World War. This was followed by a declining phase, which started in the Twenties and was at first nearly imperceptible but then became quite evident from the Thirties on.[16] Castelnuovo (1930, 614) characterizes the *school* as follows:

> The characteristic mark of the school is embodied by the founding father Cremona, who taking up the questions from the beginning, builds a new foundation for projective geometry and elevates it to a science, perfecting its methods and *meshing geometric intuition in the most skillful way with certain fundamental algebraic results* [fondendo nel modo più abile la intuizione geometrica con alcuni risultati fondamentali tolti dall'algebra]. And he was able to use these procedures with such sagacity as to allow the new geometric algebra to discover, often effortlessly, hidden properties that classical algebra, weighed down by the baggage of formulas, was only able to find with difficulty.[17] (*emphasis added*)

The three foremost exponents of the school – Castelnuovo, Enriques, and Severi – continued to follow their geometric intuition and search for mathematical simplicity. This was the primary characteristic feature of the Italian school and was its main strength.[18] But, as we shall see, it was also the weakness that ultimately led to its decline.

Although its members shared a common outlook on mathematics, the Italian school was by no means unified. For instance, there were multiple instances of disagreement between Enriques and Severi. The one we are about to present is especially interesting due to its importance, robustness, and sheer duration.

3.2 The Controversy

In the Introduction, we cited a passage by Mumford (1966, vii):

[16] See (Brigaglia and Ciliberto 1998, 300).

[17] Our translation from Italian.

[18] This is not, according at least to Severi and Enriques, incompatible with mathematical rigor. Indeed, both mathematicians identified two types of rigor, one connected with the possibility of always adding more details and another, consisting in describing the mathematical facts faithfully. For a discussion, see (De Toffoli and Fontanari 2022).

> The goal of these lectures is a complete clarification of one "theorem" on Algebraic surfaces [...]. Until 1960, no algebraic proof of this purely algebraic theorem was known. (*Although an endless and depressing controversy obscured this fact*). (*emphasis added*)

In his discussion of the *Fundamental Theorem of Algebraic Surfaces*, Mumford adopts the modern mathematical language introduced by Grothendieck – a language that was not available to the members of the Italian school. Here is a brief (and inevitably very technical) description – this is just to give an idea of the description of the result under discussion in modern terms.

The *Fundamental Theorem of Algebraic Surfaces* computes the dimension of the so-called *Picard variety*, parameterizing isomorphism classes of line bundles on a given algebraic surface S which are algebraically equivalent to the trivial line bundle 0_S. This algebraic variety is naturally a commutative group scheme; hence it is reduced in characteristic zero, according to a theorem of Cartier. It follows that its dimension equals the dimension of its Zariski tangent space at the origin, which turns out to be the first cohomology group $H^1(S,0_S)$ of the trivial line bundle 0_S. According to the terminology of classical algebraic geometry, the dimension of $H^1(S,0_S)$ is called the *irregularity* of the surface S.

For an attempt at a less technical statement of the *Fundamental Theorem of Algebraic Surfaces*, we refer to (Babbitt and Goodstein 2011, 240).[19] This theorem is part of the general effort of the Italian school towards the study and the classification of algebraic surfaces over the complex numbers – being surfaces, these have complex dimension two, but they correspond to real varieties of dimension four, and therefore they are not easily visualizable.

Very roughly, the Fundamental Theorem is a tool that allows us to analyze (and, to a certain extent, classify) algebraic surfaces (over the complex numbers) which are *irregular*. It is impossible to define irregular surfaces without entering technical details that will take us too far afield. However, just to give an idea, this is done by appealing to different notions of *genus*. Algebraic curves over the complex numbers correspond to surfaces over the real numbers and can be classified using topological genus alone. The same does not hold for algebraic surfaces. In this latter case, it is common to use two different notions: *arithmetic* and *geometric* genus. It is precisely the gap between these two notions that determines their *irregularity*. The Fundamental Theorem then tells us that a geometric space naturally associated with a surface of irregularity q (its so-called *Picard variety*) has dimension exactly q.

The following quotation from 1935 by Oscar Zariski, Mumford's dissertation advisor, still provides a good description of the situation and of the absence of an algebraic proof of the Fundamental Theorem:

> the quantitative specification that an algebraic surface of irregularity q possesses complete continuous systems consisting of ∞^q distinct linear systems [this is the formulation of the Fundamental Theorem], is a fundamental result of

[19] They discuss a related result, the *Theorem of Completeness of the Characteristic Series*. The Fundamental Theorem under discussion is one of the main consequences of such Completeness Theorem.

> the theory of surfaces, due to the combined efforts of several geometers (Humbert, Castelnuovo, Enriques, Severi, Picard, Poincaré). [...] A proof of the above fundamental result has been first proposed by Enriques (1904). Immediately after, another proof was proposed by Severi (1905). Severi himself has later pointed out that neither proof is entirely rigorous (1921). [...] A rigorous proof of the existence on a surface of irregularity q of continuous systems consisting of ∞^q linear systems was given for the first time by Poincaré (1910). Poincaré's proof is analytic and has been subsequently simplified by Severi (1921) and by Lefschetz (1924). (Zariski 1935, 98–102)

Indeed, Zariski's formulation of the *Fundamental Theorem* in terms of "the existence on a surface of irregularity q of continuous systems consisting of ∞^q linear systems" is just the traditional way to claim that the irregularity q is the number of parameters on which the Picard variety depends on. Following Zariski, we can trace the main events that gave rise to this long controversy:

> 1904: Enriques publishes an incorrect p-proof of the theorem.
> 1905: Severi publishes another incorrect p-proof of the theorem.
> 1910: Poincaré proposes an analytic proof.
> 1921: Severi criticizes both his own and Enriques's p-proofs of 1904–1905 for lack of rigor.

The disagreement arose from the fact that both Enriques and Severi thought their own p-proofs were correct. What was the problem with Enriques's and Severi's p-proofs? After some time, Severi realized that both arguments relied on a crucial assumption, whose available p-proof did not work in full generality – the p-proof, in other words, contained a significant gap. More precisely, Zariski explains:

> In both proofs we have, at the start, a linear system of curves [...] and then on the curves of this system a certain number of algebraic non-linear conditions is imposed [...]. Both proofs make use of the assumption that in the algebraic system of curves thus obtained [...] the linear system of curves infinitely near to a generic curve of the system is *complete*. Severi's criticism (1921) is to the effect that the available algebro-geometric proof of this assumption fails *if the characteristic series of the considered continuous system is special*. (*ibid., emphasis added*)

The problematic assumption concerned the completeness of a certain algebraic system of curves and disqualified both Enriques's and Severi's p-proofs as rigorous.

If Poincaré provided a rigorous proof in 1910, why wasn't the issue settled then? The reason is that both Poincaré's original proof and its subsequent simplifications applied anaytical tools in order to solve a genuine algebraic problem. This approach was not considered satisfactory. That is why both Enriques and Severi continued (in vain) to pursue the discovery of a self-contained proof within the conceptual framework of algebraic geometry. Ultimately, this goal was accomplished by Mumford in the 1960s, relying on the highly sophisticated and technically heavy theory of algebraic schemes developed by Grothendieck.

When Zariski published his treatise in 1935, the “endless and depressing controversy” mentioned by Mumford was still waiting to explode in the forthcoming mathematical duel to be fought in 1942 on the pages of the *Commentarii Mathematici Helvetici* (Enriques 1942; Severi 1942). Severi’s (1944) reconstruction of the controversy was published shortly after – he begins by considering another fallacious p-proof, published in 1938 by one of his former students (Segre 1938):

> some criticism by F. Enriques accompanied by an attempt to simplify the proof of the fundamental theorem led me to return to the age-old and very delicate question, [...]. But then I became aware of a much more serious fact: namely that the proof of the fundamental theorem on regular systems [the Fundamental Theorem] given by B. Segre (which I had tried to extend to superabundant systems) fell (as I wrote) due “to unexpected phenomena [...] which show up under a closer scrutiny of the question”. Hence, I deduced that the question remained unmoved at the point where my 1921 contribution had left it [.][20] (Severi 1944)

Enriques’s point of view is expressed in a private letter sent to his colleague Beniamino Segre in 1945:[21]

> Personally, I am especially interested in the problem of a continuous system on irregular surfaces, and this in view of the publication of my lectures on surfaces, written in 1942 and now in press. The question is extremely delicate. *I was unable at that time to reconstruct the proof that you had indicated on the basis of the information that you had given me* before my departure for Paris. Severi, with whom you have had more interaction, believed he had finally succeeded in giving a proof. His exposition seemed obscure to me and therefore dubious; I believed (in the *Commentarii Helvetici* paper) to have overcome the difficulty. (Enriques 1942, in: Babbitt and Goodstein (2011), *emphasis added*)

In this private epistolary conversation, Enriques criticizes Severi for lack of rigor. However, he is also aware of his own inability to solve the issue. His letter continues thus:

> In reality, my proof was erroneous, but this realization also pointed out the error in Severi’s proof. At that time, I was not allowed to add anything to my paper in the Commentarii Helvetici although Severi was allowed to write a note in which he said that he had derived a more general theorem (he referred to the case of algebroid entities instead of clarifying the matter in the easiest case). But shortly afterward, Severi himself, who was expounding that theory in his lectures at the Institute of Higher Mathematics, realized that his proposed proof

[20] Our translation from Italian.

[21] This letter is stored in the Beniamino Segre Archives at the California Institute of Technology and reproduced (and partially translated) by Babbitt and Goodstein (2011).

> was flawed due to a radical error. I really wish that this thing could be settled. (*ibid.*)

But the issue, unfortunately, remained unsettled:

> I have reexamined my earlier proof based on infinitely close curves of various orders and I believe it is substantially right, even if it is not rigorously complete. (*ibid.*)

While Severi's approach (following Segre) turned out to be irreparably defective, Enriques's claim that his argument is "substantially right, even if it is not rigorously complete" is indeed accurate. As Mumford would later put it, "he certainly had the correct ideas about infinitesimal geometry, though he had no idea at all how to make precise definitions" (Mumford 2011, 250).

> Here, then, are the further steps of the controversy:
> 1938: Segre proposes another incorrect p-proof (which is irreparably defective).
> 1942: Enriques proposes another incorrect p-proof (which lacks rigor but is based on correct ideas).
> 1942: Severi (after himself being criticized) criticizes Enriques's p-proof and proposes another incorrect p-proof (which contains a radical error).

The controversy continued even after Enriques's death in 1946 (see Enriques 1949 and Severi 1958).

4 Discussion

While it is common to hear of instances of PP-DISAGREEMENT in mathematics, this case is particularly significant because it is about a core result. Moreover, it lasted multiple decades. It is an instance of recalcitrant PP-DISAGREEMENT over a series of p-proofs of the same theorem.

This long disagreement shows how the lack of rigor of the Italian school of algebraic geometry led its members astray. It also exhibits the instability of this mathematical community. Arguably, *speculative methods* have an important place in the growth of mathematics, but they should be openly labeled as such. We shall see that, in this respect, Severi's and Enriques's attitudes greatly diverge (Sect. 5). But first, we are going to discuss what are the conditions that make recalcitrant PP-DISAGREEMENTS possible. To do so, we are going to start discussing (i) the links between formalization, rigor, and correctness (Sect. 4.1), (ii) the difference between a general criterion of rigor for p-proofs and specific working criteria of acceptability for rigorous p-proofs – and how these can be used to characterize *well-functioning* mathematical communities (Sect. 4.2). Afterward, we turn to (iii) the nature of PP-DISAGREEMENT within the Italian school of algebraic geometry (Sect. 4.3).

4.1 Formalization, Rigor, and Correctness

4.1.1 Formalization and Consensus

In order to understand what went wrong in the Italian school of algebraic geometry, it will be useful to have on the table a picture of why things generally go right. Why do contemporary mathematicians tend to agree so much about the correctness of their p-proofs? According to Wagner (2017, 67),

> The consensus among mathematicians about the validity of [putative] proofs has a lot to do with formalization. By formalization I do not mean the translation of an entire proof into a strictly formal language (which is almost never done, and is in fact impossible for finite humans to achieve in the context of typical research mathematics). By formalization I mean a gradual process of piecemeal approximation of formality that is conducted only as far as required to resolve a given dispute.

A problem with linking the phenomenon of mathematical consensus over the correctness of p-proofs with formalization is that even the local (and gradual) formalization is often unavailable in practice.[22] Another problem is that it is not at all clear what *formalization* means at all: Are we referring to a specific formal system? Do we intend potential or actual formalization? How are the steps of the informal p-proof related to the ones of its formal counterparts?[23]

We do not aim to answer these questions. That would take us too far afield from our present concern. Let us just briefly mention how Wagner tackles these issues in a later work (2022). He does so by way of analogy: he suggests thinking about formalizations along the lines of Courts of Appeal in the juridical system. This analogy, says Wagner, can help us make sense of how formalizations can be used to explain the absence of recalcitrant PP-DISAGREEMENT even if they are seldom appealed to in the practice of mathematics (and even if the very term *formalization* is vague). But what are Courts of Appeal exactly? The U.S. government describes them as follows:

> The U.S. Courts of Appeal hear appeals from lower courts of both civil and criminal trials, but do not investigate the facts of a case. Rather, the Appeals Courts investigate whether or not the law has been fairly and correctly applied by the lower courts.[24]

In the U.S., a court of appeal is then invoked when there are doubts about the correctness of the juridical proceedings of a lower court. The rough idea is that, in the same

[22] Note that the situation might change radically due to technological innovations related to new computer proof assistants. These tools are making the formalization of mathematics more and more manageable (Avigad 2018).

[23] These and related questions are addressed in (Burgess and De Toffoli 2022).

[24] This definition is taken from the governmental website: https://www.usa.gov/federal-agencies/u-s-courts-of-appeal.

vein, a certain level of formalization is invoked when there are doubts arising about the correctness of a less formal p-proof.

In order to better appreciate Wagner's analogy, here are the main five features he focuses on.[25] (a) Appeals Courts are only seldom involved; similarly, formalizations are not a common resort to resolve disagreements (they are just appealed to in special cases). (b) There is a variety of both formalizations and Appeals Courts (which are usually arranged hierarchically). (c) The issues to be resolved with the aid of Appeals Courts and formalization are local rather than global (we do not generally formalize the whole proof; likewise, the court of appeal will only be dealing with specific issues of a juridical case). (d) There is a hierarchy of authority: proofs that are more formal are (generally) more authoritative over less formal ones. Likewise higher courts of appeal (e.g., the Supreme Court) will be more authoritative over lower courts of appeal. (e) The Supreme Court does not have the final word – the law can change, etc. Similarly, complete formalization will have to be accepted by a community that can decide, in some instances, to reject it.

To be sure, this analogy has its limits, and it is rather impressionistic – it is meant to be this way. However, we agree that it is still helpful to provide a schema of how formalization can play a role in generating consensus, especially in light of the fact that formalization remains a rare practice among mainstream mathematicians. In our view, the key observation is that the focus should be on the very possibility of appealing to a higher court, or a greater formality, rather than actually appealing to it. That is, what matters most to mathematicians is *potential* rather than actual formalization.

4.1.2 Rigor and Formalization

The conceptual innovations that made formalizations possible were not available much before the turn of the twentieth century – and in fact, before that time, recalcitrant PP-DISAGREEMENT was a matter of routine:

> If we date a qualitative increase in consensus concerning the validity of mathematical proofs to somewhere around the turn of the twentieth century (give or take three or four decades), we should find what it is that changed in mathematical practice, which could account for the emergence of consensus. The obvious suspect is clear: mathematical formalization. (Wagner 2022, 11–12)

According to Wagner, formalization is what changed. However, what changed is also often called the *rigorization of mathematics* – a broad phenomenon that started well before and gradually invested all areas of mathematics. Rigor and (partial and piecemeal) formalization are indeed linked. According to a widely held view in the philosophy of mathematics, the very possibility of formalization of a p-proof is tied to its rigor.[26] More precisely, a rigorous p-proof can be formalized (if it is correct) because it has a certain level of precision that makes it possible (for the relevant practitioners) to gradually add more and more details. Although there are many variants

[25] We report them here briefly, but the interested reader should consult (Wagner 2022).

[26] See (Avigad 2021).

of the definition of rigor to choose from in the literature, many invoke some sort of formalization.[27]

Rather than proposing an original view on the matter, we only need here a working conception of rigor that can help us in the analysis of our case study. Note that the question of mathematical rigor is a hard and controversial one, especially if we want to analyze the history of mathematics. For our limited scope, however, it will do to work with a somewhat simplified story – one that aims to model only mathematics from the turn of the twentieth century. In this limited context, it is tenable to associate formalization with rigor. The basic idea is that a genuine proof is rigorous if it can be *in principle* formalized in an acceptable formal system or another.[28] In the quote above, Wagner writes: "[b]y formalization I mean a gradual process of piecemeal approximation of formality that is conducted only as far as required to resolve a given dispute" – our working assumption is that it is because proofs are rigorous that such process can take off at all.

Taking inspiration from (Burgess 2015), we can elucidate this in principle formalization as follows: *a p-proof is rigorous if it would convince an idealized mathematician with the right background of the existence of a formal proof*. A couple of remarks are in order. The idealization at play should neither abstract away from our human limitations in terms of computational and memory thresholds nor from our human fallibility. It should, however, disregard individual differences. Our idealized mathematician would have to possess the right background knowledge because proofs are generally intended for a trained audience. Moreover, proofs should convince by virtue of the correctness of their inferential steps (and not, say, because they are published by a famous mathematician). This implies that rigor is a good indicator of formalizability – it is not a perfect indicator of formalizability because the idealization at play does not abstract away from our fallibility.[29]

In the view on offer, a rigorous p-proof is one that can reliably be formalized. That is, most rigorous p-proofs can be formalized. According to some authors, *all* rigorous p-proofs can be formalized. We prefer, however, to adopt a looser notion according to which *rigor is formalizability-conducive rather than formalizability-entailing*.

It is noteworthy that the very two protagonists of our dispute have themselves written about mathematical rigor. They distinguish between two types of rigor. Severi's writes of *substantial* and *formal* rigor, while Enriques writes about *large-scale logic* [logica in grande] and *small-scale logic* [logica in piccolo].[30] In short, substantial rigor (and large-scale logic) has to do with a faithful description of the mathematical facts – that is, with the truth of the results. Formal rigor (small-scale logic) has instead to do with formalization, that is, with the thoroughness of the p-proof supporting a result and the correctness of its inferential steps. It is this last type of rigor

[27] For a pluralist conception of rigor, see (Tanswell forthcoming).

[28] There are important critiques to this view. Here are two objections that we find most relevant: (i) not all formal systems will do (a silly example of what won't do is an inconsistent system), (ii) the overgeneration problem discussed in (Tanswell 2015): to a single p-proof, many formal proofs can be associated.

[29] Note that this view is inspired by Burgess's, but departs from his view in this respect.

[30] For a discussion of these two types of rigor and their connection with different notions of objectivity in mathematics, see (De Toffoli and Fontanari 2022).

that is at issue here. Enriques and Severi's explicit discussion on the topic shows that (formal) rigor was a norm they were aware of, one that was not universally implemented but one that governed many mathematical practices of the time.

4.1.3 Correctness and Formalization

But why does this all matter? That is, why does it matter to link rigor with formalization? Well, it matters because, in our contemporary conception (as opposed to earlier historical conceptions), the correctness of a deductive argument is cashed out precisely in terms of formalization: "It has been common since the turn of the twentieth century to take correctness to be underwritten by the existence of formal derivations in a suitable axiomatic foundation" (Avigad 2021, 2377) – we will use the term correctness in this technical way from now on.

Correctness is then a criterion external to single mathematical communities but internal to the broad mathematical community after the turn of the century. After that time, correctness was equated with formalizability in an appropriate formal system. Here is Avigad again:

> According to the standard view, a mathematical statement is a theorem if and only if there is a formal derivation of that statement, or, more precisely, a suitable formal rendering thereof. When a mathematical referee certifies a mathematical result, then, whether or not the referee recognizes it, the correctness of the judgement stands or falls with the existence of such a formal derivation. (Avigad 2021, 7381)

Let us take stock. We started by discussing Wagner's appeal to formalization to explain the extreme rarity of recalcitrant PP-DISAGREEMENT in recent mathematics (from about the turn of the century). We then characterized rigor as a formalizability-conducive property of p-proofs. Lastly, we explained how the correctness of p-proofs can be fleshed out in terms of formalizability. Putting two and two together, we get that rigor, if it does not entail correctness, at least it is correctness-conducive. And this is one of the reasons why rigor is a desirable quality of p-proofs – another being that rigorous p-proofs can be shared among appropriately trained practitioners. This, in turn, implies that rigorous p-proofs can undergo a stricter verification process compared to idiosyncratic arguments that only a single mathematician can understand.

4.2 Acceptability Criteria and Well-Functioning Mathematical Practices

So far, so good. But the story is not so simple. As pointed out in (De Toffoli 2021b), the *criterion of rigor* (which, as we saw, is a general criterion) is different from the *criteria of acceptability for rigorous proofs* (i.e., the criteria at play in a specific context that determine which p-proofs are accepted as rigorous and which aren't). The problem is that while the former is correctness-conducive, the latter might not be.[31]

Acceptability criteria for rigorous proofs are the criteria that determine *in practice* what a community accepts as rigorous proof. These have to do with the specifics of a particular mathematical context. For instance, a good indication for such criteria is given by what is actually published in specialistic journals in a specific context. These criteria change considerably not only across time but also across sub-fields. For example, in topology, it is an acceptable practice to appeal to geometric intuition. This is not the case in other fields.

We can use these two notions to characterize *well-functioning mathematical practices*. These are practices in which the criterion of acceptability is aligned with the criterion for rigor. That is, acceptable proofs are reliably rigorous – in other words, acceptability is rigor-conducive. And since rigor is correctness-conducive, so is acceptability. It is for this reason that these are well-functioning mathematical practices -- because the p-proofs they accept tend to be correct (or formalizable).

Unfortunately, however, this alignment is not always in place. There are indeed mathematical practices in which the criteria of acceptability come apart from the criterion of rigor:

> The criteria of acceptability for (rigorous) proofs are supposed to track correctness; however, unlike rigor, they are not a guarantee. There are clear historical cases in which the criteria of acceptability for proofs, even in professional contexts, were shown to be inadequate and required revision. (De Toffoli 2021b, 1788)

Our controversy is one case of such a historical example.[32]

Well-functioning mathematical practices are mathematical practices in which the criterion of rigor and that of acceptability are aligned. They are stable, successful practices, linked to other mathematical practices. Moreover, they are generally not judged to be faulty by other publicly recognized social practices. Undoubtedly, this characterization does not allow us to draw sharp boundaries. For one, a temporal dimension is implicit. How should we temporarily delimit mathematical practices? A practice could transition from a period in which the criteria of rigor and acceptability are aligned to one in which they are not. Moreover, it might not always be possible to claim with certainty whether a mathematical practice is well-functioning or not.

By way of example, consider Shinichi Mochizuki's Japanese community of mathematicians. In 2012 Mochizuki proposed a p-proof of a famous number-theoretic conjecture: the *abc* Conjecture. His methods were new and relied on a vast amount

[31] The distinction between rigor and acceptability for p-proofs parallels Miranda Fricker's (1998) distinction between *indicator properties* and *working indicator properties* for good informants, which she develops in her investigation of the phenomenon of *epistemic injustice*. According to Fricker, a good informant is both competent and trustworthy – this is what constitutes *rational authority*. The indicator properties are properties that reliably indicate rational authority. Instead, working indicator properties "are those properties actually used in a given practice to indicate rational authority, and which may or may not be so reliable" (ibid. 168).

[32] For a contemporary case, one might point to symplectic geometry (Hartnett 2017).

of material he developed. The p-proof was accepted by other mathematicians, mostly Japanese, but with notable exceptions.[33] It was clear that checking the material was no easy task – especially for people not already acquainted with Mochizuki's previous work. In 2018, German mathematicians Peter Scholze and Jacob Stix proposed a simplification of Mochizuki's approach and challenged a particular Lemma. They were not given any satisfactory answer, and the disagreement reached an impasse. Mochizuki claimed that Scholze and Stix's failure to understand the Lemma derived from their lack of background knowledge and from the fact that their simplification was not faithful to his work. The German mathematicians disagreed.[34]

This is a clear case in which, although a specific problem was identified, there is no agreement on how to get over it. Although, at this time, there is evidence that a certain mathematical practice is problematic, the possibility of the disagreement persisting due to a lack of shared background knowledge is still open. Note that in this case, like in other cases of recalcitrant PP-DISAGREEMENT over a single p-proof, actual formalizations are of no help because they are out of reach since a lot of preliminary work would have to be done to even start the process of formalizing. What is more, as Aberdein (2023) observes, such preliminary work would *presuppose* rather than lead to a solution to the disagreement. This is because the said disagreement is about whether Scholze and Stix's simplification of Mochizuki's approach is faithful to his original techniques in the relevant respects.[35] More generally, formalization might also not resolve debates about the validity of p-proofs because there might still be disagreement about whether the validity of any given purported formalization also substantiates the validity of the proof it is supposed to formalize.

This example points to the fact that formalization raises delicate issues concerning the individuation of proofs (i.e., whether one proof is the same as one of its formal counterparts). Another upshot is that we might not be able, at a time, to determine whether a mathematical practice is well-functioning or not. However, it might be possible to do that with hindsight. Let us then go back to our case and evaluate what the problem was with Enriques's and Severi's mathematical community.

4.3 Recalcitrant PP-disagreement in the Italian School

We suggested that the very possibility of recalcitrant PP-DISAGREEMENT over p-proofs of the same theorem is rare or even absent in well-functioning mathematical practices. This is because these are stable practices in which acceptable p-proofs tend to be rigorous and thus only very seldom fail to be correct.

Recalcitrant pp-disagreement is characteristic of unstable mathematical practices. They can be compared to Kuhnian revolutionary phases, in which the accepted methods and the concepts deployed by scholars are fluid. Precisely because of their insta-

[33] One such exception being Fesenko (2019).

[34] See (Klarreich 2018) for a description of the case and (Aberdein 2023) for an insightful philosophical analysis.

[35] This is related to the problem of how we should identify proofs, which is related to CONCEPTION OF PROOF-DISAGREEMENT. See Footnote 8. Thanks to one of the anonymous referees for this clarification.

bility, mathematical communities going through such phases are *not* well-functioning (to use Kuhn's terminology, they are not part of "normal science"). This analogy has its limits. By mentioning Kuhn, we do not want to imply that different mathematical paradigms are incommensurable. Remember that we are focusing on relatively recent mathematics, and that we suggested that the standard of correctness cashed out in terms of formalizability is generally accepted. For example, Mochizuki's claim that his p-proof is correct does in indeed imply that it is formalizable – the same holds for Enriques's or Severi's p-proofs. This indicates that although there might not be an external standard of correctness for all mathematics (from ancient to contemporary times),[36] there is such a standard for recent mainstream mathematics – one that is external to single mathematical practices.

The problem with the Italian school of algebraic geometry is that, in at least some of its phases, the criteria of acceptability and of rigor were not aligned. Our case study shows clearly that different incorrect p-proofs were accepted as rigorous. Some of the accepted results were false. Others were true but not properly justified by rigorous p-proofs. As a matter of fact, the PP-DISAGREEMENT over the various p-proofs proposed was due either to the falsity of their results or to their incorrectness and lack of rigor, in particular to the presence of unjustified assumptions. This leads us to another distinction.

Besides distinguishing between evanescent and recalcitrant PP-DISAGREEMENT over one or multiple p-proofs, we can also differentiate between *global* and *local* PP-DISAGREEMENT over a single p-proof: that is, arising from a general problem with the result (such as the presence of a counterexample) or a specific problem with the p-proof (such as an incorrect step). In turn, we can differentiate among the latter between local PP-DISAGREEMENT over a p-proof *that derives from errors* and local PP-DISAGREEMENT *that derives from gaps* – errors and gaps are thus different ways in which a p-proof can be found to be incorrect. Schematically:

Global PP-DISAGREEMENT mostly arises from the presence of counterexamples. If the convertibility norm is in place, it should be possible to convert it to local disagreement. Local PP-DISAGREEMENT that arises from errors tends to be evanescent since, when an error is spotted, it is generally easy to convince everyone of its presence. It is not so simple with gaps. This is because they constitute more elusive deficiencies of p-proofs. When is a gap acceptable? It is not an easy question to answer.[37] Moreover, any answer will vary with the context. One thought is that disagreement over the presence of large gaps can become a global type of disagreement, violating the convertibility norm. This is an indication that the mathematical practice at issue is not well-functioning.

Let us return to our case study. We can identify three different phases of the Italian school of algebraic geometry. The first phase started when the Italian school began, around 1860. It lasted until the end of the 1920s. This phase was characterized by heavy use of geometric intuition and methods that only later were shown to be unrigorous. It is well epitomized by the words pronounced by Castelnuovo (1928, 201) in 1928:

[36] We do not want to imply there is not but simply that our view does not entail there is one.

[37] See (Fallis 2003).

> renouncing geometric intuition, the only thing that has allowed us to orient ourselves in this intricate territory so far, would mean extinguishing the tenuous flame that can guide us in the dark forest [vorrebbe dire spegnere la tenue fiammella che può guidarci nell'oscura foresta].[38]

The second phase spanned the 1930s and the 1940s. It is here that disagreement started to arise. This was a period in which disagreement about both gaps and actual errors was common. Doubts over the rigor of the methods used by the members of the Italian school were emerging. The members themselves began to realize that their results were wanting. It was in this phase that the Italian school lost its apparent status as a well-functioning mathematical practice since the working criterion of acceptability started to diverge from the one of rigor. By way of example, consider again the words that Enriques addressed to Segre in 1945:

> I have reexamined my earlier proof based on infinitely close curves of various orders and I believe it is substantially right, even if it is not rigorously complete. (Babbitt and Goodstein 2011)

The third phase is the one in which the methods of the Italian school were rejected. The rigorization of algebraic geometry exploded in the 1950s with the explicit, sharp criticism raised towards the foundations and methods of the Italian school by the bold younger exponents of the French school. Just to mention one well-known episode at the International Congress of Mathematicians held in Amsterdam in 1954, Severi was not even able to conclude his talk due to the repeated argumentative interruptions by Pierre Samuel and André Weil.[39] It was during this phase that Mumford's *Lectures on Curves on an Algebraic Surface* were published.

Our case of recalcitrant PP-DISAGREEMENT spans all three phases of the school. The first p-proofs proposed by Enriques and Severi from the first years of the twentieth century are ultimately incorrect because they make use of an unjustified assumption. However, they were accepted for about fifteen years before Severi unveiled such an unjustified assumption in 1921. Afterward, in the 1940s, Severi proposed a p-proof that contained an irreparable error, while Enriques's final attempt at proving the Fundamental Theorem was on the right track but lacked precision and thus contained a significant gap.

5 Rigorous and Speculative Mathematics

We now turn to the analysis of Severi's and Enriques's different attitudes concerning the role of speculative, non-rigorous mathematics. We connect our discussion to the so-called "theoretical mathematics" debate initiated by Arthur Jaffe and Frank Quinn (1993) with an article in the Bulletin of the American Mathematical Society and protracted by a series of replies by various mathematicians (Atiyah et al. 1994).

[38] Our translation from Italian.

[39] See (Ciliberto and Sallent Del Colombo 2018, 15).

5.1 Severi's Loss of Rigor

Ciro Ciliberto and Emma Sallent Del Colombo (2018, 2) point out that Severi's intellectual outlook changed radically with the great war. While in the pre-Bellum period Severi's mathematical texts were careful and rigorous, in the postwar years, he developed an "autoreferential trust" that led him to endorse a speculative style:

> His scientific production, compared to the pre-war period, suffers from an excessive self-referential trust that leads him to less attention to rigor and therefore to sometimes erroneous results.[40]

The problem is that what Severi kept calling *proofs* failed to be rigorous:

> In fact, in the post-war period, Severi, while not losing his great inspiration, his originality of views, and producing real pearls here and there, has more *conjectured* (to use a euphemism that hides various statements of dubious exactness) than *demonstrated*. […] These works are characterized by a growing "self-confidence," which often led Severi to formulate unjustified assertions, taking them as established facts, and theories developed on them collapsed as houses of cards in the face of harsh criticism.[41]

It is because of his high position in Italian mathematics, first as the only mathematician member of the *Accademia d'Italia*, and then after 1939, as the founder and lifetime president of the *Istituto Nazionale di Alta Matematica*, that he could impose his results over the community. After having gained a reputation as a brilliant and rigorous mathematician, he managed to publish and teach his dubious (and often outright wrong) results based on incorrect p-proofs. This is a dangerous case of abuse of epistemic authority because it can hinder or even block the process of self-check that mathematical practices tend to implement on themselves and that is essential for their stability and success.

A typical example of Severi's autoreferential trust is provided by the lithographed notes of Severi's (1948) lecture course on algebraic geometry held at the Scuola Normale Superiore in Pisa in February 1947. After recalling the unlucky history of the theorem of completeness of the characteristic series in a footnote to § 49 (*ibid.*, 122: "many attempts were made in Italy and abroad […], but they did not have a happy outcome"), it is triumphantly asserted (ibid., 123):

> We will add that in April 1947, when these lessons were over, Severi was able, 43 years after he had first enunciated the theorem, to prove it with only algebraic-geometric means for semiregular curves.[42]

[40] Our translation from Italian.

[41] Our translation from Italian.

[42] Our translation from Italian.

It is reasonable to suspect that this claimed algebraic-geometric p-proof was the one to be later reproduced in the last treatise by Severi (1958). Unfortunately, such a p-proof is far from being unconditional, as Severi himself admitted (Severi 1958, 197), and turns out to be essentially just a plausible argument:

> [W]e established in n. 23 the remarkable 'breaking principle' by B. Segre contained in the Memory of the Annals of Mathematics, 1938, but under the hypothesis, added here, that a component of the limit of the curve that tends to break is an ordinary irreducible curve […]. As this principle only considers the geometric kind of the surface and not the irregularity, [...] it seems natural to suppose that the principle itself is always true.[43]

It is plausible to think that it was Severi's own attitude that made the controversy possible. As we saw, still in 1945, Enriques wrote: "His [Severi's] exposition seemed obscure to me and therefore dubious." The main problem is that Severi failed to separate rigorous mathematics from speculative mathematics.

In their much-discussed, polemical paper on what they called "theoretical mathematics" (meaning non-rigorous or speculative mathematics), Jaffe and Quinn (1993, 10) argued that speculative mathematics plays a vital role in mathematical practice but should always leave space for a more rigorous counterpart:

> Theoretical work should be explicitly acknowledged as theoretical and incomplete; in particular, a major share of credit for the final result must be reserved for the rigorous work that validates it.

Their critique was leveled at those who were practicing speculative mathematics but presented it as though it were rigorous mathematics:

> the failure to distinguish carefully between the two can cause damage both to the community of mathematics and to the mathematics literature. *(ibid.*, 12)

Jaffe and Quinn incited replies from many mathematicians who, more or less explicitly, fell within the paper's target. It is relevant for us to note that the case of the Italian school was mentioned by the authors as one of their "cautionary tales."

5.2 Enriques the Visionary

Enriques's attitude was very different from Severi's. Castelnuovo, who published Enriques's (1949) posthumous monograph after his sudden death in 1946, identified a prudential quality in Enriques's approach:

> The Author himself takes care to warn right from the preface that the treatise, rather than expounding an already static and crystallized doctrine, aspires to arouse in the reader the desire to bring additions and improvements to various

[43] Our translation from Italian.

> theories. And where the ground is less solid, the Author warns the scholar. [...] A brilliant intuition led Enriques in 1904 to enunciate and establish this characteristic property of irregular surfaces; subsequent fundamental research relied on it. However, careful criticism has shown, several years later, that Enriques's demonstration was not satisfactory. The property is true, at least under some restrictions, as shown by Henri Poincaré's transcendent research. But all the attempts made later by Enriques and others to demonstrate it through algebraic-geometric considerations have run into difficulties hitherto unsurpassed. This is explicitly stated in the aforementioned Chapter IX, in which the Author also gives suggestions on the way to try to reach the goal.[44] (Castelnuovo, in: Enriques 1949, vi)

As pointed out in (Fontanari 2023), this was precisely the intellectual habit of Castelnuovo: a mixture of free open-mindedness and strict moral rigor,[45] typical of his personal attitude towards both mathematical research and teaching. For instance, in a footnote to his paper, he claimed:

> On the other hand, we are going to apply a still unproven principle in order to solve a difficult problem since we believe that such attempts may be useful to the progress of science, provided one explicitly declares what is admitted and what is proven.[46] (Castelnuovo 1889, 130)

Certainly, Enriques was bolder than Castelnuovo but remained able to avoid Severi's worse missteps. The difference between Enriques's and Severi's style of mathematical speculation is clear in the already quoted letter to Segre, where Enriques observes that Severi "referred to the case of algebroid entities instead of clarifying the matter in the easiest case." Indeed, Enriques's taste for concrete examples rather than abstract generalizations guided his intuition to safer paths. For a complete rescue of his reasons, however, Enriques had to wait for an unexpected recantation, dictated to Mumford (2011, 260) by his intellectual honesty and published in 2011:

> In my own education, I had assumed they were irrevocably stuck, and it was not until I learned of Grothendieck's theory of schemes and his strong existence theorems for the Picard scheme that I saw that a purely algebraic-geometric proof was indeed possible. […] As I see it now, Enriques must be credited with a nearly complete geometric proof using, as did Grothendieck, higher-order infinitesimal deformations. In other words, he anticipated Grothendieck in understanding that the key to unlocking the Fundamental Theorem was understanding and manipulating geometrically higher order deformations. *Let's be careful: he certainly had the correct ideas about infinitesimal geometry, though he had no idea at all how to make precise definitions.* If you compare his ideas here with, for example, the way Leibniz described his calculus, the level of

[44] Our translation from Italian.

[45] These can be numbered among "mathematical virtues" (Aberdein et al. 2021)

[46] Our translation from Italian.

> rigor is about the same. To use a fashionable word, his "yoga" of infinitesimal neighborhoods was correct, but basic parts of it needed some nontrivial algebra before they could ever be made into a proper mathematical theory. (...) In short, *Enriques was a visionary. And, remarkably, his intuitions never seemed to fail him (unlike those of Severi, whose extrapolations of known theories were sometimes quite wrong).* Mathematics needs such people[.] (*emphasis added*)

6 Conclusion

In this paper, we have shown that controversies in mathematics are not limited to the choice of axioms. Indeed, we distinguished six species falling under the genus of *mathematical disagreement*:

1) AXIOMS-DISAGREEMENTS
2) PHILOSOPHICAL-DISAGREEMENT
3) SOCIAL PRACTICE-DISAGREEMENTS
4) HEURISTIC-DISAGREEMENTS
5) CONCEPTION OF PROOF-DISAGREEMENT
6) PP-DISAGREEMENT

We then focused on PP-DISAGREEMENT, namely, on disagreement over the correctness of putative proofs. When they arise in well-functioning mathematical communities, these disputes are quickly resolved. For instance, an error was spotted and recognized as such by all parties in the case of Kempe's p-proof of the 4-color Conjecture or Wiles's initial p-proof of Fermat's Last Theorem.[47] There are, however, also cases of *recalcitrant* PP-DISAGREEMENTS that resist swift resolutions. They can arise from persistent disagreement over a single p-proof (as in the contemporary case involving Japanese mathematician Mochizuki) or over the correctness of a series of p-proofs (as in our case study and in Lakatos's historical case). Recalcitrant PP-DISAGREEMENTS tend to arise within mathematical communities lacking the resources for solving controversies arising from their inferential practices. One such community was the Italian school of algebraic geometry that straddled the nineteenth and twentieth centuries. We appealed to the gap between an abstract criterion of rigor and local criteria of acceptability to distinguish communities of this kind from well-functioning mathematical communities.

The Italian school of algebraic geometry is a prime example of a mathematical community lacking the resources to satisfactorily solve PP-DISAGREEMENTS. Although the disputes over single erroneous p-proofs did not themselves give origin to long disagreements, several incorrect p-proof of the same result were at different times accepted by the community. This reveals that the local criteria of acceptability of the Italian school did not align with an abstract criterion for rigor.

[47] It was then fixed by Andrew Wiles in about one year with the help of his former student Richard Taylor.

The decades-long controversy between Enriques and Severi, two of the most prominent members of the Italian school of algebraic geometry, is especially significant because it did not develop in far times but in the twentieth century and was not over some marginal result but over a central theorem. Some of the methods accepted among the members of the Italian school were speculative. This, we suggested, is not bad per se, but it can lead to negative results if those speculative methods are presented as rigorous, as in Severi's case. Serious predicaments arise when non-rigorous mathematics is disguised as rigorous mathematics and accepted as such.

Acknowledgements We wish to thank the two anonymous referees for their detailed comments that helped us clarify our positions. Thanks are also due to Fenner Tanswell for insightful suggestions. A recent draft was discussed during a reading group at IUSS Pavia – we are particularly grateful to Andrea Sereni and Guido Tana for their feedback. This research project was partially supported by GNSAGA of INdAM and by PRIN 2017 "Moduli Theory and Birational Classification."

References

Aberdein A (2023) "Deep Disagreement in Mathematics." Global Philosophy 33(17):1–27

Aberdein A, Rittberg C, Tanswell F (2021) "Virtue theory of mathematical practices: an introduction." Synthese 199(3–4):10167–10180

Atiyah M, Borel A, Chaitin GJ, Friedan D, Glimm J, Gray JJ, Hirsch MW, Mac Lane S, Mandelbrot B, and RuelleD (1994) "Responses to: A. Jaffe and F. Quinn, 'Theoretical mathematics: toward a cultural synthesis of mathematics and theoretical physics'." Bulletin of the American Mathematical Society 30 (2): 178–207

Avigad J (2018) "The Mechanization of Mathematics." Notices of the American Mathematical Society 65(6):681–690

Avigad J (2021) "Reliability of mathematical inference." Synthese 198:7377–7399

Azzouni J (2006) "How and Why Mathematics Is Unique as a Social Practice." In 18 Unconventional Essays on the Nature of Mathematics, edited by Reuben Hersh, 201–219. Springer

Babbitt D, Goodstein J (2011) "Federigo Enriques's Quest to prove the 'Completeness theorem.'" Notices of the American Mathematical Society 58:240–249

Barany M (2018) "The Fields Medal should return to its roots." Nature 553:271–273

Bell J, Hellman G (2006) "Pluralism and the Foundations of Mathematics." In Scientific Pluralism, edited by Kenneth Waters, Herbert Feigl, Stephen H. Kellert and Helen Longino, 64–79. University of Minnesota Press

Brigaglia A, Ciliberto C (1998) "Geometria Algebrica." In La matematica italiana dopo l'unità. Gli anni tra le due guerre mondiali, edited by Simonetta di Sieno, Angelo Guerraggio and Pietro Nastasi, 185–320. Marcos y Marcos

Burgess J (2015) Rigor and structure. Oxford University Press

Burgess J, De Toffoli S (2022) "What Is Mathematical Rigor?" Aphex 25:1–17

Castelnuovo G (1889) "Numero delle involuzioni razionali giacenti sopra una curva di dato genere." Rendiconti R. Accademia dei Lincei 5(4)

Castelnuovo G (1928) "La geometria algebrica e la scuola italiana." Congresso Internazionale dei Matematici: 3–10 settembre 1928, 1928

Castelnuovo G (1930) "Luigi Cremona nel centenario della nascita." Rendiconti R. Accademia dei Lincei 6(12):613–618

Christensen D (2009) "Disagreement as evidence: the epistemology of controversy." Philosophical Compass 4(5):756–767

Ciliberto C, Sallent Del Colombo E (2018) "Francesco Severi: il suo pensiero matematico e politico prima e dopo la Grande Guerra." Preprint, ArXiv: 1807.05769

Clarke-Doane J (2020) Morality and Mathematics. Oxford University Press

De Toffoli S (2021a) "Groundwork for a fallibilist account of mathematics." Philosophical Quarterly 71(4):823–844

De Toffoli S (2021b) "Reconciling Rigor and Intuition." Erkenntnis 86:1783–1802
De Toffoli S, Fontanari C (2022) "Objectivity and Rigor in Classical Italian Algebraic Geometry" Noesis 38:195–212
De Toffoli S (2022) "Intersubjective Propositional Justification." In Propositional and Doxastic Justification: New Essays on their Nature and Significance, edited by Paul Silva Jr and Luis R. G., Oliveira, 241–262. Routledge
De Toffoli, S (2023) "Who's Afraid of Mathematical Diagrams?" Philosopher's Imprint 23(1): 1–20
Easwaran K (2015) "Rebutting and undercutting in mathematics." Philosophical Perspectives 29(1):146–162
Elga A (2007) "Reflection and Disagreement." Nous 41(3):478–502
Enriques F (1942) "Sui sistemi continui di curve appartenenti ad una superficie algebrica." Commentarii Mathematici Helvetici 15:227–237
Enriques, F (1949) Le superfici algebriche. Zanichelli
Fallis D (2003) "Intentional Gaps in Mathematical Proofs." Synthese 134:45–69
Fesenko I (2019) "About certain aspects of the study and dissemination of Shinichi Mochizuki's IUT Theory." Manuscript
Fontanari C (2023) "Guido Castelnuovo and his heritage: geometry, combinatorics, teaching." In Algebraic Geometry between Tradition and Future, edited by G. Bini. Springer
Fricker M (1998) "Rational authority and social power: Towards a truly social epistemology." Proceedings of the Aristotelian Society 98 (2): 159–177
Gödel K (1964) "What is Cantor's Continuum Problem?" In Philosophy of Mathematics: Selected Readings, edited by Paul Benacerraf and Hilary Putnam, 258–273. Englewood Cliffs: Prentice-Hall
Goldstein C (2013) "Routine controversies: Mathematical Challenges in Mersenne's correspondence." Revue d'histoire des sciences 66(2):249–273
Hartnett K (2017) "A fight to fix geometry's foundations." Quanta Magazine
Jaffe, A., Quinn, F. (1993) "'Theoretical mathematics': Toward a cultural synthesis of mathematics and theoretical physics." Bulletin of the American Mathematical Society 29(1):1-13
Kelly T (2010) "Peer disagreement and higher order evidence." In Social Epistemology: Essential Readings, edited by Alvin I. Goldman and Dennis Whitcomb, 183–217. Oxford University Press
Klarreich E (2018) "Titans of Mathematics Clash over Epic Proof of ABC Conjecture." Quanta Magazine
Lakatos I (1976) Proofs and refutations. Cambridge University Press
Maddy P (2011) Defending the Axioms: on the philosophical foundations of set theory. Oxford University Press
Mumford D (1966) Lectures on curves on an algebraic surface. Princeton University Press
Mumford D (2011) "Intuition and rigor and Enriques's quest." Notices of the American Mathematical Society 58:250–260
Nasar S, Gruber D (2006) "The Clash Over the Poincaré Conjecture" The New Yorker
Paseau A (2015) "Knowledge of Mathematics without Proof." The British Journal for the Philosophy of Science 66(4):775–799
Paseau A (2016) "What's the Point of Complete Rigour?" Mind 125(497):177–207
Poincaré H (1889) "La logique et l'intuition." L'enseignement mathématique 1(5):157–162
Rittberg C, Tanswell F, Van Bendegem JP (2020) "Epistemic injustice in mathematics." Synthese 197(9):3875–3904
Segre B (1938) "Un teorema fondamentale della geometria sulle superficie algebriche ed il principio di spezzamento." Annali di Matematica Pura ed Applicata 17(1):107–126
Severi F (1942) "Intorno ai sistemi continui di curve sopra una superficie algebrica." Commentarii Mathematici Helvetici 15:238–248
Severi F (1944) "Sul teorema fondamentale dei sistemi continui di curve sopra una superficie algebrica." Annali di Matematica 23:149–181
Severi F (1948) Fondamenti di geometria algebrica. CADAM
Severi F (1958) Geometria dei sistemi algebrici sopra una superficie e sopra una varietà algebrica. Edizioni Cremonese
Tanswell F (2015) "A problem with the dependence of Informal Proofs on formal proofs." Philosophia Mathematica 23(3):295–310
Tanswell F (forthcoming) Mathematical Rigour and Informal Proof. Cambridge Elements: Cambridge University Press
Tymoczko T (1979) "The four-color problem and its philosophical significance." The Journal of Philosophy 76(2):57–83

Voevodsky V (2014) “The Origins and Motivations of Univalent Foundations.” The Institute Letter (The Institute for Advanced Studies)
Wagner R (2017) Making and breaking Mathematical sense: histories and philosophies of Mathematical Practice. Princeton University Press
Wagner R (2022) “Mathematical consensus: a research program.” Axiomathes 32(3):1185–1204
Zariski O (1935) Algebraic surfaces. Second Supplemented Edition ed. Springer