Abstract

Throughout history, mathematicians have expressed preference for solutions to problems that avoid introducing concepts that are in one sense or another “foreign” or “alien” to the problem under investigation. This preference for “purity” (which German writers commonly referred to as “methoden Reinheit”) has taken various forms. It has also been persistent. This notwithstanding, it has not been analyzed at even a basic philosophical level. In this paper we give a basic analysis of one conception of purity—what we call topological purity—and discuss its epistemological significance.

“. . . arithmetica arithmetice, geometrica geometrice doceantur”

—Ramus

1. Introduction

There is a long tradition in mathematics of preferring the pure to the impure in proofs of theorems and solutions of problems. “Purity” in mathematics has generally been taken to signify a preferred relationship between the resources used to prove a theorem or solve a problem and the resources used or needed to understand or comprehend that theorem or problem. In this sense, a pure proof or solution is one which uses only such means as are in some sense intrinsic to (a proper understanding of) a theorem proved or a problem solved.

Neither the supposed character nor the presumed value of this intrinsicality is very clear, though. Nor does the history of mathematics suggest predominant views concerning them. All that is clear is that there has been a widespread, persistent tendency to think that there are ways in which the resources used in a proof/solution can match or fail to match the theorem proved/problem solved, and that when proper match is achieved it in some way(s) adds to the value of the proof/solution produced.

This in itself is reason enough to seek a better understanding of purity and its supposed worth, and these are our chief goals here. We
would make it clear at the outset, though, that we do not claim to offer a general account of purity—an account that does justice to the full variety of forms it has assumed and the ends it has been taken to have throughout its history. Our aim is the more modest one of clarifying the character and epistemic significance of what we think is the or at least a central conception of purity—a conception we call the top-
cal conception.

The plan of the paper is as follows. In the next section we provide some historical background. In the third section, we then introduce, motivate, and characterize the notion of purity featured here—the topical conception. In the fourth and fifth sections we present and discuss two examples, and in the final section we offer a few closing thoughts.

2. A brief history of purity

The chief ancient proponent of purity was Aristotle, who in the Posterior Analytics wrote: “...you cannot prove anything by crossing from another kind—e.g. something geometrical by arithmetic.” Such was Aristotle’s influence in these matters that his statement deserves to be called the classical statement of purity. Ancient and medieval mathematicians generally adhered to Aristotle’s prohibition and, as a result, they generally sought to avoid appeal to arithmetic devices in their geometrical work, and vice-versa.

Aristotle’s injunction against kind-crossing was a consequence of his ontology, in which there was a hierarchy of kinds, the salient ordering of which the knowing agent was to recapitulate in her “scientific demonstrations.” Given the general decline in the influence of Aristotle’s ontological and epistemological ideas in the modern era, the extent to which proof which does not cross lines of genus or subject-matter remained an ideal among mathematicians of the modern period is perhaps surprising.

It did retain its appeal, though, even among some who favored the growing use of algebraic methods in geometrical reasoning. One clear example of this was Newton who, despite his development of the infinitesimal calculus, and his masterful use of it in physics, nevertheless defended purity as an ideal of reasoning and decried the increasingly widespread use of algebraic methods in geometry:

Equations are Expressions of Arithmetical Computation, and properly have no Place in Geometry, except as far as Quantities truly Geometrical (that is, Lines, Surfaces, Solids, and Propositions) may be said to be some equal to others. Multiplications, Divisions, and such sort of Computations, are newly received into Geometry, and that unwarily, and contrary to the first Design of this Science... Therefore these two Sciences ought not to be confounded. The Antients did so industriously distinguish them from one another, that they never introduced Arithmetical Terms into Geometry. And the moderns, by confounding both, have lost the Simplicity in which all the Elegancy of Geometry consists.

By “the moderns”, Newton was referring to Viète, Descartes, Wallis and others who made liberal use of algebraic methods in solving geometrical problems. He took exception to the common view that algebraic methods simplified geometrical reasoning. Indeed, he regarded

---

2. Using the Greek expression for “crossing over”, Aristotle’s purity constraint is sometimes called his “injunction against *metabasis*”.
the classical methods as simpler in the most important respects, and he saw this simplicity as ultimately tied to justificative issues.

... that is Arithmetically more simple which is determin'd by the more simple Equations, but that is Geometrically more simple which is determin'd by the more simple drawing of Lines; and in Geometry, that ought to be reckon'd best which is Geometrically most simple.5

This repeated the view, common in Newton’s day, that the construction of a geometrical figure is in an appropriate sense its efficient cause. When combined with the then-equally-common view that we know a thing best when we know it through its cause, this led to the belief that construction of geometrical objects supports better knowledge of them than does algebraic reasoning.6

There were those who did not take this view of classical construction, however. A particularly interesting example was Wallis, who offered a more Platonistic (i.e. less construction-dependent) view of our knowledge of geometrical figures.

... beside the supposed construction of a Line or Figure, there is somewhat in the nature of it so constructed, which may be abstractly considered from such construction; and which doth accompany it though otherwise constructed than is supposed.7

Wallis thus saw geometrical knowledge as more concerned with construction-invariant features of geometrical figures than with features deriving from their construction. In addition, he saw algebraic methods as offering striking gains in simplification over their classical counterparts. Here his views seem directly opposed to Newton’s. In particular, Wallis was not persuaded that the only type of simplicity relevant to geometrical reasoning was what Newton described as “the more simple drawing of Lines” (loc. cit.). As he saw it, the use of algebraic methods commonly afforded more efficient discovery of convincing reasons for (as distinct from proper proofs or demonstrations of) geometrical truths, and in his view, the “official” preferences of traditional geometers had not properly reflected the value of this efficiency.8

Newton’s mathematical interpreter, Colin MacLaurin, followed the lead of his (Newton’s) philosophical interpreter Locke in cautioning against the uncritical use of Wallis’ infinitistic algebraic methods. He did this, however, while acknowledging their efficiency.9 The appeal to the simplicity of classical methods thus seems to have been weaker in him than in Newton.

Mr. Lock ... observes, “that whilst men talk and dispute of infinite magnitudes [where MacLaurin has ‘magnitudes’, Locke had ‘space or duration’], as if they had as compleat and positive ideas of them as they have of the names they use for them, or as they have of a yard, or an hour, or any other determinate quantity, it is no wonder if the incomprehensible nature of the thing they discourse of, or reason about, leads them into perplexities and contradictions ...” Mathematicians indeed abridge their computations by the supposition of infinites; but when they pretend to treat them on a level with finite quantities, they are sometimes led into such doctrines as verify the observation of

---

5. Idem, p. 230
6. Knowing via minimal or simplest cause would have been better still, of course, since it could have been expected to distill a kind of “pure” cause by separating what is essential to the construction of a figure from what is accidental.
7. John Wallis, A Treatise of Algebra, both historical and practical: shewing the original, progress, and advancement thereof, from time to time, and by what steps it hath attained to the height at which it now is (London: John Playford, 1685), p. 291
8. Wallis’ preference is therefore not well described as simply a preference for the simpler over the less simple. Rather, it was a preference for a particular type of simplicity—discovermental simplicity—that he believed had been traditionally undervalued.
this judicious author . . . These suppositions [suppositions concerning the infinite] however may be of use, when employed with caution, for abridging computations in the investigation of theorems, or even of proving them where a scrupulous exactness is not required . . . Geometricians cannot be too scrupulous in admitting of infinites, of which our ideas are so imperfect.\footnote{MacLaurin, op. cit., pp. 45–47}

MacLaurin thus repeated the common observation that algebraic methods did not have the same value for demonstrating truths that traditional, constructional methods had. Wallis believed this to be compatible with their (i.e. algebraic methods') having great value as instruments of discovery or investigation. MacLaurin conceded and even repeated this point. Properly controlled, algebraic methods had distinct value as methods of investigation. It was important, however, not to conflate methods of investigation with methods of demonstration and thus to overlook the limitations of algebraic methods as methods of demonstration. Such at any rate were MacLaurin’s views.\footnote{A reviewer rightly noted that despite the fact that MacLaurin was Newton’s interpreter, the opposition between his (MacLaurin’s) and Wallis’ views is not the same as that between Newton’s and Wallis’. Newton’s opposition to the use of algebraic methods in geometry was on grounds of purity. MacLaurin’s was primarily on grounds of confidence or security. In fact, the fuller truth is more complicated still. As we will see shortly, Wallis’ preference for algebraic methods can be seen as partially based on considerations of purity. He believed that geometry was (or ought to be) about certain construction-transcendant invariances, and that algebraic methods more purely reflect these invariances than do traditional geometrical methods. In addition, it is important to bear in mind that though Newton “officially” preferred geometrical to algebraic methods in geometrical investigations, in practice he made extensive use of the latter.}

Despite Newton’s reservations concerning algebraic methods, mathematicians of the eighteenth century and later generally followed Descartes and Wallis in sanctioning the relatively free use of “impure” algebraic methods in geometry.\footnote{This is truer of mathematicians on the continent than of those in England. For more on the reception of algebraic methods in England, see Helena M. Pycior, Symbols, impossible numbers, and geometric entanglements (Cambridge: Cambridge University Press, 1997) who discusses the reception of Newton’s views among British mathematicians, and its effects in slowing the adoption of the new algebraic methods for a century and a half.} This should not be taken to suggest, however, a general decline in the importance of purity as an ideal of mathematical reasoning. Even (perhaps especially) views like Wallis’ encouraged retention of purity as an ideal of proof. What set these views apart was a different conception of the subject-matter of geometry. It remained geometrical figures, but these were not conceived in the usual way. In particular, they were not seen as essentially tied to characteristic means of construction. Rather, their essential traits were taken to be those which were invariant with regard to the method(s) of construction. Or so it may be argued.\footnote{For more on what was meant by invariance in early modern practice, cf. Michael Detlefsen, “Formalism,” in Stewart Shapiro (ed.), Handbook of the Philosophy of Mathematics and Logic (Oxford University Press, 2005), Section 4.2.1.}

The distinctive feature of the algebraists was therefore not a move away from purity as an ideal of proof, but a move away from the traditional conception of geometrical objects. Specifically, it was a move away from a view which saw geometrical objects as being given or determined by their (classical) methods of construction, and a move towards a view which saw the essential properties of geometrical figures to be those which were invariant with respect to their means of construction. For Wallis, these were mainly arithmetical or algebraic features.

Purity was also an expressly avowed ideal of such later figures as Lagrange, Gauss, Bolzano, von Staudt and Frege. It figured particularly prominently in Bolzano’s search for a purely analytic (i.e. non-geometric) proof of the intermediate value theorem.\footnote{The intermediate value theorem states that if a real function \( f \) is continuous on a closed bounded interval \([a, b]\), and \( c \) is between \( f(a) \) and \( f(b) \), there is an \( x \) in the interval \([a, b]\) such that \( f(x) = c \).} the chief motive behind which he described as follows:

[I]t is . . . an intolerable offense against correct method to derive truths of pure (or general) mathematics (i.e. arithmetic, algebra,
analysis) from considerations which belong to a merely applied (or special) part, namely, geometry. Indeed, have we not felt and recognized for a long time the incongruity of such *metabasis eis allo genos?* Have we not already avoided this whenever possible in hundreds of other cases, and regarded this avoidance as a merit? . . . [If] one considers that the proofs of the science should not merely be *certainty-makers* [Gewissmachungen], but rather *groundings* [Begründungen], i.e. presentations of the objective reason for the truth concerned, then it is self-evident that the strictly scientific proof, or the objective reason, of a truth which holds equally for all quantities, whether in space or not, cannot possibly lie in a truth which holds merely for quantities which are in space.15

For Bolzano, then, truly scientific proof was demonstration from objective grounds and this was in keeping with the demands of purity.

Frege’s logicist program also called for purity—specifically, the purification of arithmetic from geometry.

Even in pre-scientific times, because of the needs of everyday life, positive whole numbers as well as fractional numbers had come to be recognized. Irrational as well as negative numbers were also accepted, albeit with some reluctance—and it was with even greater reluctance that complex numbers were finally introduced. The overcoming of this reluctance was facilitated by geometrical interpretations; but with these, something foreign was introduced into arithmetic. Inevitably there arose the desire of once again extruding these geometrical aspects. It appeared contrary to all reason that purely arithmetical theorems should rest on geometrical axioms; and it was inevitable that proofs which apparently established such a dependence should seem to obscure the true state of affairs. The task of deriving what was arithmetical by purely arithmetical means, i.e. purely logically, could not be put off.16

In expressing his concern that proofs reveal “the true state of affairs”, Frege echoed Bolzano. Both seem to have believed in some type of objective ordering of mathematical truths a distinct exposition (or at least reflection) of which was proof’s highest calling.

Purity remained a guiding ideal of twentieth-century thinking as well. An example is the search for an “elementary” proof of the prime number theorem, which states that the prime number theorem states that the number of primes up to an integer \( n \) is approximately \( \frac{n}{\log n} \). The prime number theorem was first proved by Hadamard and de la Vallée Poussin in 1896. Because their proofs used methods from complex analysis, however, they were widely regarded as imperfect.17 In 1949, Paul Erdös18 and Atle Selberg19 found proofs that avoided such methods. This discovery was believed to be of such importance as to contribute significantly to Selberg’s earning a Fields Medal.

But though Selberg and Erdös pursued and valued purity, they offered neither a clear general characterization of it nor a statement of why it should be valued. The *Bourbakiste* Jean Dieudonné did better.

. . . an aspect of modern mathematics which is in a way complementary to its unifying tendencies . . . concerns its capacity for

---

15. Bernard Bolzano, “Purely analytic proof of the theorem that between any two values which give results of opposite sign there lies at least one real root of the equation,” in William Ewald (ed.), *From Kant to Hilbert* (Oxford University Press, 1999), p. 228; our translation is slightly different from the translation there.


18. Paul Erdös, “On a new method in elementary number theory which leads to an elementary proof of the prime number theorem,” *Proceedings of the National Academy of Sciences, USA* 35 (1949)

sorting out features which have become unduly entangled…. It may well be that some will find this insistence on “purity” of the various lines of reasoning rather superfluous and pedantic; for my part, I feel that one must always try to understand what one is doing as well as one can and that it is good discipline for the mind to seek not only economy of means in working procedures but also to adapt hypotheses as closely to conclusions as is possible.\textsuperscript{20}

Dieudonné evidently saw “closeness” or “proximity” of topic between the premises and conclusion of a proof as an important quality of it. He seems also to have believed that working to maximize this closeness typically furthers the intellectual development of the prover.\textsuperscript{21}

Here we see appreciation not of the epistemic value of purity, but of what, following traditional usage, we might call its intervenient value. In taking such a view, Dieudonné joins a long tradition of thinkers who have praised the study of mathematics as a means of improving capacity for reasoning. A well-known exponent of this view, Francis Bacon, memorably compared the benefits of doing of pure mathematics to those of playing tennis.

In the Mathematics I can report no deficience, except it be that men do not sufficiently understand the excellent use of the Pure Mathematics, in that they do remedy and cure many defects in the wit and faculties intellectual. For if the wit be too dull, they sharpen it; if too wandering, they fix it; if too inherent in the sense, they abstract it. So that as tennis is a game of no use in itself, but of great use in respect it maketh a quick eye and a body ready to put itself into all postures; so in the Mathematics, that use which is collateral and intervenient is no less worthy than that which is principal and intended.\textsuperscript{22}

Many other thinkers have made similar claims.\textsuperscript{23} Among these was Bolzano, who wrote:

It is quite well known that in addition to the widespread use which its application to practical life yields, mathematics also has a second use which, while not so obvious, is no less useful. This is the exercise and sharpening of the mind: the beneficial development of a thorough way of thinking.\textsuperscript{24}

We call attention to this view of purity as an intervenient or developmental virtue in order to distinguish it from the treatment of purity we offer here. Intervenient value is broadly pragmatic in character. We are interested, by contrast, in the distinctively epistemic or justificative value of purity, one variety of which we will discuss in §§ 3, 4.

For us, then, the question is: “What is the epistemic value of purity?” Here there is relatively little to go on in the literature. The little there is we will now briefly survey.

In his 1900 Problems address, Hilbert made a statement that suggested that he saw purity as an epistemic virtue. Specifically, he claimed that in solving a mathematical problem we ought to stay as close as possible to the conceptual resources used in stating, or, perhaps better, understanding that problem. As he put it:

It remains to discuss briefly what general requirements may be justly laid down for the solution of a mathematical problem. I should say first of all, this: that it shall be possible to establish

\textsuperscript{20} Jean Dieudonné, \textit{Linear algebra and geometry} (Boston, Mass.: Houghton Mifflin Co., 1969), p. 11

\textsuperscript{21} Hans Freudenthal, “Review of Dieudonné, \textit{Algèbre linéaire et géométrie élémentaire},” \textit{American Mathematical Monthly} 74/6 (1967), pp. 744–748 criticizes Dieudonné on these points.

\textsuperscript{22} Francis Bacon, \textit{The two booke of Francis Bacon. Of the proficience and advancement of learning, divine and humane} (London: Henrie Iomes, 1605), Bk. 2, VIII, 2, emphasis added

\textsuperscript{23} See e.g. the anonymously written dedication to Nicholas Saunderson, \textit{The elements of algebra} (Cambridge University Press, 1740), xiv–xx.

\textsuperscript{24} Bernard Bolzano, “Preface to Considerations on some Objects of Elementary Geometry,” in William Ewald (ed.), \textit{From Kant to Hilbert} (Oxford University Press, 1999), p. 172
Purity Of Methods

out a previously determined method and by employing certain limited means.\(^{29}\)

The pre-determined or limited means mentioned \textit{might} represent a restriction to \textit{pure} methods, but it need not represent only that. Restriction to pure proof is only one type of restriction. And though it may offer something to learn, so do other types of restrictions. That there is more to be learned from a restriction to pure methods than from other types of restrictions is an unwarranted assumption. Accordingly, a systematic preference for pure proof is likewise unwarranted. This at any rate is what Hilbert seems to have been asserting in the passage from the \textit{Grundlagen} just noted.

Purity has retained a following as an ideal of proof to the present day. Gel’fond and Linnik, for example, see it as representing a “natural desire” to attain elementary solutions to elementary problems.

There can be no question of renouncing transcendental methods in modern number theory. Yet the investigator feels a \textit{natural desire} to seek more arithmetic approaches to the solution of problems which can be stated in elementary terms. Apart from the obvious methodological value of such an approach, it is important by reason of giving a simple and natural insight into the theorems obtained and the causes underlying their existence.\(^{30}\)

There are other cases we might mention as well.\(^{31}\) We trust, though, that we have said enough to motivate our interest in purity. In the remaining sections, we will attempt to say more clearly what purity is and how it might function as an epistemic virtue. We will also present and discuss some examples.

\textbf{25}\ David Hilbert, “Mathematische Probleme,” \textit{Archiv der Mathematik und Physik (3rd series)} 1 (1901), p. 257, emphasis added


\textbf{28}\ In classifying preference for pure proof as \textit{subjective} Hilbert seems to have had something like the following in mind: we generally stand to learn as much from impure proof as from pure; therefore, a systematic preference for pure proof is unjustified (and in that sense “subjective”).


3. Specific Ignorance & Its Relief

One important goal of epistemic development is to reduce ignorance of which an investigator is aware and relief of which she in some important sense pursues. Among the different types of ignorance investigators may seek to alleviate is one we call “specific ignorance.” This is ignorance of solutions to specific problems of which we are aware and whose solution we desire.32 How, exactly, relief of ignorance ought to be thought of, and how, so conceived, it and its relief ought best to be measured are subtle and complicated matters. In this paper we operate from a basic but only partially developed conception of ignorance and its relief, one which sees specific ignorance as capable of relief even in cases where global ignorance (measured, for example, by the number of problems we do not know how to solve, or by the proportion of problems we know to which we have solutions) may not be.

We should at this point, say a few words concerning the terminology of “relief.” We say “relief” rather than “reduction”, “decrease” or something similar because we believe that alleviating a case of specific ignorance may be an epistemic good even if it does not, overall, result in a lasting reduction of the extent of our ignorance.33 At the same time, however, we do not believe that absolutely all ways of eliminating instances of specific ignorance should count as relieving it. Specifically, we would not count the elimination of a case of specific ignorance as relief if application of the means of elimination itself systematically produced further cases of specific ignorance not eliminated by such application. So, even though relief of specific ignorance does not itself necessarily entail reduction, it does require the absence of systematic effects of replenishment.

Relief of specific ignorance and pure problem solution are thus joined at the hip in our account. Relating them in this way requires that we make certain assumptions, of course. Specifically, it requires that we assume that (i) knowledge-seekers, or what we will generally refer to as investigators, can be aware of cases of specific ignorance, and that (ii) they are within their rights (i.e. they operate within the purview of their roles as investigators) to seek to relieve it.34 On the view we will now present, the basic epistemic virtue of pure problem solutions is that they are particularly effective means of relieving specific ignorance.

---

32. We are not assuming any specific basis or type of basis for this desire. Specifically, we are not assuming that it has to represent what might be thought of as purely epistemic reasons or motives.

33. We intend this to apply to both cardinal conceptions of extent (i.e. the sheer number of instances of our specific ignorance) and to proportional conceptions of extent (i.e. the proportion of instances of specific ignorance that have been relieved).

34. We note that on our view investigators are not generally to be thought of as individuated along the same lines as persons or human beings. Rather, they are to be thought of as conductors of investigations. Generally speaking, one and the same human person can and typically will be involved in the conduct of more than one investigative role at a given time.
3.1 Problems
We’ll generally refer to the attempts of investigators to relieve specific ignorance as “directed investigations.” We’ll call the questions or problems towards which these investigations are directed “directing questions” or “directing problems.” The directing problems with which we shall primarily occupy ourselves are yes-no in form: for example, “Are there infinitely many primes?”. We adopt this focus not only for convenience, but also because we believe the yes-no form to be perhaps the basic form of question or problem.

We represent the yes-no problems that direct directed investigations as ordered triples $P = (\gamma_{y/n}, P, \phi)$, where $\gamma_{y/n}$ stands for what we’ll call a “yes-no” interrogative attitude; $P$ stands for a propositional content; and $\phi$ stands for a formulation of $P$ (from whatever pertinent formulative resources are available to the investigator in question). There are thus three elements that determine the identity of a directing yes-no problem—its attitude, its content, and the formulative means by which the content is represented to the investigator.  

3.2 Solutions
On our analysis, directing problems are solved by directed investigations. More accurately, they’re solved by the results of directed investigations. Here by “result”, we mean the body of evidence produced by a directed investigation for the purpose of answering its directing question. An “answer”, on our analysis, is thus a “yes” or “no” response backed by evidence, and not merely a “yes” or “no” response.

To answer a question is one way to properly terminate an investigation. Another way is to dissolve the question. This type of “solution” is in fact suggested by the very etymology of the verb ‘to solve’, which is from the Latin solvere—to loosen, release, unbind. We take this “releasing” aspect of solution seriously, if not entirely literally. If problems are solved in order to relieve specific ignorance, then, generally speaking, a solution to a problem is anything that eliminates the specific ignorance it represents. Accordingly, we believe, there are two broadly different paths to the solution (in the sense just described) of a directing problem:

(1) to provide an answer to it, and

(2) to rationally dissolve it.

The idea behind (2), of course, is that once a problem ceases to be a directing problem for an investigator, it no longer represents a locus of specific ignorance for her. To repeat what we said above, we take the epistemic virtue represented by pure problem solutions—or, more particularly, what we will call topically pure solutions—to be their special potential for relieving specific ignorance. That they have this capacity is

---

35. Questions of other interrogative types (such as “what” or “why” problems) do not ordinarily express a proposition, but instead only propositional functions (cf. John Searle, *Speech Acts* (Cambridge University Press, 1969), p. 31). For instance, consider the question “What is $\int_0^1 x \, dx$?”. Its content is the propositional function $\int_0^1 x \, dx = X$, for $X$ a propositional variable ranging over some intended domain (which we will take to be the complex numbers). So we may restate the question “What is $\int_0^1 x \, dx$?” as “Does there exist an $X$ such that $\int_0^1 x \, dx = X$?” While it is true that this restatement results in a different question, the change affects the interrogative type and formulation of the question but not its content; and indeed the formulation changes only to reflect the shift in interrogative type. For this reason we consider this reformulation adequately close to the question as originally formulated for our analysis. This permits problems of non-“yes-no” interrogative type to be represented as yes-no problems by ordered quadruples $(\gamma_{y/n}, P(X, \phi), D)$, where $X$ denotes finitely many propositional variables occurring in $P(X)$ ranging, respectively, over the domains $D$. Yes-no problems are thus a special case where $X$ (and thus $D$) are all empty.

36. The thinking behind the notion of dissolution is that when a problem becomes a problem for an investigator it does so partly in virtue of commitments made by the investigator. These commitments determine the character and content of the problem. To the extent that this is so, retraction of these commitments would have the net effect of eliminating the original problem. Retraction of commitments might therefore affect not only a problem’s solution, but the very problem itself. More on this later.
due to their distinctive potential to dissolve those problems that they do not perdurantly solve.

3.3 **Co-finality, stable solution & dissolution**

Our analysis of problem solution centers on a relationship between problems and their solutions that we call “co-finality”. Roughly, a problem \( P = (\alpha, n, P, \phi) \) and its solution by an investigation \( I_\alpha \) are “co-final” for \( \alpha \) when the result \( E \) of \( I_\alpha \) retains its status as a solution of \( P \) for \( \alpha \) for as long as \( P \) continues as the content of \( P \) for \( \alpha \), and so remains the problem it is for her. Co-finality thus implies that a solution will perdure as a solution to a given problem as long as that problem itself remains the problem it is.

A co-final solution of a problem thus ensures reduction of specific ignorance regarding it. That this is so is due to the fact that co-final solution ensures one of two outcomes. Either it will survive as a solution, or it will not. If it does, it reduces specific ignorance by solving a problem that continues to be a problem. If it does not, then neither will the problem to which it was originally a solution. Whatever modification changes the original solution to a non-solution will also remove it from the register of problems to be solved. Either way, the specific ignorance that is represented by the original problem’s being a problem will be reduced. In the one case, it will be reduced by the addition of an enduring solution. In the other case, it will be reduced by the deletion of the problem from the class of problems that are to be solved.

We will now look more closely at the key notions involved in the statements just given.

**Co-finality**

\( E \) is co-final with \( P = (\alpha, n, P, \phi) \) for \( \alpha \) just in case any retraction of \( E \) by \( \alpha \) would dissolve \( P \) for her. Here by retraction we mean a change in \( \alpha \)’s attitude according to which (i) there is a premise or inference of \( E \) that \( \alpha \) no longer accepts, and (ii) what remains of \( E \) once this premise or inference is removed (call it \( E^- \)) is no longer a justification of \( P \) (or its denial) for \( \alpha \). The general line of thinking here is the following: (i) that among the things which go towards making a problem \( P \) the problem that it is are various commitments or beliefs, (ii) that such commitments might be among the commitments made by a given solution of \( P \) and therefore (iii) that among the effects that retracting a solution of a problem might have is the dissolution of the problem itself.

One the account developed here, co-finality is that trait of pure problem solution which chiefly underlies its epistemic value—namely, its unique capacity to relieve specific ignorance.

**Stable Solution**

A directed investigation \( I_\alpha \) stably solves its directing problem \( P = (\alpha, n, P, \phi) \) for an investigator \( \alpha \), when it provides evidence \( E \) (what we call the result of \( I_\alpha \), or the solution of \( P \)) such that the following two conditions hold:

(i) (a) \( E \) justifies belief in \( P \) for \( \alpha \) (according to appropriate standards of justification) or (b) \( E \) justifies belief in not-\( P \) for \( \alpha \), and

(ii) \( E \) is co-final with \( P \) for \( \alpha \).

The significance of stability, in the present sense, is that a stable solution to a given problem lasts as long as the problem does. Overall, this seems to be a desirable property for solutions to have. It also seems to be a dictate of the “logic” of the notion of a (genuine) problem solution. We take this logic to imply that rational confidence that \( \sigma \) solves \( P \) cannot exceed rational confidence that so long as \( P \) remains the problem it is, \( \sigma \) will continue to be a solution to it.\(^{37}\)

Stable solutions of problems promise to relieve specific ignorance by reducing the stock of problems that agents can represent to themselves (and on which account alone they can be regarded as genuine problems for them), but which they cannot solve. Such reduction can take the form of either a solution of a retained problem, or the dissolution, and hence the non-retention, of a problem. Either way, specific

\(^{37}\) In this respect, problem solution resembles knowledge. I cannot know that \( P \) but also believe that at some future time, \( P \) will be false or unjustified. Similarly, I cannot rationally believe that \( \sigma \) is a solution to \( P \) while also believing that at some future time it will not be.
ignorance is relieved because both solution and dissolution result in the elimination of a problem \( P \) that is (a) retained, but (b) unsolved.

Dissolution

A problem \( P = (\alpha/y, P, \phi) \) is dissolved for \( \alpha \) when it is either “attitudinally”, “contentually” or “formulatively” dissolved for her. \( P \) is attitudinally dissolved for \( \alpha \) when reasons sufficient to warrant her adopting \( ?_{\alpha/y} \) towards \( P \) are no longer available to her. \( P \) is contentually dissolved for \( \alpha \) when her formulative resources no longer include that part of \( \phi \) that allowed her to represent \( P \) to herself.³⁸

The central idea behind dissolution is that the commitments that are given up in the retraction of a pure problem solution are among those that determine the content of the problem or its interrogative attitude. This being so, retraction can change not only the investigator’s commitment to the problem’s solution, but the identifying characteristics of the problem itself. It does not of course “destroy” content that was once among the investigator’s problemistic contents. Neither does it change the fact that a particular interrogative attitude was once taken towards that content. What it does change, though, is the particular problems that lie in the investigator’s “active” file—what we will call her investigative corpus. These are the problems that represent the (instances of) specific ignorance the investigator’s directed investigations are concerned to alleviate.³⁹

Is or should dissolution of a problem be sensitive to the differences between the different types of reasons for retraction? We don’t know a general answer to this question. Neither, though, does our argument require that we do. In claiming that retractions of solutions may lead to the dissolution of problems, we are only saying that certain beliefs (what we will call problem-determining beliefs) are commonly among the elements that determine our grasp or understanding of mathematical problems. In linking purity of problem solution with problem dissolution, we are calling attention to a notion of purity (namely, topical purity) which seeks to constrain the solution of a problem by requiring that it make use of only problem-determining beliefs. Strictly speaking, not even this is true. We are only obliged to require that the retractable beliefs used in a pure solutions are all problem-determining beliefs.

³⁸. The condition for formulative dissolution is the same as for contentual dissolution. The notion of formulative dissolution is different from the notion of contentual dissolution, however. Formulative dissolution is constituted by a change in an investigator’s commitments which results in \( \phi \)’s no longer qualifying as a formulation of a content \( P \). This could be either because the commitments in question constitute changes in what qualifies an expression as a formulation of a content in general, because they change what qualifies an expression as a formulation of \( P \), or because they amount to removing a condition which was needed to qualify \( \phi \) as a formulation of \( P \) under a given view of what such qualifying conditions are.

³⁹. Intuitively, \( \alpha \)’s investigative corpus is intended to be the family of problems to which she wants solutions and towards whose solution her investigative efforts are directed.

Purity Of Methods

What we are calling (resolutive) retractions can, of course, take different forms and can be made for different reasons. Different forms include changes from belief to non-belief and from belief to various types and/or degrees of disbelief. Different reasons include conviction that something formerly believed to be true is false, or that something formerly believed to be true is in fact inconsistent or rationally absurd. Generally speaking, there will be as many ways to dissolve a problem as there are ways for resolutive retractions to imply contentual, interrogative or formulative changes in it.⁴⁰

For this to be so, of course, requires that beliefs should commonly be among the things that determine problems, and that whatever exact acts and attitudes may be covered by “resolutive retraction”, a significant portion of this range should be comprised of problem-determining beliefs.⁴¹

⁴⁰. Is or should dissolution of a problem be sensitive to the differences between the different types of reasons for retraction? We don’t know a general answer to this question. Neither, though, does our argument require that we do. In claiming that retractions of solutions may lead to the dissolution of problems, we are only saying that certain beliefs (what we will call problem-determining beliefs) are commonly among the elements that determine our grasp or understanding of mathematical problems. In linking purity of problem solution with problem dissolution, we are calling attention to a notion of purity (namely, topical purity) which seeks to constrain the solution of a problem by requiring that it make use of only problem-determining beliefs. Strictly speaking, not even this is true. We are only obliged to require that the retractable beliefs used in a pure solutions are all problem-determining beliefs.

⁴¹. It has been suggested that this commits us to the implausible view that

\[ \Delta : \text{if we change from believing a proposition } P \text{ to believing it to be false, the content of } P \text{ or our grasp of it must change too.} \]

We are not committed to \( \Delta \) and, indeed, we have sympathy with those who regard it as implausible. To think that we are committed to \( \Delta \) is to fail to distinguish two quite different ideas—namely, (a) that rejecting a problem-determining belief changes its (i.e. the belief’s) contents, and (b) that rejecting a problem-determining belief changes the content of the problem it (partially) determines. We are committed to (b), not to (a).
3.4 The Dynamics of Purity: An Illustration

To make the above theoretical description of the effects of and interrelations between the various elements of purity more concrete, it may be useful to consider an example. We will therefore briefly describe a case which arises from a geometric problem originally posed by J.J. Sylvester\(^{42}\), and later formulated by Erdős\(^{43}\) as follows: “Let \( n \) given points have the property that the straight line joining any two of them passes through a third of the given points. Show that the \( n \) points lie on a straight line.”

Many different solutions to Sylvester’s problem have been offered. We are particularly interested in a “metrical” solution given by Kelly\(^{44}\). We call this proof “metrical” because it assumes that there is a shortest line between every line and point not on that line. That is, it uses a metrical notion of distance.

Since a straight line (segment) is, by definition, the shortest distance between two points, it makes sense to think that distance is relevant to the problem.\(^{45}\) Despite this, however, metrical definitions of straight line have not been the rule in recent geometry. An illustration of the attitudes involved here is given Hilbert’s discussion of his fourth problem.\(^{46}\) He rejected the metrical definition of straight line because he believed that it ought to be a theorem rather than a definition.

Coxeter agreed on the grounds that distance is “essentially foreign to this problem, which is concerned only with incidence and order”.\(^{47}\)

Supposing that Hilbert and Coxeter are correct, accepting this metrical assumption is not crucial to the ability to understand the problem. As a result, one could “retract” one’s commitment to Kelly’s assumption that there is a shortest line between every line and point not on that line, without changing one’s understanding of Sylvester’s problem. It would follow that Kelly’s metrical proof is not a co-final solution to Sylvester’s problem.

For present purposes, the critical assumption of Kelly’s solution was that there is a shortest line between every line and point not on that line. Had Kelly retracted this premise, his understanding of Sylvester’s problem would not have changed.\(^{48}\) As a result, his solution should not be considered a stable solution of Sylvester’s problem,


\(^{45}\) Archimedes, Leibniz and Legendre, for example, all accepted such a definition at one time or another.

\(^{46}\) Hilbert, op. cit.

\(^{47}\) Cf. Coxeter, op. cit., p. 27. We explain briefly what Coxeter means by “incidence and order”. Incidence in geometry concerns just points, straight lines, the incidence of points on straight lines, and the intersections of straight lines (in points). The surface grammar of Sylvester’s problem indicates straightforwardly that it is a problem concerning incidence: it mentions points, straight lines, and the incidence of points on straight lines. That it also concerns order is less obvious. Coxeter elsewhere gave the following argument for this claim:

The essential idea [for problems like Sylvester’s] is intermediacy (or ‘betweenness’), which Euclid used in his famous definition: “A line (segment) is that which lies evenly between its ends.” This suggests the possibility of regarding intermediacy as a primitive concept and using it to define a line segment as the set of all points between two given points.” (Cf. H. S. M. Coxeter, Introduction to Geometry, second edition (Wiley, 1989), p. 176)

Coxeter’s view is that a proper definition of straight line involves order, as Euclid’s definition brings out. Here Coxeter relies on an unusual translation of Euclid’s definition I.4; by contrast, Heath’s translation reads, “A straight line is a line which lies evenly with the points on itself”, avoiding reference to order. For a careful “reverse mathematical” comparison of axiomatic systems sufficient for formalizing various solutions to Sylvester’s problem, including Kelly’s as well as Coxeter’s order-theoretic proof, cf. Victor Pambuccian, “A reverse analysis of the Sylvester-Gallai theorem,” Notre Dame Journal of Formal Logic 50/3 (2009).

\(^{48}\) The status of metrical assumptions in geometry has long been controversial, with geometers since von Staudt seeking to eliminate the metrical from geometry, particularly from projective geometry. The above-mentioned remark of Hilbert’s that the metrical definition of line should be regarded as a theorem instead ought to be understood along these lines. We may thus see this move away from the metrical in geometry amongst nineteenth-century geometers as examples of retraction.
and hence not a pure solution in the sense that is our focus here. Such, at any rate, is our view.

3.5 Topical purity

Generally speaking, a purity constraint restricts the resources that may be used to solve a problem to those which determine it. We may treat the measure of such determination as a parameter, and investigate particular measures and the purity constraint each such measure induces. In this paper, our focus is “topical determination”. The topically determining commitments of a given problem \( \phi \) are those which together determine what its content is for a given investigator.

In mathematics, among those things that determine contents are definitions, axioms concerning primitive terms, inferences, etc. We’ll generally refer to such items as commitments. What we are calling the topic \( T_P \) of the problem \( P \) (or, equally, of the investigation \( I_P \)) is a set of commitments. Specifically, it is that set of commitments each element of which is such that were \( a \) to retract it, the content of \( \phi \) would not be what it is for her.

We say that a solution \( E \) of \( P \) is topically pure when it draws only on such commitments as topically determine \( P \).

The epistemic significance of topical purity derives from the stability it brings to problem solutions. Every topically pure solution \( E \) to a problem \( P \) is stable in the sense that were \( a \) to retract a premise or inference from \( E \), the content of \( P \) would change for her. In other words, her retraction would contentually dissolve \( P \) for her (i.e. \( E \) would be co-final with \( P \) for her). By contrast, if \( E \) were a topically impure solution to \( P \), there would be premises or inferences in \( E \) that the investigator could retract without contentually dissolving \( P \). In that case, \( E \) would not be stable in the aforementioned sense.


To make a theory of topical purity complete, we would of course need an account of how topics of problems are generally determined. This is a difficult task that is beyond what we presently know how to do. We can make a start on it, though, by giving some cases we think illustrate topical purity. It is to this that we now turn.

4. Further Examples

Consider the infinitude of primes problem (IP) (that is, the problem whether, for all natural numbers \( a \), there is a natural number \( b > a \) such that \( b \) is prime) taken at “face value.” By this we mean an understanding of IP determined by such commitments as the following:

1. axioms for successor (for a natural number \( n \), written \( S(n) \))
2. (first-order) induction axioms for making precise the view that the natural numbers ‘begin’ with 1 and ‘continue’ onward thereafter
3. definitions and axioms for an ordering on the natural numbers which, following typical practice, would specify a linear discrete ordering
4. a usual conception of primality such as that \( a \) is prime if and only if \( a \neq 1 \) and the only numbers dividing \( a \) are 1 and \( a \)
5. and the accompanying definitions of divisibility and multiplication (e.g. \( a \) divides \( b \) (written \( a|b \)) if and only if there exists \( x \) such that \( a \cdot x = b \))

The first-order Peano axioms for the natural numbers provide a reasonable formulation of these commitments, augmented by the definition of primality and divisibility just described.

What is perhaps the most widely known solution of IP is that given in Elements IX.20. The argument there proceeds essentially as follows. If \( a = 1 \), then since \( 2 = S(1) \) is prime, we know that there is a prime greater than \( a = 1 \). So suppose that \( a > 1 \). Let \( p_1, p_2, \ldots, p_n \) be all the primes less than or equal to \( a \), and let \( Q = S(p_1 \cdot p_2 \cdots p_n) \). Note that \( Q \) has a prime divisor \( b \). For each \( i, b \neq p_i \); if not, then \( b|(p_1 \cdot p_2 \cdots p_n) \) and \( b|S(p_1 \cdot p_2 \cdots p_n) \), and so \( b = 1 \), contradicting the primality of \( b \). Finally, either \( b > a \), or \( b \leq a \), but since \( b \leq a \) contradicts that the \( p_i \)
were all of the primes less than or equal to \( a \), we may conclude that \( b > a \).

This proof has several steps that themselves require proof. Examples are the step that consists in the assertion that if \( b | \{ p_1 \cdot p_2 \cdots p_n \} \) and \( b | S(p_1 \cdot p_2 \cdots p_n) \), then \( b = 1 \), or the step in which it is asserted that if \( a | b \) and \( a | S(b) \), then \( a = 1 \). The typical proofs of these results (e.g. in elementary number theory textbooks) can be carried out from the Peano axioms, so it is reasonable to think of these axioms together with the aforementioned definitions of primality and divisibility as at least approximating the topic of IP.

A solution to IP that clearly leads outside the topic just identified is a topological proof offered in 1955 by Furstenberg,

1. The set \( \{ B_{a,b} : a, b \in \mathbb{Z}, b > 0 \} \), where \( B_{a,b} \) denotes the arithmetic progression \( \{ a + bn : n \in \mathbb{Z} \} \), is a basis generating a topology on the integers. [Proved by elementary point-set topological and arithmetic means.]

2. For all \( a, b \in \mathbb{Z}, b > 0 \), \( B_{a,b} \) is both open and closed. [Proved by elementary point-set topological and arithmetic means.]

3. In a topology, unions of finitely many closed sets are closed. [Proved by elementary point-set topological means, in particular using boolean operations on sets.]

4. The union of finitely many \( B_{a,b} \) is closed. [By (2) and (3).]

5. Every integer \( m \) besides \( \pm 1 \) has a prime factor, i.e. for some prime \( p \) and some integer \( n, m = pn \). [By the Fundamental Theorem of Arithmetic.]

6. Every integer besides \( \pm 1 \) is contained in some \( B_{0,p} \) for \( p \) prime. [By (5) and the definition of \( B_{0,p} = \{ pn : n \in \mathbb{Z} \} \].

7. Let \( A = \bigcup_{p} B_{0,p} \) for \( p \) prime. Then \( A = \mathbb{Z} - \{-1,1\} \). [By (6).]

8. (a) Suppose there are only finitely many primes, so that \( A \) is a union of finitely many \( B_{0,p} \).

(b) Then \( A \) is a closed set in our topology. [By (4).]

(c) Then \( \{-1,1\} \), being the complement of a closed set, is open. [By the definition of closed set.]

(d) The basic open sets \( B_{a,b} \) are all infinite. [By the infinitude of \( \mathbb{Z} \).

(e) Each open set is a superset of some basic open set. [By the definition of basis.]

(f) This contradicts the finitude of \( \{-1,1\} \). [By (8c), (8d), (8e).]

9. Hence there are infinitely many primes, and so for every \( a \), there is a prime \( b > a \).

---

50. The Peano axioms include axioms for addition. However, addition is not explicitly mentioned in IP as formulated here. Consequently one might maintain that a pure solution to IP cannot draw on additive resources. A fuller discussion of this point would take us too far afield to be developed here, and will be given elsewhere.


52. Proof: we will check the two necessary conditions for being a basis of a topological space on the integers. Firstly, each integer \( x \) must be contained in a basis element \( B_{a,b} \). For this we may take \( B_{a,b} \) for any difference term \( b > 0 \) we like. Secondly, if an integer \( x \) belongs to the intersection of two basis elements \( B_{a,b} \) and \( B_{c,d} \), then it must belong to a third basis element \( B_{e,f} \) such that \( B_{e,f} \subset B_{a,b} \cap B_{c,d} \). For this we may take \( e = x \) and \( f = \text{lcm}(b,d) \) (where \( \text{lcm}(x,y) \) denotes the least common multiple of \( x \) and \( y \)), so that \( B_{e,f} = \{ x + \text{lcm}(b,d) \cdot n : n \in \mathbb{Z} \} \). Clearly \( x \in B_{e,f} \). We must show that if \( y \in B_{e,f} \) then \( y \in B_{a,b} \cap B_{c,d} \). We have that \( y = x + \text{lcm}(b,d) \cdot n \) for some \( n \). Then \( y = x + bn' \) for some \( n' \), and \( y = x + dn'' \) for some \( n'' \); and hence \( y \in B_{a,b} \cap B_{c,d} \) since \( x \in B_{a,b} \cap B_{c,d} \). By assumption, and since any member of an arithmetic progression may be taken as its initial term, we also have that \( B_{a,b} = B_{a,b} \) and \( B_{c,d} = B_{c,d} \). Therefore, \( y \in B_{a,b} \cap B_{c,d} \).

53. Proof: note that

\[
B_{a,b} = \mathbb{Z} - \bigcup_{1 \leq i \leq b-1} B_{a+i,b}
\]

i.e. the complement of the union of the other arithmetic progressions with the same difference \( b \) as \( B_{a,b} \). Since the union of open sets \( B_{a+i,b} \) is open (by the definition of topological space), \( B_{a,b} \) is the complement of an open set and hence is closed.
5. Discussion of Furstenberg’s Solution

In our view, Furstenberg’s solution to IP is topically impure. The first step of his proof establishes that certain arithmetical progressions form a basis for a topological space. Accepting this engenders several set-theoretic commitments. It also requires commitments to definitions of a topological space and a topological basis. The second step adds further commitment to definitions of open and closed sets in a topology. As we see it, retraction of any of these commitments would not in itself require a corresponding change in our understanding of IP.

A few words on the retraction of definitions are in order. When we accept a definition, we represent the concept it defines as being formulated by that definition. Retraction of a definition entails neither that no other definition of that concept can be accepted, nor that nothing falls under that concept. We saw this earlier when thinking of retracting the metrical definition of straight line: it was replaced by other definitions of the same concept, for instance Coxeter’s order-theoretic definition.54

54. This raises the question of why commitments to definitions would ever be retracted, which we now briefly address. Perceived incoherence is one reason for retracting, but there are other reasons. Broadly speaking, “fruitfulness” is another (class of) reason(s). Mathematicians can and have withdrawn (i.e. changed) definitions because they think the new definition represents a more fruitful way of looking at what it is they’re interested in (cf. Jamie Tappenden, “Extending Knowledge’ and ‘Fruitful Concepts’: Fregean Themes in the Foundations of Mathematics,” Nous 29 (1995) for details). Alternately, it may simply be that a new definition, even one extensionally identical to the one retracted, is more economically formulated than the one retracted. Another reason comes to the fore in the projective geometric tradition of the nineteenth century. As we have already pointed out, metrical considerations (including metrical definitions of straight line, for instance) were retracted, and replaced with purely incidence-theoretic considerations on grounds of generality or basicness. To study geometry with metrical assumptions in force is to limit the study’s generality to only a special case of geometry, where geometrical concepts could be developed in a more general or basic way through non-metrical projective considerations. These then are some reasons why retraction, and in particular retraction of definitions, might happen.

In the particular case at hand, when I retract my commitment to a definition of topological space, I no longer represent the concept of topological space to myself in the way formulated by that definition. Since my understanding of IP does not require that concept to be represented to myself, this retraction does not dissolve the problem.

It follows that Furstenberg’s solution of IP is topically impure, lacking the special connection with IP that is required in order for it to be co-final with it.55 An explanation for this is that Furstenberg’s solution seems to be inspired by the view that, at bottom, arithmetic (or at least that part of it which has most to do with the IP) is really topological in character.

Colin McLarty raised essentially this point in correspondence. As he sees it, there is a preferred way of understanding of IP, and on this understanding one is committed to topological principles such as those used by Furstenberg. In McLarty’s view, then, the content of IP is both topological and arithmetic and Furstenberg’s proof should not therefore be regarded as impure simply because it appeals to topological principles.

In taking this view, McLarty is aligning himself with the Bourbaki tradition of arithmetic research, a tradition to which Furstenberg’s work also belongs.56 In 1940 Chevalley published a seminal paper on...
class field theory\textsuperscript{57} containing work that he considered arithmetic, despite the central role of topology in it. Chevalley considered the involvement with topology to be \textit{axiomatic} or \textit{set-theoretic} in character, and not of the sort more typical of the Poincaré-Lefschetz conception of topology according to which essential use is made of continua such as the real or complex lines. Chevalley judged topology in the latter sense to be non-arithmetic.

\textit{Historical Digression}

One reason this case is a good one is that its main protagonists, Chevalley and the \textit{Bourbakistes}, recognized how “revolutionary” their views concerning arithmetic practice were, and as a result published explicit comments to this effect. Bourbaki thus described this work in 1948 as showing that, in an “astounding way, topology invades a region which had been until then the domain \textit{par excellence} of the discrete, of the discontinuous, viz. the set of whole numbers”.\textsuperscript{58} They believed this combination of topology and arithmetic to yield a “deeper” understanding of arithmetical problems generally.

Where the superficial observer sees only two, or several, quite distinct theories, lending one another ‘unexpected support’ through the intervention of a mathematician of genius, the axiomatic method teaches us to look for the deep-lying reasons for such a discovery, to find the common ideas of these theories, buried under the accumulation of details properly belonging to each of them, to bring these ideas forward and to put them in their proper light.\textsuperscript{59}

And further:

\begin{itemize}
\item \textsuperscript{57} Claude Chevalley, “La théorie du corps de classes,” \textit{Annals of Mathematics} 41 (1940)
\item \textsuperscript{58} Nicholas Bourbaki, “The Architecture of Mathematics,” in William Ewald (ed.), \textit{From Kant to Hilbert} (Oxford University Press, 1999), §5
\item \textsuperscript{59} \textit{Idem}, §2
\end{itemize}

\textit{Purity Of Methods}

The axiomatic method has shown that the ‘truths’ from which it was hoped to develop mathematics were but special aspects of general concepts, whose significance was not limited to these domains. Hence it turned out, after all was said and done, that this intimate connection, of which we were asked to admire the harmonious inner necessity, was nothing more than a fortuitous contact of two disciplines whose real connections are much more deeply hidden than could have been supposed \textit{a priori}.\textsuperscript{60}

Applied to the specific case of IP, the view would be that Furstenberg’s work reveals previously unnoticed topological and set theoretical elements of IP’s content.\textsuperscript{61}

\textit{End Digression}

60. \textit{Idem}, §7

61. We observe that in many other cases, the dynamics of topical change are not as clear as they are in this case. One could think of the change from Viète to Descartes concerning the algebraic interpretation of multiplication (cf. Henk J. M. Bos, \textit{Redefining Geometrical Exactness} (New York: Springer-Verlag, 2001), chapters 8 and 21). There is a natural sense of dimension for geometrical quantities: a line segment is one-dimensional, while a rectangle is two-dimensional. Viète followed this sense in interpreting multiplication: the product of two magnitudes resulted in a magnitude of higher dimension, so that the product $ab$ designated a higher dimensional magnitude than $a$ or $b$ alone. So, for example, if $a$ and $b$ designated line segments, the product $ab$ was taken to designate a rectangle.

Viète’s algebra accordingly enforced a dimensional “homogeneity” requirement, according to which only terms of the same dimension could be combined in an arithmetic operation (cf. François Viète, \textit{The Analytic Art} (Kent State University Press, 1983), p. 15). In Descartes’ view (as developed in the \textit{Géométrie}), such restriction was not necessary. In his view, the product $ab$ of two line segments $a$ and $b$ was simply another line segment. The advantages of Descartes’ approach over Viète’s were evident almost immediately, and the dimensionality of magnitude soon receded from view.

Since, for Viète, dimensionality and the commitment to homogeneity were foundational to his understanding of algebraic multiplication, it is not unreasonable to see this Cartesian shift as one concerning the topic of algebraic multiplication. Unlike the Furstenberg case, though, this shift was not a matter of reformist fervor but one of accommodation to the power of Cartesian methods. Similar examples from more recent practice could be cited. The identification of topics for such cases may thus depend upon less dramatic methodological shifts than in the Furstenberg case.
Mclarty’s view is thus that, properly understood, the topic of IP makes room for topological elements, and that it is therefore wrong to classify Furstenberg’s proof as impure simply because it makes use of such elements. The point is a serious one, but, in the end, one with which we disagree.

In our view, the conceptual resources that underlie our grasp of IP do not implicate topological elements in the way McLarty suggests. The central truth, as we see it, is this: our basic understanding of IP would not change simply because we were presented with reasons for retracting, say, the definition of a topological space that Furstenberg uses in his solution of IP. In other words, our comprehension of IP—as presented, say, by Euclid in Book IX of the Elements—would not dissolve just because we were presented with a reason to retract the definition of a topological space that figures in Furstenberg’s proof.

This suggests a distinction of some delicacy between certain notions of “depth” and certain notions of “basicness” of proof or problem solution. A topological solution of IP may provide what is in a certain sense its deepest solution without thereby providing its most basic solution. Here by “most basic” we mean something like most rudimentary, that is, possessing in highest degree that quality which solutions have when the conceptual resources they employ correspond most closely to those which are needed to grasp the problem.

Topical purity, as we see it, is intended to achieve rudimentariness rather than depth of solution. This is so, at any rate, so long as depth is conceived as McLarty seems to conceive of it—namely, as a feature of a solution which reflects the fact that the conceptual resources of which it makes use are of such types as might, or perhaps ought, for reasons of simplicity, to be used to solve all, most or at least an impressively comprehensive range of mathematical problems.

Pursuing depth of this sort may be, as Bourbaki has suggested, a means of achieving “considerable economy of thought”.

62. Bourbaki, op. cit., §5

over a pure solution to IP. None of this, however, is reason to identify purity with depth in the present sense, or to regard Furstenberg’s proof as (topically) pure.

6. Conclusion

We have focused on one conception of purity, what we have called the *topical* conception. In our view, the epistemic significance of this type of purity is that it provides a particularly stable means of reducing specific ignorance in directed investigations. Topically pure solutions of problems endure as solutions as long as their directing problems remain the problems they are. This ‘stability’ of pure solutions is what we take to be their primary epistemic advantage. Improvement of knowledge sometimes takes the form of reduction of specific ignorance. When it does, that improvement is better which, other things being equal, more effectively reduces the specific ignorance involved. We have argued that stable solutions of directing problems promise more effectively to reduce the ignorance represented by those problems than do unstable solutions. That this is so is due to the fact that stable solutions, unlike unstable solutions, promise to reduce ignorance through dissolution even should their capacity to reduce it via solution fail.

We have offered a model to explain how topical purity may be a virtue of proof. We have not argued that it is the greatest of all virtues of proof. There are a variety of epistemic virtues that problem solutions may have or lack, and we see little reason to think that single solutions generally realize them all. To the extent that this is so, there may be good reason for the common practice of seeking multiple solutions of the same problem.

63. Idem, §1

Acknowledgements

The authors are grateful to Marcus Giaquinto, Michael Hallett, Paolo Mancosu, Colin McLarty, Marco Panza, Doug Patterson and Jamie Tappenden for useful conversations on various occasions. The first author thanks the Agence Nationale de la Recherche, France (ANR), the Université Denis-Diderot Paris VII, the Université Nancy 2 and the Collège de France for their generous support of his senior chaire d’excellence during part of which he worked on this project. The second author also thanks the ANR for their support during work on this project. The authors together thank the audience at the conference on mathematical understanding held in Paris at the Université Denis-Diderot Paris VII in June 2008. We also thank Ivahn Smadja, the organizer of the conference, and the ANR, who provided funds for it.

References


Bacon, Francis The two bookes of Francis Bacon. Of the proficience and advancement of learning, divine and humane. London: Henrie Tomes, 1605


Detlefsen, Michael “Formalism” In Stewart Shapiro (ed.), Handbook of the Philosophy of Mathematics and Logic. Oxford University Press, 2005

Erdős, Paul “Problem 4065” American Mathematical Monthly 50/1 (January 1943): 65

Fregé, Gottlob “Formal Theories of Arithmetic” In Brian McGuinness


Frege, Gottlob “Formal Theories of Arithmetic” In Brian McGuinness...


Ingham, A.E. The distribution of prime numbers. Cambridge University Press, 1932


Livesey, Steven Metabasis: the Interrelationship of the Sciences in Antiquity and the Middle Ages. Ph. D. diss University of California, Los Angeles 1982

MacLaurin, Colin A Treatise of Fluxions. Ruddimans, 1742


Nathanson, Melvyn B. Elementary methods in number theory. New York: Springer-Verlag, 2000


Saunderson, Nicholas The elements of algebra. Cambridge University Press, 1740


Sylvester, James Joseph “Mathematical Question 11851” Educational Times 59 (1893): 98


Viète, François The Analytic Art. Kent State University Press, 1983 Translated by T. Witmer

A Treatise of Algebra, both historical and practical: shewing the original, progress, and advancement thereof, from time to time, and by what steps it hath attained to the height at which it now is. London: John Playford, 1685