Why Our Hand is not the Whole Deck: Embrace, Acceptance, or Use of Limitations

Robert S.D. Thomas
St John's College and University of Manitoba

Follow this and additional works at: https://scholarship.claremont.edu/jhm

Part of the Arts and Humanities Commons, and the Mathematics Commons

Recommended Citation

©2020 by the authors. This work is licensed under a Creative Commons License.
JHM is an open access bi-annual journal sponsored by the Claremont Center for the Mathematical Sciences and published by the Claremont Colleges Library | ISSN 2159-8118 | http://scholarship.claremont.edu/jhm/

The editorial staff of JHM works hard to make sure the scholarship disseminated in JHM is accurate and upholds professional ethical guidelines. However the views and opinions expressed in each published manuscript belong exclusively to the individual contributor(s). The publisher and the editors do not endorse or accept responsibility for them. See https://scholarship.claremont.edu/jhm/policies.html for more information.
Why Our Hand is not the Whole Deck: Embrace, Acceptance, or Use of Limitations

Cover Page Footnote
I acknowledge with thanks the encouraging discussion of the first ancestor of this paper (something like Sections 1 and 4) by Staffan Angere, Leon Horsten, James Ladyman, and John Mayberry. Also of the second version by Alex Paseau and Dan Isaacson (also for telling me about Leslie Tharp). Inspiration for this version came from reading [32] and [2]. Jean-Pierre Marquis has helped. I owe the Pycior reference to Chris Hollings, the Boole reference to Dirk Schlimm, awareness of Kohlenback to Douglas Bridges, and the Guinée and l'Hôpital references to Jeanne Peiffer. Thanks also to Douglas Bridges, Julian Cole, Benedict Eastaugh, Hardy Grant, and Hugh Thomas. Section 5 was presented at the annual meeting in Regina of the Canadian Society for History and Philosophy of Science in May, 2018, and Sections 2, 3, and 6 were presented at the annual meeting in Montréal of the Canadian Society for History and Philosophy of Mathematics in June, 2018.

This report is available in Journal of Humanistic Mathematics: https://scholarship.claremont.edu/jhm/vol10/iss1/12
Why Our Hand is not the Whole Deck: Embrace, Acceptance, or Use of Limitations

Robert S. D. Thomas

Department of Mathematics, University of Manitoba, Winnipeg R3T 2N2, CANADA
robert.thomas@umanitoba.ca

Synopsis

I investigate with examples limitations that mathematicians accept in research: aesthetic, scientific, and practical, in particular, on what bases such choices are made. Discussion is partly in terms of the ideal agents that Philip Kitcher and Brian Rotman used to analyse mathematical writing.

1. Introduction

There is an insulting expression “not playing with a full deck,” which one would hardly apply widely to research mathematicians. But it is common in research to accept limitations, sometimes severe, on the means to be used. This report is the result of my inquiring why this should be so, initially thinking of constructive mathematics but with the topic of the final section also in mind. I set out to find facts, and I am intending to indicate what I

---

1 I acknowledge with thanks the encouraging discussion of the first ancestor of this paper (something like Sections 1 and 4) by Staffan Angere, Leon Horsten, James Ladyman, and John Mayberry. Also of the second version by Alex Paseau and Dan Isaacson (also for telling me about Leslie Tharp). Inspiration for this version came from reading [33] and [3]. Jean-Pierre Marquis has helped. I owe the Pycior reference to Chris Hollings, the Boole reference to Dirk Schlimm, awareness of Kohlenback to Douglas Bridges, and the Guisnée and l'Hôpital references to Jeanne Peiffer. Thanks also to Douglas Bridges, Julian Cole, Benedict Eastaugh, Hardy Grant, and Hugh Thomas. Section 5 was presented at the annual meeting in Regina of the Canadian Society for History and Philosophy of Science in May, 2018, and Sections 2, 3, and 6 were presented at the annual meeting in Montréal of the Canadian Society for History and Philosophy of Mathematics in June, 2018.
found rather than to prove any point. I have found three different kinds of limitation that mathematicians accept in research, finding different reasons why they are acceptable to some. These kinds and reasons are aesthetic, scientific, and practical. Making no claim to be exhaustive, I hope that others will be interested enough to find other kinds and more easily other examples of my kinds.

An interest that I find in these limitations is the reasons for them, and I am going to use the word “justify” for the relation between a reason and a limitation. I am specifically not intending philosophical justification, whatever that would mean. I mention philosophical justification, however, because Helen Billinge [5] shows that the constructivism of Errett Bishop [6, 8] is not philosophically justified by him, and much the same question is still being addressed about constructivism [18].

In order to discuss limitations, it will be convenient to use the a term introduced by Philip Kitcher [28] and refined and elaborated by Brian Rotman [46]. A lot of mathematics concerns what can be done. The classic example is the use of straightedge and collapsing compasses in Euclidean geometry. Humans cannot draw Euclidean lines or circles, and one can think of an ideal agent that can do these things with the required (exact) precision. Such agents are plainly non-human and disconnected from the mathematical reader (or writer for that matter). A writer, on the other hand, may have an ideal engaged reader in mind, to whom sentences like “Consider triangle $ABC$” are addressed. It is not the non-human agent that considers the triangle; its job is to draw it, and it responds to its instructions. I shall mention this ideal reader occasionally; assumptions may be made of the reader by the writer. Both receive instructions in a typical piece of mathematics. A piece of intuitionistic mathematics may be written by a non-intuitionist assuming that the engaged reader is an intuitionist and so cares about intuitionistic proof. It is the presumed reader that needs to be an intuitionist,

\[ \text{Justification is used with different meanings in two famous sayings from the Renaissance, “Justification by faith alone” from Martin Luther and “justify God’s ways to man” from John Milton. Luther’s justification is making one to be just or righteous, a divine action consequent upon faith, whereas Milton’s justification is only to show to men that God’s ways are just not to make them just. The difference is between the act of making right and demonstration of being right. This difference is so large and there are so many sorts of justification in between that the notion needs a book not a footnote.} \]
not the writer. Restrictions are placed on presumed readers and on ideal agents in much mathematical writing. Such restrictions have always been part of geometry, the only old mathematics sufficiently formalized to consider in this way. The early books of Euclid’s *Elements* allow the agent the use of only straightedge and collapsing compasses in a context where famous problems like the duplication of the cube could be solved by other means. Having been done for ever is a reason for this particular restriction and potentially for other such restrictions. A significant amount of mathematics is the fairly systematic exploration of the effects of such restrictions, often by choice of axioms, which usually restrict the agent, and of reasoning style, which usually applies to the reader. This seems to be how we explore the wisdom of time-honoured restrictions.

In the *Elements*, Euclid allows the agent to draw circles, given centre and a point on the circumference, and to join points with straight lines and extend them. When you think of what you can do with an idealized pencil, that is amazingly restrictive. And those restrictions have to be relaxed in the later three-dimensional books. It is customary to misstate these capacities of the agent in terms of existence. “There is a line between or through two points,” and “there is a circle through a given point with a given centre”. But the constructions require not just that these things exist but also that they can be drawn. A goal of the *Elements* is the polyhedron constructions of Book XIII.

I think that restrictions on the reader alone were not contemplated until the twentieth century. Brouwer both restricted the reader with intuitionistic logic and allowed it (she or he?) to produce choice sequences, rendering his intuitionism incompatible with mainstream mathematics. As well as a philosophical stance, Brouwer gave rise to intuitionistic mathematics, which one need not be an intuitionist to find interesting to study. There is the example of Philip Scowcroft [47, 48], who both studied intuitionistic mathematics

---

3Rotman calls the engaged reader the “subject” in his detailed analysis of mathematical writing. According to Stephen Pollard [41, page 92], Moritz Pasch as long ago as 1919 [36, page 21] wrote in the as-if mode of Hans Vaihinger [62] of an idealized combinatorial reasoner blessed with “eternal life and unlimited memory” with which to discuss the possibility of certain constructions as expressed by his axiom systems [37, 38].

4See [49] for a careful examination of the agent’s restrictions in Book I–VI of the *Elements*. 
without being an intuitionist and could make it interesting to an audience of non-intuitionists (personal experience). One is not bound to mimic the engaged reader to read a text any more than one inherits the powers or limits of the implied agent.

In the twentieth century both Turing machines and actual physical digital computers have been studied as obviously restricted agents. One scientific interest in agents is the equivalence of powers of differently described agents. As James Ladyman pointed out to me, a reason for so much interest in Turing machines is that what they can do is provably equivalent to various other agents’ capacities. Such equivalence classes have the characteristics of natural kinds, perhaps more so than natural kinds. Likewise, what can be found by constructions with straightedge and collapsing compasses in the Euclidean plane is the same as what can be found with arithmetic and square roots in the Cartesian plane. Again the equivalence-class boundary has been studied a great deal with things proved to be inside or outside that common capacity. Computer science is much concerned with the limitation embodied in Turing machines; this seems to be an example of scientific interest in and justification of those limitations (in the narrow sense I in which I am using “justification”).

More generally, restrictions on what the implied agent can do with the mathematical objects involved, usually specified by axioms, and on the logic the engaged reader can use are standard in the past century or so. In the early nineteenth century an agent could manipulate symbols, but the symbols had to represent numbers unless you were an advanced thinker like George Boole [9]. As an editor, I run into the restriction to first-order logic all the time. And group theory is a restriction of field theory; the agent can add or multiply but not both.

2. Aesthetics

I find a tension between two aesthetic attitudes to mathematics associated with G. H. Hardy. He maintained that a pure mathematician is observing a given landscape and notes down observations (proof is successfully pointing to a feature) [20, page 18]. But on the other hand he also said that there is no permanent place for ugly mathematics [21, page 85]. One can hardly demand that the unalterable landscape one is observing has any particular
attributes. As I have said elsewhere [58], the impermanent may last a very
long time if something ugly is sufficiently important. But Hardy's ugliness
or beauty is that of the written/spoken/performed mathematics, not of the
landscape whose exploration is written up. There is fine travel writing about
unpleasant sights. It is unusual for the mathematical objects themselves to
be of interest for their beauty. An example is my (and others') study of weav-
ing patterns [55, 56]. Escaping the landscape metaphor, I suggest that the	
tension is between the curiosity and aesthetic motivations of the enterprise. If
one wants beautiful definitions, axioms, proofs, then one's motivation is aes-
thetic, which may limit the landscape's exploration. But if one explores the
landscape at all costs, then the costs may be aesthetic. These two projects
have quite distinct motivations beforehand and rewards afterwards. I want
to move now to the aesthetic way of looking at the matter, setting aside from
consideration that mathematics, as exploration and observation, satisfies cu-
riosity directly. This is to consider pure mathematics as an outlet for artistic
creativity satisfying aesthetic appreciation rather than the raw curiosity of
normal mathematical research, which of course it also satisfies.

As I said above, a non-intuitionist can require an intuitionist implied reader
that operates with intuitionistic logic. And a constructivist can recognize
a valid classical proof that uses the law of excluded middle (EM). In his
philosophical manifesto, Bishop [7] is quite clear about needing agency.

Our point of view is to describe the mathematical operations that
can be carried out by finite beings, man's mathematics for short.
In contrast, classical mathematics concerns itself with operations
that can be carried out by God. ... The most solid foundation
available at present seems to me to involve the consideration of
a being with non-finite powers—call him God or whatever you
will—in addition to the powers possessed by finite beings. [7,
page 9]

Such an agent has always been necessary in geometry in view of our drawing
limitations, and it has been called upon in analysis more or less since its
invention. Euclid's *Elements'* agent selects by "drawing" circles and line
segments, selecting them from the infinite plane surrounding them. A far
cry from the axiom of choice, but still an infinite task. This is something
that Bishop does not seem to recognize. He ignores geometry and considers
Why Our Hand is not the Whole Deck

some classical analysis to be human mathematics supplemented by EM and its consequence the limited principle of omniscience (LPO), “If \( \{n_k\} \) is any sequence of integers, then either \( n_k = 0 \) for all \( k \) or there exists a \( k \) with \( n_k \neq 0 \),” discoverable by the agent. To give a classical theorem \( A \) what Bishop calls its “meaning,” he would say “LPO implies \( A \)” if LPO but not EM has been used in its proof or “EM implies \( A \).” I’m not much concerned with the practicalities of constructivism. What interests me here is motive. It seems to me that Bishop’s reason for pursuing the style he adopted was aesthetic. I formed this impression from his book and then had it confirmed by this sentence in his manifesto. “My feeling is that it is likely to be worth whatever extra effort it takes to prove ‘\( A \)’ rather than ‘LPO \( \rightarrow \) \( A \)’.” [7, page 14] This is aesthetic because it is an ascription of worth and is based on feeling but is about something objective, namely the difference between the proofs of \( A \) and of \( EM \rightarrow A \).\(^5\)

It had seemed to me that Bishop’s restriction of means, to the extent that it was reasonable, had to be aesthetically motivated, and so I was pleased to find that he said so. It was a matter of choice of style. In his polemical writing Bishop made out that it was more than that, as is indicated by the title of what I have referred to as his manifesto, “Schizophrenia in contemporary mathematics.” [7] The agent is allowed by some mathematical styles to complete infinite processes. This appears to be something Bishop was reacting against. I take his restrictions to be not allowing the agent to complete infinite processes nor the reasoning to use excluded middle. Reasons for these restrictions appear to be

1. Compatibility with, indeed inclusion in, mainstream mathematics and

2. The feeling, at least of Bishop’s followers, that what is accomplished within these voluntary restrictions is preferable to what is done outside—rather like some others’ avoidance of the axiom of choice.

The second reason is a good example of the aesthetic basis for some limitations that mathematicians place on themselves in their research. Had that been acknowledged instead of presenting constructivism as the cure for

\(^5\)This is by no means to claim that Bishop thought of increased meaningfulness as an aesthetic quality. He wanted the greater meaning. The reader may prefer Bishop’s stated reason to my version.
the maladies of mid-twentieth-century mathematics, it might have received
a better hearing. Most mathematicians did not feel diseased and were not
looking for a cure. It is a long time since [6], and in that time constructiv-
ivism has proved itself mathematically, as illustrated by Brian Davies [13],
not himself a constructivist.

3. Style: The musical analogy

If one considers a musical analogy for mathematics [3], it will not be as simple
as tunes for a penny whistle. More like instrumental music as a whole. As
soon as one looks at a whole art form like music or architecture, one realizes
that restrictions are inevitable. Just as when building a building one needs
a site and materials, if one is to write music one needs to consider what
medium one is writing for. Bach’s Art of the Fugue is unusual in being, I
understand, just notes (playable on a keyboard). Almost all music is written
for particular forces.

The musical analogy here is just that. It may be shallower than that to fic-
tion. I am not suggesting that there is much in common between composing
music and doing mathematical research. I have thought of just a few analo-
gies that I think are interesting. There seem to be a couple of similarities
where music has a name but mathematics does not. Genre is defined usually
by the instruments one writes for and then more finely. This corresponds
in mathematics to those choices such as Bishop’s to limit what one uses. In
both art forms such decisions do have to be made.

The vast differences between music and mathematics prevent any idea that
thinking about them can be blended. Music is about sound and mathematics
about ideas. The philosophical difference between them includes the freer
creative nature of musical composition, whereas mathematicians have this
idea that they are mapping terrain to which they have some mysterious
access. Even those more impressed negatively by the mystery than positively
by the access have to admit that the notion is borne out by one’s being able
to take up the work of others and extend it seamlessly. That does not happen
in music. It is serious work to take the gems of several mathematicians and
arrange them mosaic-like in a treatise,\(^6\) but it is done, whereas composers do

\(^6\)Mario Bunge [11, page 111] (quoted [35]) points out that mathematicians do not put
not weld smaller compositions of others into a symphony. Charles Avison’s arrangement of Domenico Scarlatti’s harpsichord sonatas into concerti grossi is a rare exception, and they are all by Scarlatti. What does get put together is mostly folk melodies.

Another similarity is again where a musical term is standard and mathematics has only a tenuous grip on the idea. A piece of music is often arranged for forces different from the original. For example a symphony can be reduced to be played on the organ (for example, Dvořák’s New World Symphony). The pieces of music are different but related in an obvious but imprecise way. Mathematicians sometimes prove a theorem from systematically altered hypotheses. It is not the same theorem, but it may have a proof that one would wish to call “the same.”\(^7\) We have no standard term for doing this when it is not generalizing.

There is some similarity between harmony in music and rigour in mathematics. The one that came instantly to my mind was that what is harmonious and what is rigorous are far from constant over time. But being harmonious is a restriction on what one can do as being rigorous is a restriction. In my experience, only hymn tunes are rewritten to make them more harmonious. After a century or so of much music without old-fashioned scales, the restriction to standard keys has not disappeared but remains. If intuitionistic logic had taken over, it would have been a rejection of the more freewheeling logic of the twentieth century not altogether unlike a rejection of atonality, which has also not happened.

Mathematics and music on the page are just symbols. In order to be what they symbolize they both need to be lifted from the page by a person or by persons in the case of compositions for a group of players.\(^8\) It seems to me to be easier to lift mathematics mentally than to do justice to a musical score by reading it, but a musician might feel oppositely. Both mathematics and music are best performed and need performance to be appreciated. This is of course well-known for music, but I think that it is underappreciated

---

7Brian Davies [14] mentions generalizing Mercer’s theorem from kernels on a one-dimensional interval using “the core of Mercer’s argument.” The “idea” of such a proof is like a melody that can be moved around, even turned upside down.

8This has been noted by others, e.g., Marquis [35].
for mathematics. Mathematicians do themselves and their mathematics no favours by paying so little attention to the fact that presentation matters. Incomprehensible colloquia are inexcusable.

There is no question that in music—in principle—anything goes. One can choose whatever forces one wishes and specify whatever notes are to be played by them (subject to only physical limitations). But there are good reasons for choosing combinations that “work,” a number of which are traditional. An exact example: how many different instruments could replace the cello of a string quartet and still produce sounds that many would be prepared to listen to more than once? I heard the premiere and only performance of Canadian composer Lawrence Ritchey’s duet for flute and bagpipe. He wrote it as a joke, of course, because while the instruments could be played antiphonally, one could not hear the flute if the bagpipe was playing too. When axioms began to be taken seriously outside geometry in the nineteenth century, Augustus De Morgan [15], following George Peacock [42] and followed by George Boole [9], said that—in principle and in algebra—anything goes. The idea of arbitrary axioms is accordingly very old as a logical possibility, but no one is interested. The idea of arbitrary axioms is plain silly, which does not restrict freedom.

Turning back to the matter of mathematical genre, if one thinks of that analogy, one no more needs to apologize for working for one set of instruments or from one set of axioms than from another. The only thing that matters is whether the mathematics produced in the genre is of value, chiefly, in my opinion, of interest. A whole genre, even an important one, may be of no interest to someone. The widespread lack of interest in set theory among mathematicians in general has somehow generated a vague hostility to set theorists. This cries out for study and understanding, not to mention correction. Bishop showed that with his forces he could do a lot of interesting mathematics. As a result his constructivism has or should have greater respectability than Brouwer’s intuitionism, from which Bishop distanced himself, especially in view of his offering classical proofs preferable at least aesthetically to non-constructive proofs.\footnote{Beyond the aesthetic value of constructivism, constructive proofs have a scientific value as informative. Benedict Eastaugh has pointed out to me that in proof theory Kreisel had a programme that is called “unwinding” in which nonconstructive proofs are turned into constructive ones to extract computational content, “allowing one to compute explicit}
that constructivism, for example, should ultimately be the norm. Bishop probably did want to do that, and he couldn’t. It was Bishop’s inclination not his mathematics that lacked not just philosophical justification but any justification. The very idea that his style had or required justification comes from his mistakenly seeking it.

4. Scientific Reasons for Limitations

A reason that has been given for some restrictions, for instance that to three dimensions in geometry, is that mathematicians explore a given (though inaccessible) Platonic realm sending back reports of what is discovered. This sort of realism, associated with Frege, imports to mathematics whatever prejudices are operative in its practitioners. While this can be presented as philosophical justification for specific axiom systems (all of which limit in some way [20, page 19]), its historical effect has often been to block the exploration of what has been deemed outside the realm. Such attitudes hampered the study of negative and complex numbers for centuries and probably geometry of more than three dimensions. Intellectual progress is better directed by mathematical and scientific (i.e., broadly pragmatic) imperatives. The idea that the restriction to two and three dimensions is not a restriction but is in any sense natural for geometry has now been out of date for two hundred years; they are, to be sure, natural foci of interest. That’s what is natural. Two-dimensional geometry is easier than three-dimensional geometry, easier in turn than four-dimensional geometry; so there is a pedagogical progression that is also as natural as inclusion.

The applied-mathematical example that raised the question of limitations in my mind is the ancient text *Spherics*, which exists only in a second-century BCE version [23, 12, 63, 60], which date just indicates when it stopped being witnesses or bounds on the growth of a function from a result that was apparently purely existential in character” (personal communication, May 4, 2018). A leading exponent of this programme is Ulrich Kohlenbach [29].

10 In the preface of his [6], Bishop referred to his “ultimate goal—to hasten the inevitable day when constructive mathematics will be the accepted norm.” He did not claim, however, that “idealistic mathematics is worthless from the constructive point of view” because “[e]very theorem proved with idealistic methods presents a challenge: to find a constructive version, and to give it a constructive proof.” (page x)
worked on because it was superseded, except as introductory, by the invention of trigonometry. The treatise is centuries older, being called upon in Euclid’s *Phenomena* [4] and probably *Elements*. *Spherics* is quite strictly deductive but appears to be pre-axiomatic since no axiom is stated or mentioned. The deductions help themselves to whatever results of plane or solid geometry are needed, so long as they do not pertain to spheres. The subject is the parallel circles on the celestial sphere that are paths of stars and some great circles including the reader’s horizon and the ecliptic, that is, the daily path of the Sun. Books II and III contain this material applicable to the spherical astronomy of the stars (no planets).

The constructions used in their proofs require occasionally, in order for the agent to draw parallel circles, its finding the pole of a circle (i.e., the point on the surface used with compasses to draw the circle) and frequently drawing a great circle. In preparation for these activities of the treatise’s agent, Book I shows how the agent can do both of these things with compasses and straight-edge, almost like *Elements* I (but the compasses must not be collapsing as they are used to transfer distances). What matters here is how Book I proves things. The agent practises, as it were, the constructions to be used in Books II and III as one would expect, because the constructions have to be proved effective. The reader meanwhile performs what my friend Bob Alexander characterized for me as thought experiments, virtual constructions, which need to be thought about but need not be performed, allowing the reader to see geometrical facts. The first proposition is as good as any to illustrate this.

On the surface of a sphere, the plane through three points $A, B, G$—determining a plane according to the *Elements*—cuts the surface of the sphere in a circle. To prove this, the reader imagines (as I am conceptualizing what is going on) a perpendicular dropped from the centre $C$ of the sphere to the plane at $D$ and sphere radii $CA$ and $CB$ and what will turn out to be circle radii $DA$ and $DB$. In right triangles $ACD$ and $BCD$, $CD$ is common and the sphere radii $CA$ and $CB$ are equal; so the third sides $DA$ and $DB$ are equal to each other and to $DG$, which is any other radius of what is by now obviously a circle. While there is no way in space to join $CA$ much less to drop a perpendicular from $C$ to $D$, this is all imagined by the reader with no action taken by the agent with the compasses, but that’s good enough it seems to me. One can discuss this in terms of two agents, one to do these improbable and unnecessary things [50, 57] while saving the other to wield compasses.
Why Our Hand is not the Whole Deck

and straightedge later as required. In that case one has unrestricted and restricted agents (imagination and construction agents). In my view the reader has to have imagination, but otherwise has no restrictions at all. The point being made here is that, whichever way you conceptualize the agent or agents, restrictions on agents have always been with us, since those in the *Elements* are well known.

(This is what the more restricted agent does. The basic construction is to find the diameter of a circle on the surface of the sphere, given three points on it, say $A, B, G$. It cleverly makes use of the fact that a circle is a planar figure as well as a spherical figure and transfers the distances $AB$, $BG$, $GA$, to a plane as $ab$, $bg$, $ga$ with the non-collapsing compasses acting as dividers. It draws the triangle $abg$ with the straightedge and constructs perpendiculars at $b$ and $g$ inward to meet at $d$, which is the other end of diameter $ad$ of the circumcircle of the triangle $abg$, equal to the circle on the sphere. Line segment $ad$ is a diameter because a diameter is what subtends right angles, and $ad$ subtends right angles at $b$ and $g$.)

Restrictions on agents are so old a feature of mathematics as to be (in part) definitive of it. Yehuda Rav points out in his famous paper “Why do we prove theorems?” that we would not need to prove theorems if we could somehow be told which statements are true. Ignorance is the great given human limitation. Without chosen limitations, mathematics would be reduced by our physical limitations to adding and multiplying whole numbers. The restriction “Smaller from larger” is introduced with subtraction.

*Spherics* illustrates one sort of restriction of mathematical work—to the limited mathematics needed for natural science. I now give a second example

---

11 Two agents of differing powers may be more common than the Kitcher–Rotman analysis suggests. If one is a numerical analyst considering rational arithmetic, one is going to do so in a context of real numbers, that is, where one can call upon an agent to make the sum of root two and root three meaningful. Or if one is going to prove that one cannot trisect an angle with arithmetic and square roots, one is going to think in the standard Cartesian plane and just restrict to the determinable points. An increasing amount of mathematics has to do with such matters, which can be thought of in terms of non-absolute Tharp modality, discussed shortly.

12 It is of course agency that is plainly present in the *Elements*, joining points, extending lines, and drawing circles, with no agent mentioned. Andrei Rodin’s recent discussion of the *Elements* makes much of agency but without an agent.
because it is so very different and yet illustrates scientific reasons for using weaker than normal axioms. Weaker axioms authorize weaker agents.

5. Scientific Interest in Mathematics Itself: Reverse Mathematics

I have already mentioned the study of agents of power equivalent to Turing machines. Other agents can be studied too. In the simplest model of ordinary mathematics, one takes axioms and a style of inference and deduces theorems from the axioms. Number theory is in principle the collection of theorems one can deduce from the first-order Peano Axioms (PA), where the induction axiom is actually a scheme of many axioms, one for each formula $\varphi(n)$:

$$[\varphi(0) \land \forall n (\varphi(n) \Rightarrow \varphi(n + 1))] \Rightarrow \forall n \varphi(n).$$

This system is an example of axioms of the self-evident variety, which we use because we believe in them. The language of PA includes the constant 0, variables, function symbols, =, logic symbols including universal and existential quantifiers, and parentheses with which to make terms, formulas, and equations. For example, subtraction of less from greater is possible:

$$(\forall \ell)(\forall g \geq \ell)(\exists d)\ell + d = g.$$ 

We call $d$ the difference $g - \ell$; it is something that the agent here can find. For what I want to say, the quantifiers are important because it is possible to recast well-formed statements into a different-looking special form. The special form, unlike what one might call the more natural form I’ve set out above, has all the quantifiers at the beginning alternating between universal ($\forall$) and existential ($\exists$) so that there are two forms possible beginning with either a universal or existential quantifier. If there are $m$ quantifiers, the first class is called $\Pi^0_m$ and the second $\Sigma^0_m$. The size of $m$ is a measure of the complexity of the property defined by the $m$ quantifiers and the quantifier-free formula at the end. Unsurprisingly, there are properties definable in $\Sigma^0_{m+1}$ that are not definable in $\Sigma^0_m$. We are now able to return near to Bishop because

\[\text{Reverse mathematics has also been done constructively. That and other aspects of constructive mathematics are discussed in [10]. Incidentally, Bishop was very much in favour of formalizing his mathematics for two reasons, for humans to understand it more precisely and for computers to be able to do it at all [7, page 14].}\]
sets of integers that are computable are defined by $\Sigma^0_1$ statements. One of Bishop’s ideas seems to have been to stick to the computable, although he hampered critics by failing to specify what he meant by “computable.” In terms of agents, Bishop wanted them to be of human powers; exactly what that entails needs to be made more precise to be studied. The quantifiers seem to have interpretation in terms of agents. The existential quantifier for Bishop must mean that his agent can find the item quantified, and the universal quantifier must indicate “for any that the agent can find.”

Very many mathematicians are interested in doing mathematics way beyond what is in any sense computable, and still others are scientifically interested in studying that mathematics for its properties. One of the simpler such properties is just how far away results of analysis are from being computable. Students of reverse mathematics use several axiom systems, of which one is PA with two changes. The induction scheme, symbolically identical to that of PA, is restricted to apply only to $\Sigma^0_1$ formulas $\varphi(n)$,

\[
[\varphi(0) \land \forall n (\varphi(n) \Rightarrow \varphi(n + 1))] \Rightarrow \forall n \varphi(n),
\]

and there is added, again only for $\varphi(n)$ in $\Sigma^0_1$, what is called recursive comprehension:

\[
\forall n (\varphi(n) \Leftrightarrow \psi(n)) \Rightarrow \exists X \forall n (n \in X \Leftrightarrow \varphi(n)),
\]

where $\psi(n)$ is a $\Pi^0_1$ formula and $X$ is not a free variable in $\varphi$. This means that if a property is definable in both of the forms I have mentioned, $\Sigma^0_1 \Pi^0_1$, then there is a set of natural numbers satisfying the formula. The relevant agent can pick out that set. RCA$_0$, as this system is called,\footnote{15} is a very restrictive system in which little can be proved; not, for instance, that a monotone bounded sequence of rational numbers has a limit (monotone-convergence theorem [53, page 82]). One might easily say that it is not mathematically justified, but it is scientifically justified by what else can be done with it.

\footnote{14}{If it is fair to say that the intuitionists find the constructive concept of a sequence generated by an algorithm too precise to adequately describe the real number system, the recursive function theorists on the other hand find it too vague.” [7, page 20].}

\footnote{15}{RCA$_0$ is apparently not an abbreviation. RC of course abbreviates “recursive comprehension.” Writers in this subject use this symbol and another I’ll mention as system names. At least one other system has as its name a symbol (WKL) that is an abbreviation of something that is not an axiom system; so WKL is definitely not the abbreviation of the name of the system.}
While on its own, among well-known theorems of analysis, it proves only the intermediate-value theorem, it allows proofs that various other theorems are equivalent to each other. For example, the monotone-convergence theorem is equivalent [53, page 82] to the Bolzano–Weierstraß theorem and the extreme-value theorem is equivalent [53, page 107] to the Heine–Borel theorem. So it is useful for comparing the strengths of theorems it cannot prove. The study of reverse mathematics, which is only forty years old, has also used stronger axiom systems than RCA. Stronger systems are particularly important, including ACA, in which many theorems of analysis can be proved, and another (WKL) between RCA and ACA in strength and characterized by an axiom equivalent to the Heine–Borel theorem. ACA and some others are characterized by stronger comprehension schemes than the recursive. In each of them some theorems can be proved and some pairs, unproved, can be proved to be logically equivalent to each other. I don’t want to belabour this topic, but I do hope to have made my point that an axiom system lacking mathematical justification can have justification for the scientific study of mathematics. I shall finish these details with two remarks. One is that the equivalence classes of theorems are surprisingly large. Lastly, my chief sources for this material are a paper that I read before its publication in Philosophia Mathematica [16] and a recent readable

---

16 An infinite set of numbers between two given real numbers has a limit point [53, page 57].
17 Every continuous real-valued function on a finite interval has both a maximum and a minimum on the interval [53, page 60].
18 Every closed finite real interval is compact [53, page 59].
19 This scientific interest in mathematics seems to have been initiated by the problem of the parallel postulate. Euclid’s other postulates enabled the proof that the parallel postulate (one and only one parallel) is equivalent to Playfair’s axiom (at most one parallel, 1795). I owe this remark to Benedict Eastaugh. Such interest must have been stimulated by Sierpiński’s work on the axiom of choice, showing what needed it and what was equivalent to it in many papers.
20 ACA has the simpler but stronger \textit{arithmetical comprehension} scheme:
   \[ \exists X \forall n (n \in X \leftrightarrow \varphi(n)), \]
where \( X \) is a set of natural numbers and \( \varphi \) is definable in the language of PA.
21 Other theorems equivalent to Heine–Borel over RCA are that the limit of a sequence of uniformly continuous functions on a closed interval is continuous, Brouwer’s fixed-point theorem, the Jordan curve theorem, and the separable Hahn–Banach theorem.
book by John Stillwell [53], to which I have referred, the standard reference being [51].

As usual, axiom systems clarify what we can say about our material without saying what we are talking about, and these axioms are different from those useful in natural science. But we can say more. I draw your attention to two different situations.

1. Reverse mathematics studies pre-existing carefully axiomatized mathematics to determine, for a result studied, exactly which hypotheses will not just give a result but are equivalent to it. This is completely precise and is a scientific justification for the mathematical axioms used.

2. The second is less precise, and so I’ll use an analogy used by Frege [17] and Hardy [20, page 18] for the result of the mysterious access to the terrain of pure mathematics: one is sending back reports of mountain peaks that one has spied in the distance. In terms of that metaphor, much constructive mathematics studies just which peaks can be not just seen but climbed by showing how to climb them.

One way to look at reverse mathematics is as determining how far from being climbable various peaks are, how far from being constructive various non-constructive theorems are [53, page 156]. The aim is a functioning climbing map, a map of what can be climbed and with what equipment. Scientifically, we are interested in what it is possible to do. Mathematically, we are interested primarily in doing it and secondarily in determining what in principle is possible for various agents. How does all this possibility fit in?22 Consider a quotation from Leslie Tharp, who died in 1981.

We have claimed that the modal propositions of arithmetic are primarily about concepts, and are about ordinary objects and activities in the indirect sense that the concepts may be applied to ordinary objects arising from ordinary activities, such as an actually constructed inscription. In particular, existential assertions such as “there is a number…” may go far beyond anything humanly feasible. The discomfort with modal treatments of mathematics is reminiscent of the everyday interchange of “can” and

22I can’t help mentioning the modal structuralism of Geoffrey Hellman [24], but this is not the place to discuss it.
“may.” One sometimes says “Herr Schmidt can drive 150 kph on the Autobahn” when he actually cannot (because, say, his Volkswagen won’t go that fast). Obviously, what one means is that he may, that is, the relevant rules permit such speed. We interpret the mathematical modalities in such a “may” sense: one may construct an inscription with $99^{99}$ strokes—the concepts undeniably permit it. [54, page 187]

Herr Schmidt is a Kitcher–Rotman agent.

I have observed a twentieth-century shift in mathematics toward consideration of possibility and efficiency, what can be done and how well. The shift seems to have begun very early in the century, since already the Académie Française offered a grand prize in 1918, won by Gaston Julia, for an investigation of global properties of iterations of rational mappings in the plane. Their exotic behaviour, now known as fractal, had already been written about by Pierre Fatou in 1906. Emphasis on what can (and of course cannot) be done takes mathematics back to its roots, if not in Europe, then elsewhere. Much of our evidence of pre-Greek middle-eastern mathematics is of problem-solving processes. Indian mathematics was focused on calculation. Chinese mathematics was algorithmic when the ancient Greeks were thinking instead of theorems; Chinese calculators invented matrices to solve linear equations about two thousand years before Europeans [22]. Arabic mathematics in the middle ages, despite being based on Greek theory, was sufficiently algorithmic to have given us the term. Once you shift your attention to how to do things rather than reporting on geography, how well you can do what you can do becomes of mathematical interest and studying it becomes justified by the practical reasons for doing what you do, solving the problems that you solve.

6. Practical Reasons

There is another kind of narrowing with reasons, which is a looser choice than a restriction that anyone would try to suggest as imperative. This kind

\footnote{J. M. Landsberg [31, 32] has written about more efficient algorithms to multiply $n \times n$ matrices. The definition requires $n^3$ multiplications, but the number of arithmetic operations has been reduced to $\mathcal{O}(n^{2.38})$.}
of choice is that of a tool to use for whatever ends rather than a topic to study by whatever means. (It is also a topic for study.) My observation of mathematical practice is that this adoption of a tool to use is common and is done with no thought that its use needs any sort of justification other than its working.\textsuperscript{24} One does with it, in each application, what one can; perhaps others will use it or something else more effectively later. One will move on to another application or to learning how to use some other tool. Substantial investment is involved; one can’t learn to use everything and then choose what to do with it all. This is analogous to physicists, who have built a cyclotron to perform certain experiments, continuing to use their cyclotron to do other things simply because they have it and are experts on its use. Not all momentum is physical.

Jean-Pierre Marquis, in a paper twenty years ago [34] and subsequently, has discussed mathematical tools and machines, using as his main example in [34] K-theory as a tool in a context of mathematical gadgets of two sorts, instruments and tools. I have seen the value of this by thinking about examples that I understand better than his. While tools manipulate things, I take instruments to be mainly denoters that say what you’re looking at (even if you don’t know exactly what it is). Letting \(x\) be the unknown quantity is an important instrument. So I have two lists illustrating my interpretation of this distinction; see Table 1 on the following page.

The last example is not a generally known technique but is mentioned by Emiliano Ippoliti [27] (not knowing I studied it fifty years ago).

Since what I am calling an “instrument” specifies and represents what one attends to, and what I’m calling a “tool” manipulates things identified by a corresponding instrument, in order to use a tool one needs to understand its use, and one also needs to have things represented in an appropriate way. The tools that Marquis lists are more sophisticated than mine, “spectral sequences, homology theories, cohomology theories, homotopy theories, K-theories, sheaves, schemes, representation theory and character theory, commutative algebra, graph theory in group theory, classical geometry in algebra,

\textsuperscript{24}Knowing about Coxeter groups one might see how that knowledge can be applied to solve problems about quivers. Or knowing about quivers one might apply them to cluster algebras. As these examples illustrate, the same body of work can be either tool or target. And of course learning can be of the tool as well as of the target.
<table>
<thead>
<tr>
<th><strong>Instruments</strong></th>
<th><strong>Tools</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>lettered diagram</td>
<td>straightedge and compasses in Euclidean geometry</td>
</tr>
<tr>
<td></td>
<td>rotation of plane curve to produce surface of revolution</td>
</tr>
<tr>
<td>using a symbol for the unknown</td>
<td>ordinary arithmetic (becomes algebra)</td>
</tr>
<tr>
<td>letting $\epsilon$ be an arbitrary real number</td>
<td>epsilon-delta proof</td>
</tr>
<tr>
<td>multi-digit numbers (binary, decimal,…)</td>
<td>addition by successive adding of up to three single-digit numbers</td>
</tr>
<tr>
<td></td>
<td>infinite-series approximations</td>
</tr>
<tr>
<td>matrices for systems of equations or inequalities</td>
<td>matrix manipulation to categorize and solve</td>
</tr>
<tr>
<td>cartesian co-ordinates to identify points and</td>
<td>solving simultaneous equations to find points of intersection</td>
</tr>
<tr>
<td>equations to identify loci</td>
<td></td>
</tr>
<tr>
<td>adjacency matricies for graphs</td>
<td>contracting to produce minors in graphs</td>
</tr>
<tr>
<td></td>
<td>growing subtrees of graph</td>
</tr>
<tr>
<td>closed braids for knots</td>
<td>systematic manipulation of braids</td>
</tr>
</tbody>
</table>

Table 1: My interpretation of the distinction between instruments and tools.
especially field theory, model theory in algebra, in combinatorics, and in analysis.” Since I studied only one of these, homotopy theory (not theories), as either undergraduate or graduate student in the 1960s, I thought it might be helpful to illustrate the idea with some older things at a lower level. Marquis explained his list to me when I inquired about it.

The main point in the list that I gave was to underline how widespread the phenomenon is in contemporary mathematics: large portions of contemporary mathematics consists in building the right tools, instruments and conceptual machines to explore numbers and spaces. More generally, one sees in the 20th century the development of a systematic technology in mathematics (and I really want to make a close parallel with the role of scientific technologies in the development of the sciences). (personal communication, November 27, 2017)

My distinction is different from his since my tools manipulate objects, whereas his (and the functors of Krömer’s book [30]) take things elsewhere the better to study them.

My interest in this matter is in seeing the adoption of a particular tool or more likely combination of tools and finding out what can be done with them as analogous to a literary or other artistic genre, like string quartets. The tools chosen define the genre, and the common justification of the tools—which need not be at the level of sophistication of Marquis’ tools—and of the genre produced is the interest that is found in what one can do. As there is no reference to anything outside of mathematics in this process, it stands as mathematical justification of the procedure.25

Let us consider an example from my table, the tool analytic geometry. There has, so far as I know, never been any suggestion that geometry ought to be done as analytic rather than synthetic. When it became possible to prove theorems analytically, this was done as an exploration of what the tool could do. Both Nicolas Guisnée (d. 1718) [19] and l’Hôpital [26] wrote books doing

25The main article [2] in the April 2018 Notices of the American Mathematical Society honours the wielding not building of Hodge theory in complex algebraic geometry by Claire Voisin. It is quite explicit that what she has done so well is to solve problems with this tool.
proofs and constructions that had been done the other way by Euclid and Apollonius. In the early eighteenth century, the justification of the analytic-geometry tool consisted of what could be done with it that could already be done and that some things could be done that had not previously been done, including developing the new infinitesimal calculus.

No one has thought probably for centuries of justifying analytic geometry because it was never put forward as exclusive, presumably on account of the enormous prestige of Euclid’s synthetic approach. If one thinks of a priori justifications, then the idea of constructions with entirely imaginary compasses and straightedges is in serious need. A pluralism that has come naturally to mathematics—first to geometry—needs to be extended to the philosophy of the subject, as Reuben Hersh wrote in his [25]. He also showed a way to do it.

As with musical composition, there may be scientific or other practical reasons to work in a particular genre, and then the justification can be either mathematical because the mathematical result is pleasing (like, for example, a film score that is worth listening to elsewhere) or scientific or practical because a need was fulfilled (a film score could be strikingly effective but tied closely to the film). A mathematical example of the latter is linear programming, which has had an enormous amount of effort put into it because it is done so much, but has had little appeal to mathematicians. I shall be interested to see further reasons for limitations, having made no claim to being exhaustive.

While this paper was being considered for publication two more limitations came to my attention at the same time in an opinion piece in the Notices of the American Mathematical Society based on facts that it is an accessible source for. Jeremy Avigad [1] discusses the property of proofs called “surveyability”—mainly being short enough that a human can view the whole as one piece. The bounds of surveyability are being tested by the classification of finite simple groups, which is expected to run to twelve volumes [52]. The obvious contrast is computer-aided proofs of which the first famous one was the four-colour theorem. Hales’s proof of the Kepler sphere-packing conjecture is another. At first such computer-aided proofs were suspect, and so

\[26\text{Accessible in two senses in being both popular writing and in being freely available from ams.org.}\]
the limitation was accepted by some to have nothing to do with them. The philosophical aspect of such proofs was early discussed by Tom Tymoczko [61]. To the extent that a computer proof is a black box, it is reasonable to distrust it, and many do. There is a wild west of undependable amateur computing that some think prove something. But this trouble has its solution, it is widely held, in verification of the correctness of a computer-aided proof. This has been done for the four-colour theorem and Hales’s proof, among others. It is a small step to thinking that all proofs should be so verified. If one takes up this view, then an almost opposite limitation is set up: to nothing but computer-verified proofs. This limitation is also embraced, for instance by the late Fields medallist Vladimir Voevodsky [64]. Both are avoiding the wild west of undependable computer proofs, and the latter avoids also undependable surveyable proofs. (Voevodsky’s interest was begun by publishing an error.) If the latter demand is only that computer proofs need verification, then the two limitations do not conflict. As with non-constructive proofs, the wild west creates raw material for seeking either surveyable or verifiable proofs depending on which limitation is embraced.

Postscript

Because I regard mathematics as one of the humanities, I remember encouraging Alvin White in his work on the Humanistic Mathematics Network and its newsletter, but my computer files do not go back that far. I contributed to Newsletter Number 6 (1991) and then became busy with my own humanistic endeavour, series three of Philosophia Mathematica, which I have edited since 1992. While I have done a little history of mathematics [4, 57, 60], for most of my life I have taught [59] and studied [55, 56] mathematics with latterly occasional ventures in philosophy of mathematics [58]. Always appreciative of thoughtful journalism, this is my first venture into that genre since junior high school.

27 As managing editor of a journal, I rejected in 1979 a submission proving an inequality involving the semi-perimeter \( s = (a+b+c)/2 \) of an arbitrary triangle, angle bisectors, and a median with computer help because an algebraic proof was available, which I provided [39]. The reason for interest in the inequality was that it strengthened one found with a computer check of 500 random triangles by J. Garfunkel and subsequently shown to be correct.
References


Robert S. D. Thomas


