Counterfactual Logic and the Necessity of Mathematics



Samuel Z. Elgin¹

Received: 5 March 2020 / Accepted: 13 May 2020 / C Springer Nature B.V. 2020

Abstract

This paper is concerned with counterfactual logic and its implications for the modal status of mathematical claims. It is most directly a response to an ambitious program by Yli-Vakkuri and Hawthorne (2018), who seek to establish that mathematics is committed to its own necessity. I demonstrate that their assumptions collapse the counterfactual conditional into the material conditional. This collapse entails the success of counterfactual strengthening (the inference from 'If *A* were true, then *C* would be true' to 'If *A* and *B* were true, then *C* would be true'), which is controversial within counterfactual logic, and which has counterexamples within pure and applied mathematics. I close by discussing the dispensability of counterfactual conditionals within the language of mathematics.

Keywords Necessity of mathematics · Counterfactual conditionals · Modality

1 Introduction

Mathematical truths necessarily obtain.¹ While it is possible for Hillary Clinton to have won the 2016 presidential election, it is necessary that 2 + 2 = 4; while the Axis powers could have won World War II, it could not be that negative numbers have real square roots; and while there are some possible worlds in which everyone has brown

Samuel Z. Elgin elgin.samuel@gmail.com

¹I would like to thank the attendees of the thirteenth Annual Graduate Conference on the Philosophy of Mathematics and Logic at Cambridge University, the Conference on Mathematical and Numerical Inspirations in Philosophy at Stanford University, and the semantics working group at UC San Diego for their helpful feedback. In particular, I would like to thank Alexander Roberts for his detailed comments, Catherine Elgin for her perennial feedback, and Juhani Yli-Vakkuri for his enlightening email correspondence.

¹ University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

eyes, there are none in which all Fermat numbers are prime. History might have progressed far differently from the way that it actually did, and the laws of physics might even have diverged wildly from what they actually are, but, the received wisdom goes, pure mathematics concerns what is necessarily true—it may even be the paradigmatic example of a realm of necessary truths.

This much is uncontroversial (or, at least, as uncontroversial as anything ever is in philosophy), but there is currently no consensus on the foundations for the necessity of mathematics. In virtue of what do these truths, rather than others, hold necessarily? Are we justified in our collective confidence that they could not have been otherwise? Is there a division of labor, such that mathematics provides the truths and philosophy the necessity, or is mathematics itself committed to the necessity of its claims?

Numerous proposals are available in the literature. According to one, the necessity of mathematics is secured by the strength of our intuitions.² Perhaps conceivability is a guide to possibility; the fact that it is conceivable that p is evidence that it is possible that p, and the fact that it is inconceivable that p is evidence that it is impossible that p. If so, then our inability to conceive of a way for two and two to make five is evidence that it is impossible for two and two to make five. If all mathematical falsehoods are similarly inconceivable, we can be confident in the necessity of mathematical truths. Of course, this strategy does not determine the metaphysical basis for the necessity of mathematics, but it could explain why our belief in that necessity is justified. Alternatively, according to neologicists—who maintain that arithmetic is reducible to logic-the necessity of mathematics results from the necessity of logic.³ Arguably, the necessity of logic is as reasonable a starting point as any in modal inquiry, so if arithmetic is reducible to logic, then logical truths generate arithmetic truths that necessarily obtain. However, in light of Gödel's incompleteness theorem, neologicists typically aim only to establish the necessity of a fragment of mathematics.⁴ Still others argue that we ought not be nearly so confident in the necessity of mathematics as we currently are.⁵ Mathematicians are standardly content to prove

²See [2]. For a more general discussion of the connection between conceivability and possibility (especially in light of the [27] development of the necessary *a posteriori*), see [13].

³See, for example, [17].

⁴Yli-Vakkuri and Hawthorne claim "the neologicist strategy has inherent limitations. It can, at best, establish only the necessity of those mathematical truths that are provable in whatever axiomatic system it uses. By Gödel's first incompleteness theorem, we know that these cannot even include all truths of first-order arithmetic" (pg. 4). I find this modesty premature. It is worth recalling, as philosophers are often prone to forget, that arithmetic is only incomplete on the assumption that its axioms ought to be decidable---that is, that an infinitely large computer with an infinite amount of time ought to be able to determine whether a given formula is an axiom. There are numerous complete, albeit undecidable, axiomatizations of arithmetic. Whether decidability is an appropriate restriction depends largely on our theoretical aims. I see no reason why axioms ought to be decidable when the subject is the reduction of arithmetic to logic; all that is required is that each axiom be a principle of logic. For example, the ω -rule, according to which one may infer $\forall x F x$ after infinitely many steps determining that Fa, Fb, ... is undecidable but arguably a principle of logic (minimally, it seems as plausibly a principle of logic as Hume's Principle, according to which the number of Fs = the number of Gs just in case there is a one-to-one correspondence between the Fsand Gs, something neologicists often assume). I suspect that this humility arises because neologicists are typically committed not only to the reduction of arithmetic to logic in general, but to [12]'s derivation in particular. This strategy inevitably inherits the incompleteness of Peano arithmetic.

⁵See, for example, [18].

that something is true; they seldom bother to prove that it is necessarily true. Indeed, terms like 'necessity' and 'possibility' are conspicuously absent from the vast majority of mathematical texts. Philosophers, some claim, step in when mathematicians' work is complete and (perhaps erroneously) attribute necessity to the results of their theorems.

Yli-Vakkuri and Hawthorne [46] provide a novel defense for the necessity of mathematics. They argue that counterfactual logic and mathematical practice jointly entail that mathematics is committed to its own necessity: that, for any sentence S within the language of pure mathematics, if S is true, then S is necessarily true. Their assumptions entail that mathematics is committed not merely to the necessity of its claims but to an **S5** modal logic in particular.⁶ Its modal commitments run deep.

When I first encountered this paper, I was captivated by its result. It seemed to me that—at long last—we had no need to rely on the strength of intuition or the dubious program of neologicism. A rigorous derivation could take their place. Perfectly innocuous assumptions about counterfactual logic entail that mathematics is committed to its own necessity. Indeed, I suspected that this would eventually be seen as one of the most significant contributions to the philosophy of mathematics in many years.

My doubts have since developed. I no longer believe that this program succeeds. This paper principally consists of two worries for Yli-Vakkuri and Hawthorne's argument and its relation to the formal system they develop. In my mind, these worries are simply that: worries. They are troubling enough to undermine confidence in this program's success-they do not ensure its failure. Nevertheless, much would need to be done to restore confidence in their result. The first problem I raise is that their assumptions entail the success of counterfactual strengthening-the inference from 'If A were true, then C would be true' to 'If A and B were true, then C would be true.' Many deny the felicity of counterfactual strengthening in ordinary modal contexts. Indeed, the [41]/[28] semantics for counterfactual conditionals, which remains dominant in the discipline at large, allows for counterfactual strengthening to fail. Whether Yli-Vakkuri and Hawthorne's assumptions are tenable depends (at least partially) on whether the problematic implications of strengthening can be derived in the language of mathematics. This requires a more precise account of what constitutes pure mathematics than is currently available. The second problem is the reemergence of an objection they address: mathematical counterfactuals are entirely dispensable. Their assumptions entail that the counterfactual conditional collapses into the material conditional—so every instance of a mathematical counterfactual can be replaced with its material analog. If mathematical counterfactuals are entirely dispensable, they seem powerless to underwrite weighty modal consequences.

2 The Necessity of Mathematics

Yli-Vakkuri and Hawthorne's program fits broadly within a reorientation occurring in metaphysics. Following the formalization of modal logic in the 1960s, and the

⁶The modal commitments of Yli-Vakkuri and Hawthorne actually surpass the commitment to **S5**. As I mention below, they are committed to **TRIV**: the strongest consistent modal logic.

apparent theoretical uses for modality that ensued, many took possibility or necessity to be primitive, and defined other modal notions (such as the counterfactual conditional) in terms of them. In contrast, some contemporary philosophers maintain that the counterfactual conditional ought to be taken as primitive, and necessity and possibility defined in terms of it.⁷ The crucial definition of necessity in terms of counterfactuality is the following:

 $\Box A =_{df} \neg A \Box \rightarrow \bot$

The claim that it is necessary that *A* amounts to the claim that if *A* were false, then the absurd would obtain. This definition receives support on several fronts. It is partially motivated by the Stalnaker/Lewis semantics for counterfactual conditionals, according to which sentences of the form 'If *A* were true, then *B* would be true' hold just in case the closest possible worlds in which *A* is true are also possible worlds in which *B* is true.⁸ But perhaps the most compelling defense of this principle occurs in [45], who demonstrates that it follows from a **K** modal logic—the weakest modal logic standardly available—and the following two principles:

NECESSITY:	$\Box(A \to B) \to (A \Box \to B)$
POSSIBILITY:	$(A \Box \rightarrow B) \rightarrow (\Diamond A \rightarrow \Diamond B)$

These assert, respectively, that if it is necessary that if A then B, then if A were to obtain then B would obtain, and that if it is the case that if A were to obtain then B would obtain, then if it is possible that A then it is possible that B. With the counterfactual definition of necessity at hand, possibility can be defined in the standard way:

 $\Diamond A = {}_{df} \neg \Box \neg A^9$

With an eye toward the necessity of mathematics, Yli-Vakkuri and Hawthorne appeal to counterfactual conditionals occurring in mathematical texts. Sentences like

⁷See [45]. This trend is in its infancy; it remains to be seen whether it will stand the test of time. Part of the motivation for this approach is that, Williamson maintains, we have more direct epistemic access to counterfactual conditionals than we have to necessity and possibility. While scientific experiments may inform us of what would happen if electrons were to pass through an open slit, it is not obvious that they inform us that water is necessarily H₂O. However, for alternate accounts of our epistemic access to modality, see, for example, [16, 21, 33].

⁸This is a rough gloss on their views, which differ in philosophically important ways. In particular, Stalnaker's similarity relation selects a unique, most-similar w' for each possible world w and determines the truth of counterfactuals by what occurs in it. Lewis, in contrast, evaluates counterfactuals by truth at the closest possible worlds (plural) and does not assume that there is a unique most-similar world. Each version has benefits over the other. For example, it is a consequence of Lewis's—but not Stalnaker's—view that the Counterfactual Excluded Middle ($A \square \to B \lor A \square \to \neg B$) fails. I take it that these debates, important though they are, have no bearing on the current project.

⁹Those operating with an intuitionist modal logic would probably reject this definition of possibility—see, for example, [5]. For the purposes of this paper, I follow Yli-Vakkuri and Hawthorne in assuming classical logic.

"[If] there were a machine computing t [then] it would have some number k of states" [6] regularly appear, and are naturally interpreted as counterfactual conditionals. Given that the truth-values of these sentences depend upon merely possible situations, mathematics is plausibly committed to a wide modal scope.

There is a natural objection to this interpretation that ought to be set aside. Arguably, counterfactual conditionals with necessary or impossible antecedents are defective. A counterfactual with a necessary antecedent may collapse into the material conditional (because the closest world in which the antecedent obtains is the actual world), and counterfactuals with impossible antecedents may be vacuously true (because there are no worlds in which the antecedent obtains).¹⁰ Given the charitable assumption that mathematicians' assertions are neither trivial nor vacuous, some might reasonably prefer alternate interpretations of Boolos, Burgess and Jeffrey's sorts of claims.

There are several ways to assuage this concern. Some of Yli-Vakkuri and Hawthorne's imagined interlocutors are those who maintain that mathematical truths are contingent; they cannot object by appealing to the inadmissibility of counterfactuals with necessary or impossible antecedents, because they do not believe that mathematical counterfactuals *have* necessary or impossible antecedents. Others, however, maintain that mathematical truths are necessary (and that mathematical falsehoods are impossible), but do not identify the source of that necessity with mathematical counterfactuals. Such philosophers might deny that $\Box A$ is equivalent to $\neg A \Box \rightarrow \bot$, on the grounds that $\neg A \Box \rightarrow \bot$ may be a false counterpossible (depending upon the content of A)—that is, a false counterfactual conditional whose antecedent could not possibly obtain. However, as will soon become clear, Yli-Vakkuri and Hawthorne do not actually depend upon the equivalence of $\Box A$ and $\neg A \Box \rightarrow \bot$ in its full generality: all that their argument depends upon is the inference from $\neg A \Box \rightarrow \bot$ to $\Box A$, and this inference is relatively uncontroversial; even if there are false counterpossibles, it may be that if $\neg A \Box \rightarrow \bot$ is true, then $\Box A$ is true as well.¹¹

Yli-Vakkuri and Hawthorne assume that the language of pure mathematics is at least equipped with sentences (which are denoted by 'A,' 'B,' etc. for individual sentences and by ' Γ ,' Π ,' etc. for collections of sentences), the classical logical connectives, the counterfactual connective $\Box \rightarrow$, the absurdity operator \bot , and a symbol for informal provability \vdash . The least familiar of these is, presumably, the notion of informal provability. Informal proofs are mathematically rigorous; the main difference between informal and formal proofs is that the results of informal provability rue, while falsehoods are formally provable in systems with false axioms or unsound inferential rules. Additionally, the notion of informal provability is sensitive to mathematical practice: the fact that mathematicians regularly license a particular kind of inference is evidence that it is admissible in informal proofs.

¹⁰This is the standard Stalnaker/Lewis line. There has, however, been a sustained defense of counterpossibles: substantive counterfactual conditionals with impossible antecedents. See, for example, [3, 7, 8, 15, 19, 34]. Nevertheless, I note that Yli-Vakkuri and Hawthorne do not avoid the collapse of the counterfactual conditional into the material conditional. As I mention below, it is provable on their assumptions that $A \square B$ iff $A \rightarrow B$.

¹¹For their discussion of this point, see [46, pg. 14].

Yli-Vakkuri and Hawthorne make the following assumptions:

CLASSICAL CONSEQUENCE	$\Gamma \vdash A$ whenever A follows from Γ by classical logic.
MODUS PONENS	$\Gamma, A \Rightarrow B, A \vdash B$ where \Rightarrow is either the counterfactual or material conditional.
CUT	If $\Gamma \vdash A_1,, A_n$ and $\Pi, A_1,, A_n \vdash B$ then $\Pi, \Gamma \vdash B$.
COUNTERFACTUAL DEDUCTION	If $\Gamma, A \vdash B$, then $\Gamma \vdash A \square \rightarrow B$.
DEDUCTION THEOREM	If $\Gamma, A \vdash B$, then $\Gamma \vdash A \rightarrow B$.

Classical Consequence, Modus Ponens, Cut and Deduction Theorem are all, they claim, uncontroversial. The novel assumption is Counterfactual Deduction. But there is plenty of textual evidence that mathematicians assume that it is true. Take, for example:

Let us designate the set of all such Gödel numbers by R, and let us suppose that R is recursively enumerable. Then, since $R \neq \emptyset$, there would exist a recursive function f(n) whose range is R. [9, pg. 78]

Davis recognizes that, under the assumption that R is recursively enumerable, it is provable that there is a function whose range is R. What he concludes, then, is a counterfactual: if R were recursively enumerable, then there would be a function with R as its range. This is an instance of Counterfactual Deduction.

Or consider an elementary proof that there are infinitely many prime numbers. Suppose, for reductio, that there were only finitely many primes. In this case, these primes would have a product n. The number n + 1 would not be evenly divisible by any prime number (except the number 1, depending on whether 1 is regarded as prime) and would therefore be prime. However, n + 1 is not a factor of n, because it is larger than n. Therefore, n would not be the product of all primes, which contradicts the former claim that it is the product of all primes.

Several counterfactuals occurred in this proof. The relevant inference occurs from what is *provable* from the claim that there are finitely many primes to what *would occur* were there finitely many primes. This too is an instance of Counterfactual Deduction. Notably, the other principles Yli-Vakkuri and Hawthorne rely upon receive no sustained defense or discussion.

With such principles at hand, the derivation of the necessity of mathematics is as follows. Let *A* be an arbitrary sentence in the language of mathematics. From Classical Consequence, we have:

$$A, \neg A \vdash \bot$$

Counterfactual Deduction then entails:

$$A \vdash \neg A \Box \rightarrow \bot$$

The counterfactual definition of necessity then gives us:

 $A \vdash \Box A$

Deduction Theorem then entails:

 $\emptyset \vdash A \to \Box A$

This does not simply assert that if a sentence is true, then it is necessarily true; it makes the stronger claim that it is *provable* that if A is true, then it is necessarily true.¹²

Replacing A with $\Box A$ and $\Diamond A$ yields:

4:
$$\vdash \Box A \rightarrow \Box \Box A$$

5: $\vdash \Diamond A \rightarrow \Box \Diamond A$

Classical Consequence, Deduction Theorem, Modus Ponens, and Cut collectively imply that:

$$\vdash (\neg A \Box \rightarrow \bot) \rightarrow A$$

From the counterfactual definition of necessity, we then have:

$$\mathbf{T:} \vdash \Box A \to A$$

Additionally, the K axiom

$$\vdash \Box(A \to B) \to (\Box A \to \Box B)$$

and the necessitation rule

$$\vdash A \rightarrow \vdash \Box A$$

are both theorems. These suffice to axiomatize **S5** modal logic. And so, Yli-Vakkuri and Hawthorne conclude, mathematics is committed not only to its own necessity but to an **S5** system in particular. But although **S5** forms a lower bound on the modal logic of mathematics (in that every theorem of **S5** expressible in the language of mathematics is a theorem in mathematics), it is not an upper bound on that logic—there are theorems of modal logic that can be proven in the language of mathematics and cannot be proven purely from **S5**. Because it is provable that $\vdash A \Leftrightarrow \Box A$, the upper bound on the modal logic of mathematics is, in fact, **TRIV**. Why maintain that **TRIV** forms the upper bound? Because it is the strongest consistent modal logic, and mathematics is consistent. And so, far from being agnostic about its modal commitments, mathematics determines the system of modal logic that governs its theorems' results.

3 A Worry Concerning Counterfactual Strengthening

I hope that this (admittedly somewhat cursory) overview conveys both the structure and the initial appeal of Yli-Vakkuri and Hawthorne's argument. This argument

¹²Note that this need not conflict with the incompleteness of various mathematical systems. There may be many sentences *A* in the language of pure mathematics such that *A* is true but $\vdash A$ is false. What this asserts is that, even in these cases, $\vdash A \rightarrow \Box A$ remains true.

is incontrovertibly valid, so any disagreement must emanate from challenging their assumptions—assumptions that strike me as *prima facie* plausible.

As it turns out, these seemingly innocuous assumptions have surprising implications. In particular, they entail that the counterfactual conditional collapses into the material conditional; within the language of pure mathematics, $A \rightarrow B'$ holds just in case $A \square \rightarrow B'$ holds. The derivation of the collapse is as follows:

1. $A \to B, A \vdash B$ Modus Ponens

- 2. $A \to B \vdash A \Box \to B$ 1, Counterfactual Deduction3. $\vdash (A \to B) \to (A \Box \to B)$ 2, Deduction Theorem4. $A \Box \to B, A \vdash B$ Modus Ponens5. $A \Box \to B \vdash A \to B$ 4, Deduction Theorem6. $\vdash (A \Box \to B) \to (A \to B)$ 5, Deduction Theorem
- 7. $\vdash (A \rightarrow B) \leftrightarrow (A \square \rightarrow B)$ 3, 6 Cut and Classical Consequence

For example, it is provable that 'If 2+2 = 4, then 2+3 = 5' obtains if and only if 'If it were the case that 2+2 = 4, then 2+3 would equal 5' obtains. While this particular example is seemingly unproblematic, the collapse has undesirable implications. In particular, it forces our hand on a contentious debate between the following three principles of counterfactual logic:

SUBSTITUTION OF EQUIVALENTS	If A is logically equivalent to B, then if $A \square \rightarrow C$ then $B \square \rightarrow C$.
SIMPLIFICATION	If $(A \lor B) \Box \to C$, then $A \Box \to C$ and $B \Box \to C$.
FAILURE OF COUNTERFACTUAL STRENGTHENING	It is not the case that $A \square \rightarrow C$ entails $(A \land B) \square \rightarrow C$.

Each of these principles has received some measure of support. The Substitution of Equivalents is often defended on theoretical grounds. If two sentences are logically equivalent, it is difficult to see how any difference between them could affect the truth-values of counterfactuals they occur within. After all, they hold in precisely the same possible situations. Additionally, it is an immediate consequence of the Stalnaker/Lewis semantics for counterfactual conditionals that the Substitution of Equivalents holds. The closest possible worlds in which a sentence obtains are invariably the closest possible worlds in which equivalent sentences obtain, so accounts that rely upon the closeness of worlds do not distinguish between equivalent expressions. Even when the commitment to a particular semantics for counterfactual conditionals is dropped, many endorse a principle allowing for the substitution of equivalent expressions.¹³

Simplification is often defended by appeal to ordinary reasoning.¹⁴ It would be strange to assert 'If Jack or Jill were to come to the party, then the party would be fun, and if Jack were to come to the party, it would not be fun.' Similarly, it seems reasonable for someone to deny 'If it were to rain or not to rain, then the street would be wet' on the grounds that they deny 'If it were not to rain, then the street would be wet.' Both of these examples involve appeals to Simplification.

The Failure of Counterfactual Strengthening is often defended by appeal to the intuitive consistency of Sobel sequences.¹⁵ It may be that if Tim were to take the aspirin, he would be fine, but if Tim were to take the aspirin and the cyanide, he would not be fine, and it may be that if the Federal Reserve were to lower the interest rate, the economy would grow, but if the Federal Reserve were to lower the interest rate and the European markets were to collapse, the economy would not grow. If these sentences are consistent, as they seem to be, then counterfactual strengthening fails at least some of the time. Notably, this is a respect in which the counterfactual conditional appears to diverge from the material conditional. It is straightforward to establish that the material analog of Counterfactual Strengthening universally holds; that is, if $A \rightarrow C$, then $(A \land B) \rightarrow C$.

Despite these three principles' initial appeal, one must be abandoned, for they are mutually inconsistent. The conflict between them can be brought out in the following way:

1.	$A \dashrightarrow C$	Supposition
2.	$A \vee (A \wedge B) \square \rightarrow C$	1, Substitution of Equivalents

3. $(A \land B) \Box \rightarrow C$ 2, Simplification

If Substitution of Equivalents and Simplification are both true, it follows that counterfactual strengthening universally succeeds. The two collectively entail that if 'If Sarah were to work hard, she would get a raise' is true, then 'If Sarah were to work hard and slap her boss, she would get a raise' is true as well.

While it is indisputable that these principles are incompatible, what we ought to do in light of this incompatibility is a matter of heated debate. Arguably, the most popular option is to retain the Substitution of Equivalents and the Failure of Counterfactual Strengthening, and to abandon Simplification. This option is forced upon us by the Stalnaker/Lewis semantics for counterfactuals. As mentioned before, this semantics licenses the Substitution of Equivalents, because equivalent expressions are true in

¹³For an extended discussion of how substitution coheres with natural-language modals, see [23-26].

¹⁴This was independently noticed by [10] and [36] in response to [28]. For a response to Nute, see [31], and for the ensuing discussion about disjunctive antecedents in counterfactual conditionals more generally, see [1, 30, 37].

¹⁵See [40]. For canonical discussions of Sobel sequences, see [28, 29, 41].

the same possible situations. It also provides an intuitive explanation for the Failure of Counterfactual Strengthening. It may be that the closest worlds in which Sarah works hard are ones in which she gets a raise, but the closest worlds in which Sarah both works hard and slaps her boss are not ones in which she gets a raise, because the closest worlds in which she works hard are not ones in which she slaps her boss. Simplification fails when only one disjunct is relevant to the most-similar possible worlds. Perhaps all of the closest worlds in which either Jack or Jill come to the party are ones in which Jill comes to the party. In this case, the closest worlds in which Jack comes to the party, then the party would be fun.' Admittedly, abandoning Simplification is a theoretical cost, but the pertinent cases can arguably be accommodated pragmatically, rather than semantically.¹⁶

Others disagree. Fine [11] provides a hyperintensional semantics for counterfactual conditionals-one that preserves both Simplification and the Failure of Counterfactual Strengthening and abandons the Substitution of Equivalents. Santorio [39] advocates abandoning both the Substitution of Equivalents and Simplification, but preserving the Failure of Counterfactual Strengthening. Kocurek [22] provides independent reasons to abandon the Substitution of Equivalents. All counterpossibles have equivalent antecedents, and few license the substitution of any impossible antecedent with another. If substitution principles fail for counterpossibles, it is reasonable to expect them to fail for ordinary counterfactuals as well. And there is a growing literature by those who reject the Failure of Counterfactual Strengthening. von Fintel [42], Gillies [14] and Willer [44], for example, argue that the order in which Sobel sequences are expressed often affects how felicitous they seem.¹⁷ Some might claim 'If Luke were to go to the concert, he would see the singer' before claiming 'If Luke were to go to the concert and stand behind someone tall, he would not see the singer,' but changing the ordering of the sentences sounds odd—it would be strange to claim 'If Luke were to go to the concert and stand behind someone tall, he would not see the singer' and then assert 'If Luke were to go to the concert, he would see the singer.' But the Stalnaker/Lewis semantics for counterfactuals is insensitive to the order in which these sentences occur; if a Sobel sequence is admissible, the reverse sequence ought to be as well. On the alternate, dynamic semantics that von Fintel, Gillies and Willer provide, Counterfactual Strengthening universally succeeds, and the inadmissibility of Sobel sequences is explained by contextual shifts. And so, debate rages on. While the Stalnaker/Lewis line remains prominent (minimally, given the enduring popularity of this semantics, it is an option many are tacitly committed to), the fact that it forces our hand on this debate counts among its most controversial implications.

Yli-Vakkuri and Hawthorne's assumptions also force our hand in this debate, *but* force it differently than Stalnaker and Lewis do. Due to the collapse of the counterfactual conditional into the material conditional, their assumptions entail that the

¹⁶For pragmatic accounts of this phenomenon, see, for example, [20].

¹⁷For a response to this type of argument, see [35]. Others attempt to account for the relative felicity of Sobel sequences pragmatically—see [32, 43].

Substitution of Equivalents and Simplification are both true. Consequently, these assumptions entail that Counterfactual Strengthening universally succeeds.

The derivation of the Substitution of Equivalents is as follows:

Suppose that A is logically equivalent to B.

1.	$B \vdash A$	Classical Consequence
2.	$A \longmapsto C, A \vdash C$	Modus Ponens
3.	$A \longmapsto C, B \vdash C$	1, 2, and Cut
4.	$A \square \to C \vdash B \square \to C$	3, Counterfactual Deduction
5.	$\vdash (A \Box \rightarrow C) \rightarrow (B \Box \rightarrow C)$	4, Deduction Theorem

The derivation of Simplification is as follows:

6.	$A \vdash A \lor B$	Classical Consequence
7.	$(A \lor B) \Box \rightarrow C, A \lor B \vdash C$	Modus Ponens
8.	$(A \lor B) \Box \rightarrow C, A \vdash C$	6, 7, and Cut
9.	$(A \lor B) \square \to C \vdash A \square \to C$	8, Counterfactual Deduction
10.	$\vdash ((A \lor B) \Box \to C) \to (A \Box \to C)$	9, Deduction Theorem

As with the derivation of the necessity of mathematics, not only does it follows that the Substitution of Equivalents and Simplification hold, but it is always provable that they hold. As we have already seen, these principles, in turn, entail the success of Counterfactual Strengthening. Therefore, Yli-Vakkuri and Hawthorne's assumptions entail that every instance of Counterfactual Strengthening expressible in the language of pure mathematics succeeds.

How troubling is this result? Presumably, this depends (at least partially) on what can be expressed within the language of pure mathematics. If Sobel sequences are expressible, then the assumptions may have untenable implications. Determining the viability of this program thus requires an account of what constitutes pure mathematics. Without one, it is challenging to determine whether Sobel sequences can be expressed. However, there is reason to suspect that they face such a worry—that the language they are concerned with has the resources to express Sobel sequences.

Some purely mathematical cases seem innocuous. It seems reasonable to accept that 'If 2 were prime, there would be an even prime' and 'If 2 were prime and 3 were prime, there would be an even prime' are both perfectly true. However, other cases are much more suspect. Consider, for example, the simple arithmetic statement 'If 6 were added to 7, the result would be 13.' This sentence does not entail 'If 6 were added to 7 and 5 were subtracted, the result would be 13.' After all, the result would be 8, not 13.

And yet the second sentence is the strengthened version of the first; if counterfactual strengthening holds, the latter ought to be true if the former is. Consider, also, the relation between 'If there were a Turing machine in state T, then two steps later it would be in state T'' and 'If there were a Turing machine in state T and a 0 were changed to a 1, then two steps later it would be in state T'.' The first of these sentences may be true while the second may be false. Once again, however, the second sentence is a strengthened version of the first, so if counterfactual strengthening holds, then the truth of the first ought to guarantee the truth of the second.

These examples are even more troubling than Sobel sequences typically are. As previously mentioned, there are some who endorse Counterfactual Strengthening on the grounds that the order in which Sobel sequences are expressed impacts their felicity.¹⁸ While someone might reasonably claim 'If kangaroos lacked tails, they would fall over, and if kangaroos lacked tails and had crutches, they would not fall over,' it would be odd to assert 'If kangaroos lacked tails, they would fall over, and if kangaroos lacked tails, they would fall over,' if would be odd to assert 'If kangaroos lacked tails and had crutches, they would not fall over, and if kangaroos lacked tails, they would fall over.' If reverse Sobel sequences are inadmissible, then perhaps Counterfactual Strengthening never fails after all. However, this type of response is inapplicable to examples of Sobel sequences that occur in the language of pure mathematics; the felicity of such sentences does not turn on their ordering. 'If 6 were added to 7 and 5 were subtracted, the result would be 13' is false regardless of what sentences occur before or after it.

But perhaps Yli-Vakkuri and Hawthorne would claim that these examples miss their mark. On a somewhat narrow reading of their paper, they do not defend the necessity of mathematics as such but merely claim that the practice of mathematicians commits them to the necessity of mathematical results. If there is to be a debate about whether mathematical truths are contingent, this is a debate about whether mathematical practice ought to be revised.¹⁹ Perhaps these examples indicate that mathematicians ought to revise their practice, but arguably have no implications about what the commitments of practicing mathematicians actually are.

This reply concedes too much. If mathematical practice were committed to Counterfactual Deduction, Classical Consequence, Modus Ponens, Cut, and Deduction Theorem, then that practice would be committed to Counterfactual Strengthening—to the inference from 'If there were a Turing machine in state T and a 0 were changed to a 1, then two steps later it would be in state T'' to 'If there were a Turing machine in state T, then two steps later it would be in state T'.' But mathematicians are not committed to this inference. Therefore, mathematicians are not committed to the conjunction of Counterfactual Deduction, Classical Consequence, Modus Ponens, Cut, and Deduction Theorem.

When I first considered this problem, it seemed to me there was an additional interpretive puzzle for Yli-Vakkuri and Hawthorne: it is not at all obvious

¹⁸See, again, [14, 42, 44].

¹⁹They claim, for example, "Granted, certain proponents of the view that there are false counterpossibles may take the further radical step of rejecting COUNTERFACTUAL DEDUCTION. We are not going to engage with such radicalism, except to remind the reader that we have argued that COUNTERFACTUAL DEDUCTION is a part of the practice of mathematics, and so, if we are right, the radical step is tantamount to revisionism about that practice" (pg. 26).

what the semantics underlying mathematical counterfactuals is. The Stalnaker/Lewis semantics requires Simplification to be false, while Yli-Vakkuri and Hawthorne's assumptions entail that Simplification is true. The Stalnaker/Lewis approach remains the dominant interpretation of counterfactuals in philosophy (and beyond). Without the ability to appeal to it, another ought to take its place. Absent any semantics for mathematical counterfactuals at all, it is unclear what, precisely, they mean.

However, there is an interpretation of these counterfactuals that has the logical attributes Yli-Vakkuri and Hawthorne desire. This interpretation is referred to as a 'strict counterfactual implication,' and predates the Stalnaker/Lewis semantics.²⁰ On this view, the counterfactual conditional $A \square \rightarrow B$ is synonymous with $\square(A \rightarrow B)$. The claim that if *A* were true, then *B* would be true amounts to the claim that it is necessary that if *A* is true, then *B* is true. This interpretation can then be supplemented by the standard Kripke semantics for necessity and possibility (or any other such semantics) to furnish a semantics for counterfactual conditionals. The reason this interpretation is amenable to Yli-Vakkuri and Hawthorne's program arises firstly from the fact that the counterfactual conditional collapses into the material conditional (i.e., that $(A \square \rightarrow B)$) iff $(A \rightarrow B)$) and, secondly, from the fact that every truth within their language is necessarily true (i.e., that $(A \rightarrow B)$ iff $\square(A \rightarrow B)$).

Although this approach renders mathematical counterfactuals meaningful, it does so at a cost. The semantics underlying mathematical counterfactuals differs from the semantics of ordinary counterfactual conditionals. Those who deny the felicity of Counterfactual Strengthening in ordinary modal contexts must claim that mathematical counterfactuals operate differently from ordinary counterfactuals. After all, many reject the strict interpretation precisely because it licenses Strengthening: it entails 'If Julia were to take the bus, she would save money' implies that 'If Julia were to take the bus and buy a Ferrari, she would save money.' After all, in order for the first sentence to be true, it must be necessary that if Julia takes the bus, then she saves money. This approach thus requires the semantics for mathematical counterfactuals to come apart from the semantics of ordinary counterfactuals-the two types of expressions mean different things. What's more, this is not peculiar to the strict interpretation of counterfactual conditionals. Because Yli-Vakkuri and Hawthorne's assumptions entail that the logic of mathematical counterfactuals differs from the logic of ordinary counterfactuals, any semantics for mathematical counterfactuals must likewise diverge from a semantics for ordinary counterfactuals.²¹ So while it may be possible for Yli-Vakkuri and Hawthorne to supplement their view with a semantics, any they appeal to will require that counterfactuals occurring in mathematical contexts mean something different from counterfactuals occurring in ordinary contexts. This raises difficulties for the interpretation of sentences that conjoin a purely mathematical counterfactual with an ordinary counterfactual, as there is a shift in meaning between the two.

²⁰This type of view was endorsed by, for example, [38]. How extensive its history is is a matter of debate. Peirce attributes this sort of view to the Hellenistic logician Philo. However, [4] prefers an alternate interpretation of Philo's work.

²¹More precisely, any semantics for which their assumptions are both sound and complete will not be a semantics for which the logic of ordinary counterfactual conditionals is sound and complete—at least on the assumption that Counterfactual Strengthening fails for ordinary counterfactuals.

4 A Worry Concerning the Dispensability of Mathematical Counterfactuals

Perhaps surprisingly, in Yli-Vakkuri and Hawthorne's view, counterfactual conditionals are dispensable to mathematics. Recall that their assumptions entail that the counterfactual conditional collapses into the material conditional; for arbitrary sentences A and B within the language of pure mathematics, $A \rightarrow B$ holds just in case $A \square \rightarrow B$ holds as well. Therefore, every mathematical counterfactual could be replaced with its material analog *salva veritate*.²²

This is borne out by examples that Yli-Vakkuri and Hawthorne discuss. Every mathematical counterfactual that they mention could be replaced by a material conditional without impacting the mathematical results. Some examples (chosen effectively at random) are the following:

"If there were another system of the conjugate Sylow *p*-groups, then its members would be transformed into each other by **B** in systems of transitivity whose degree would be divisible by p" (pg. 6)

could be rephrased as

If there is another system of the conjugate Sylow p-groups, then its members are transformed into each other by **B** in systems of transitivity whose degree would be divisible by p.

While

"Let us designate the set of all such Gödel numbers by R, and let us suppose that F is recursively enumerable. Then, since $R \neq \emptyset$, there would exist a recursive function f(n) whose range is R" (pg. 9)

contains the same information as

Let us designate the set of all such Gödel numbers by R, and let us suppose that F is recursively enumerable. Then, since $R \neq \emptyset$, there is a recursive function f(n) whose range is R.

And

"If it were the case that, for some plane separated into contiguous regions, there is no function f from the regions to a four-membered pure set such that, for all pairs x, y of contiguous regions $f(x) \neq f(y)$, then it would be the case that \perp " (pg. 17)

amounts to the claim that

If it is the case that, for some plane separated into contiguous regions, there is no function f from the regions to a four-membered pure set such that, for all pairs x, y of contiguous regions $f(x) \neq f(y)$, then it is the case that \perp .

²²And, indeed, every material conditional could be replaced with its counterfactual analog instead.

Why is the dispensability of mathematical counterfactuals significant? At least partially because others have argued that at least some mathematical counterfactuals are *in*dispensable—that mathematics employs counterfactual conditionals that cannot be replaced with their material counterparts. One example occurs in "Counterpossibles in Science: The Case of Relative Computability" [19]. Jenny argues that counterpossibles are integral to the theory of relative computability—a branch of mathematics that addresses whether solutions to some problems would generate solutions to others.²³

A set is said to be algorithmically decidable just in case there is some algorithm that, after finitely many steps, determines whether an arbitrary entity is a member of that set. For example, the set of propositional tautologies is algorithmically decidable (in that there is an algorithm that determines in finitely many steps whether an arbitrary sentence of propositional logic is a tautology), while the set of predicatively valid sentences is not (in that there is no algorithm that determined in finitely many steps whether an arbitrary sentence of predicate logic is valid). This is referred to interchangeably as the 'validity problem' and the 'decision problem.' Additionally, the set of arithmetic truths is undecidable (in that there is no algorithm that determines whether an arbitrary sentence is an arithmetic truth) as is the set of algorithms that eventually halt—(i.e., there is no algorithm that determines whether an arbitrary algorithm will halt after finitely many steps).

Although neither the validity problem nor the halting problem is decidable, mathematicians investigate what would happen if they were. In particular, they investigate what would follow if there were an algorithm that solved the validity problem. As it turns out, if the validity problem were decidable, then the halting problem would be as well. Any solution to the validity problem could be used to generate a solution to the halting problem. The sentence 'If there were a solution to the validity problem, then there would be a solution to the halting problem' appears to be a counterfactual and, Jenny argues, is indispensable to relative computability theory. However, counterfactuals in relative computability theory cannot be replaced by their material analogs, precisely because they do not allow for counterfactual strengthening. Consider, for example, the sentence:

Even if the validity problem were decidable, then the arithmetic problem would not be decidable.

Mathematicians maintain that this sentence is true on the grounds that, while a solution to the validity problem could be used to generate a solution to the halting problem, a solution to the validity problem could not be used to generate a solution to the arithmetic problem. However, mathematicians deny a strengthened version of this counterfactual:

Even if the validity problem were decidable and the arithmetic problem were decidable, then the arithmetic problem would not be decidable.

²³For an introduction to relative computability theory, see [9].

After all, if there were a solution to the arithmetic problem, then there would trivially be a solution to the arithmetic problem—arithmetic truth is computable relative to itself. The counterfactual conditional diverges from the material conditional in this case—the inference from 'If the validity problem is decidable, the arithmetic problem is not decidable' to 'If the validity problem is decidable and the arithmetic problem is decidable, then the arithmetic problem is not decidable' is legitimate: both are true because they are material conditionals with false antecedents. If Jenny is correct (in claiming that at least some mathematical counterfactuals are distinct from their material analogs), then Yli-Vakkuri and Hawthorne err in collapsing the counterfactual conditional to the material conditional.

The dispensability of mathematical counterfactuals is troubling for further reasons. If every instance of a mathematical counterfactual conditional could be replaced by a material conditional, it is difficult to see what or how these counterfactuals add to our body of knowledge. Nothing modally substantive follows from material conditionals used by mathematicians, and Yli-Vakkuri and Hawthorne's assumptions ensure that mathematicians could exclusively employ the material conditional. So, why think that these dispensable conditionals ought to be used to derive modally substantive results?

Although they do not address the collapse of one conditional to the other, Yli-Vakkuri and Hawthorne discuss the dispensability objection at length. They claim:

Counterfactuals are absolutely indispensable to what mathematics contributes to our total body of knowledge...Note first that myriad applications of mathematics to the hustle and bustle of both everyday life and engineering require our knowing that mathematical truths would remain true even if things had gone differently in various ways. For example, in justifying a particular engineering solution, one often appeals to mathematical truths in reasoning about how things would have gone if one had opted for an alternative solution. In doing so one assumes—and if one is successful, one knows—that those mathematical truths would have been true even if one had opted for the alternative solution. Note second that, as the queen of the sciences, mathematics is primed for application in any area of objective inquiry, whether it be the science of electromagnetism, the theory of rook and pawn endings, or natural language semantics. (pg. 14)

This passage strongly suggests that mathematical counterfactuals occur in disciplines ranging from engineering to electromagnetism to natural-language semantics. After all, if the language of mathematics is incapable of expressing these counterfactuals, how could they lend support for the indispensability of counterfactuals in mathematics? Sobel sequences are derivable in every discipline they mention. An engineer may claim that, if a pulley were to double in size, it could lift a heavy box, but may deny that if a pulley were to double in size and be made of twine, it could lift a heavy box. A physicist might conclude that if an electron were to be placed in a field, it would accelerate, but deny that if an electron were to be placed in a field and an equal-but-opposite force were to be introduced, it would accelerate. Both the engineer and physicist thus deny the felicity of Counterfactual Strengthening in the counterfactuals they appeal to. Yli-Vakkuri and Hawthorne's assumptions, which entail that Counterfactual Strengthening succeeds, are at odds with this practice. Of course, it might be claimed that sentences occurring in engineering, physics, and the like are not pure mathematics. Yli-Vakkuri and Hawthorne's argument for the indispensability of mathematical counterfactuals is arguably an appeal to *applied* mathematics—not to pure mathematics. As such, these sentences fall outside of the scope of their program; they need not claim that Counterfactual Strengthening is admissible in these types of cases, because these sentences are not within the language of pure mathematical, the dispensability objection returns. After all, it may be that counterfactuals are indispensable to applied mathematics, but are they indispensable to pure mathematics? And so, we arrive at a crossroads. If counterfactual conditionals that appear in applied mathematics fall within the scope of Yli-Vakkuri and Hawthorne's project, then their assumptions err in licensing counterfactual strengthening. And if these counterfactuals do not fall within that scope, what reason is there to think that mathematical counterfactuals are indispensable?

5 Conclusion

I have presented two related concerns for Yli-Vakkuri and Hawthorne's program: that their assumptions illicitly give rise to counterfactual strengthening, and that counterfactuals are entirely dispensable to mathematics. Both arise from the collapse of the counterfactual to the material conditional, which is provable on the assumptions that they make.

And for both, the same response is available. Because there is no account of what pure mathematics consists of, Yli-Vakkuri and Hawthorne may simply say that the intended examples fall outside of their intended scope. But with each retreat, the project becomes less ambitious. Problematic examples arise in arithmetic, computability theory, and numerous applied fields. In order for counterfactual logic to secure the foundations for the necessity of mathematics, we require an account of what pure mathematics consists of—an account that evades the problems generated by counterfactual strengthening and that accounts for how it is dispensable counterfactuals furnish the resources for the necessity of mathematics.

References

- 1. Alonso-Ovalle, L. (2006). Counterfactuals, correlatives and disjunction. *Linguistics and Philosophy*, 32, 207–44.
- Bealer, G. (2002). Modal epistemology and the rationalist renaissance. In Gendler, T., & Hawthorne, J. (Eds.) *Conceivability and possibility* (pp. 71–126): Oxford University Press.
- 3. Bjerring, J.C. (2013). On Counterpossibles. Philosophical Studies, 190, 327-53.
- 4. Bobzien, S. (2011). Dialectical school. Stanford Encyclopedia of Philosophy.
- Bobzien, S., & Ian, R. (2020). Intuitionism and the modal logic of vagueness. *Journal of Philosophical Logic*, 49, 221–48.
- 6. Boolos, G., Burgess, J., Jeffrey, v. (2007). Computability and logic. Cambridge University Press.
- 7. Brogaard, B., & Salerno, J. (2013). Remarks on counterpossibles. Synthese, 190, 639-60.

- 8. Cohen, D. (1987). The problem of counterpossibles. *Notre Dame Journal of Formal Logic*, 29(9), 91–101.
- 9. Davis, M. (1958). Computability and unsolvability. McGraw-Hill.
- 10. Fine, K. (1975). Review of Lewis's counterfactuals. *Mind*, 84, 451-8.
- 11. Fine, K. (2012). Counterfactuals without possible worlds. Journal of Philosophy, 109(3), 221-46.
- 12. Frege, G. (1884). The foundations of arithmetic. Oxford University Press.
- 13. Gendler, T., & Hawthorne, J. (2002). Conceivability and possibility. Oxford University Press.
- 14. Gillies, A. (2007). Counterfactual scorekeeping. Linguistics and Philosophy, 30, 329-60.
- Goodman, J. (2004). An extended Lewis/Stalnaker semantics and the new problem of counterpossibles. *Philosophical Papers*, 33, 35–66.
- Hale, R. (2003). Knowledge of possibility and necessity. Proceedings of the Aristotelian Society, 103, 1–20.
- 17. Hale, R., & Wright, C. (2001). The Reason's proper study: essays toward a Neo-Fregean philosophy of mathematics. Oxford University Press.
- 18. Hodges, W. (Forthcoming). Necessity in mathematics.
- Jenny, M. (2018). Counterpossibles in science, the case of relative computability. Noûs, 52(3), 530– 60.
- Klinedinst, N. (2009). (Simplification of) disjunctive antecedents. MIT Working Papers in Linguistics 60.
- 21. Kment, B. (2018). Essence and modal knowledge. Synthese, 1-23.
- 22. Kocurek, A. (Forthcoming). On the substitution of identicals in counterfactual reasoning. Noûs.
- 23. Kratzer, A. (1981). The notional category of modality. In Eikmeyer, H.J., & de Gruyter, H. R. (Eds.) *Words, worlds and contexts: new approaches to world semantics.*
- Kratzer, A. (1981). Partition and reviesion: the semantics of counterfactuals. *Journal of Philosophical Logic*, 10(2), 201–16.
- Kratzer, A. (1986). Conditionals. Chicago Linguistics Society: Papers from the Parasession on Pragmatics and Grammatical Theory, 22(2), 1–15.
- Kratzer, A. (1991). Modality. Semantics: an international handbook of contemporary research (pp. 639–50).
- 27. Kripke, SA. (1980). Naming and necessity. Harvard University Press.
- 28. Lewis, D. (1973). Counterfactuals. Blackwell Publishers.
- Lewis, D. (1973). Counterfactuals and comparative probability. *Journal of Philosophical Logic*, 2(4), 418–46.
- Lewis, D. (1977). Possible-world semantics for counterfactual logics: a rejoinder. *Journal of Philosophical Logic*, 6(1), 359–63.
- Loewer, B. (1976). Counterfactuals with disjunctive antecedents. *Journal of Philosophy*, 73(16), 531– 7.
- 32. Lowe, E.J. (1983). A simplification of the logic of conditionals. *Notre Dame Journal of Formal Logic*, 24(3), 357–66.
- 33. Lowe, E.J. (2012). What is the source of our knowledge of modal truths? Mind, 121(484), 919-50.
- 34. Mares, E. (1997). Who's afraid of impossible worlds? *Notre Dame Journal of Formal Logic*, 38, 535–72.
- 35. Moss, S. (2012). On the pragmatics of counterfactuals. Noûs, 46(3), 561-86.
- 36. Nute, D. (1975). Counterfactuals and the similarity of worlds. Journal of Philosophy, 72(21), 773-8.
- Nute, D. (1980). Conversational scorekeeping and conditionals. *Journal of Philosophical Logic*, 9(2), 153–66.
- 38. Peirce, C. (1896). Regenerated logic. The Monist, 7(1), 19-40.
- 39. Santorio, P. (2018). Alternatives and truthmakers in conditional semantics. *Journal of Philosophy*, 115(10), 513–49.
- 40. Sobel, H. (1970). Utilitarianisms: simple and general. Inquiry, 13(4), 2008.
- 41. Stalnaker, R. (1968). A theory of conditionals. In Recher, N. (Ed.) *Studies in logical theory*: Oxford University Press.
- von Fintel, K. (2001). Counterfactuals in a dynamic context. In Kenstowicz, M. (Ed.) Ken Hale: a life in language (pp. 123–52): MIT Press.

- 43. Warmbröd, K. (1981). An indexical theory of conditionals. Dialogue, 20(4), 644-64.
- 44. Willer, M. (2017). Lessons from Sobel sequences. Semantics and Pragmatics, 10(4), 1-57.
- 45. Williamson, T. (2007). The philosophy of philosophy. Blackwell Publishing.
- 46. Yli-Vakkuri, J., & Hawthorne, J. (2018). The necessity of mathematics. Noûs, 1-28.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.