Potential Infinity
a modal account

Øystein Linnebo and Stewart Shapiro

March 1, 2016
Beginning with Aristotle, almost every major philosopher and mathematician before the nineteenth century rejected the notion of the actual infinite. They all argued that the only sensible notion is that of potential infinity—at least for scientific or, later, non-theological purposes.

In *Physics* 3.6 (206a27-29), Aristotle wrote, “For generally the infinite is as follows: there is always another and another to be taken. And the thing taken will always be finite, but always different (2o6a27-29).”

As Richard Sorabji [38] (322-3) puts it, for Aristotle, “infinity is an extended finitude”. (see also [19],[20])
Aristotle, along with ancient, medieval, and early modern mathematicians, recognized the existence of certain *procedures* that can be iterated indefinitely, without limit. Examples are the bisection and the extension of line segments.

Ancient mathematicians made brilliant use of such procedures. For example, the method of exhaustion, a kind of forerunner to integration, was employed to calculate the areas of curved figures in terms of rectilinear ones.

What was rejected are what would be the end results of applying these procedures infinitely often: self-standing points and infinitely long regions.
In *On generation and corruption*, Aristotle writes:

> For, since no point is contiguous to another point, magnitudes are divisible through and through in one sense, and yet not in another. When . . . it is admitted that a magnitude is divisible through and through, it is thought that there is a point not only anywhere, but also everywhere in it: hence it follows that the magnitude must be divided away into nothing. For there is a point everywhere within it, so that it consists either of contacts or of points. But it is only in one sense that the magnitude is divisible through and through, viz. in so far as there is one point anywhere within in and all its points are everywhere within it if you take them singly. (317a3-8)
Closely related to the notion of potential infinity, for Aristotle, is that of potential existence. For Aristotle, points just are the limits of line segments (and line segments just are the edges of plane figures, and plane figures just are the (flat) surfaces of physical objects).

The points interior to a line segment only exist potentially. They are places where the line can be broken. But if the line is not broken there, the point only exists potentially.

The same goes for the parts of the line segment themselves. As a continuous magnitude, the line segment is a unity. Its parts exist only potentially.
Jonathan Lear [19] argues that it is not the existence of iterated procedures that makes for Aristotelian potential infinity. The matter concerns the *structure* of geometric magnitudes:
… it is easy to be misled into thinking that, for Aristotle, a length is said to be potentially infinite because there could be a process of division that continued without end. Then it is natural to be confused as to why such a process would not also show the line to be actually infinite by division. . . . [I]t would be more accurate to say that, for Aristotle, it is because the length is potentially infinite that there could be such a process. More accurate, but still not true . . . Strictly speaking there could not be such a process, but the reason why there could not be is independent of the structure of the magnitude: however earnest a divider I may be, I am also mortal. . . . even at that sad moment when the process of division does terminate, there will remain divisions which could have been made. The length is potentially infinite not because of the existence of any process, but because of the structure of the magnitude. (p. 193)
According to Lear, then, a line segment is potentially infinite because there are infinitely many places where it *can be* divided. So, no matter how many times one divides a line, there will still be some of the line left.

Lear concludes that Aristotle’s thesis is “that the structure of the magnitude is such that any division will have to be only a partial realization of its infinite divisibility: there will have to be possible divisions that remain unactualized” (p. 194).
On either Lear’s reading of Aristotle or the above gloss concerning iterated procedures, potential infinity invokes both modality and the activities of a perhaps idealized mathematician.
There is a closely related matter. It is generally agreed that Euclid’s *Elements* captures at least the spirit of geometry during Plato’s and Aristotle’s period. Most of the language in the *Elements* is dynamic, talking about what a (presumably idealized) geometer *can do*.

For example, the First Postulate is “To draw a straight line from any point to any point”, and the Second is “To produce a finite straight line continuously in a straight line”.

Or consider the infamous Fifth:

*That, if a straight line falling on two straight lines make the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than the two right angles.*
Plato was critical of the geometers of his day, arguing that this dynamic language is inconsistent with the nature of the true subject matter of geometry:

[The] science [of geometry] is in direct contradiction with the language employed by its adepts . . . Their language is most ludicrous, . . . for they speak as if they were doing something and as if all their words were directed toward action . . . [They talk] of squaring and applying and adding and the like . . . whereas in fact the real object of the entire subject is . . . knowledge . . . of what eternally exists, not of anything that comes to be this or that at some time and ceases to be. (Republic, VII)
This, of course, is a nice Platonic thought—for those who like such thoughts.

Aristotle rejected this orientation and, we suggest, the dynamic language employed in Ancient geometry better reflects his views.

The matter of infinity is tied to this. For Aristotle, we never have infinite collections of points, objects, or anything else, and we never have infinitely long lines or infinitely large regions of space or time. Because of the structure of the geometric realm, we have procedures that can be iterated indefinitely, and we speak about what those procedures could produce, or what they will eventually produce if carried sufficiently (but only finitely) far.
In holding that these geometric procedures can be iterated indefinitely, Aristotle again follows the mathematical practice of the time, this time in opposition to his other major opponents, the atomists (see [26]), who postulate a limit to, say, bisection.
Views like Aristotle’s were standard throughout the medieval and early modern period, through most of the nineteenth century. The greatest mathematical minds insisted that only the potentially infinite makes sense. Leibniz, for example, wrote:

*It could . . . well be argued that, since among any ten terms there is a last number, which is also the greatest of those numbers, it follows that among all numbers there is a last number, which is also the greatest of all numbers. But I think that such a number implies a contradiction . . . When it is said that there are infinitely many terms, it is not being said that there is some specific number of them, but that there are more than any specific number.* (Letter to Bernoulli, [21], III 566, translated in [24], 76-77, 87)
The contradiction here might be the so-called “Galileo paradox”, that, with an infinite collection, the part can be the same size as the whole. Today, of course, this is regarded as a feature of the infinite, not a bug.

Leibniz:

... we conclude ... that there is no infinite multitude, from which it will follow that there is not an infinity of things, either. Or [rather] it must be said that an infinity of things is not one whole, or that there is no aggregate of them. ([22], 6.3, 503, translated in [24], 86)
Yet M. Descartes and his followers, in making the world out to be indefinite so that we cannot conceive of any end to it, have said that matter has no limits. They have some reason for replacing the term “infinite” by “indefinite”, for there is never an infinite whole in the world, though there are always wholes greater than others ad infinitum. As I have shown elsewhere, the universe cannot be considered to be a whole. ([23], 151)

For Leibniz, as for Aristotle, as for a host of others, the infinite just is the limitlessness of certain processes; no actual infinities exist. The only intelligible notion of infinity is that of potential infinity—the transcendence of any (finite) limit.
For at least the cases of interest here—regions, natural numbers, and the like—Georg Cantor argued for the exact opposite of this, claiming that the potentially infinite is dubious, unless it is somehow backed by an actual infinity:

\[
\text{I cannot ascribe any being to the indefinite, the variable, the improper infinite in whatever form they appear, because they are nothing but either relational concepts or merely subjective representations or intuitions (imaginationes), but never adequate ideas ([6], 205, note 3).}
\]
... every potential infinite, if it is to be applicable in a rigorous mathematical way, presupposes an actual infinite ([7], 410–411).

We think it safe to say that this Cantorian orientation is now dominant in the relevant intellectual communities, especially concerning the mathematical domains mentioned above, with various constructivists as notable exceptions.
It should be noted that, on the surface, at least, Cantor was not consistent in his rejection of the potential infinite. Sometimes he ascribed to so-called “absolutely infinite”, or what he dubbed “inconsistent multitudes” (e.g., the ordinals), features closely analogous to those of the potentially infinite.
In a much quoted letter to Dedekind, in 1899, he wrote:

\[\text{It is necessary \ldots to distinguish two kinds of multiplicities \ldots For a multiplicity can be such that the assumption that all of its elements \textit{are together} leads to a contradiction, so that it is impossible to conceive of the multiplicity as a unity, as \textit{one finished thing}. Such multiplicities I call absolutely infinite or inconsistent multiplicities \ldots If on the other hand the totality of the elements of a multiplicity can be thought of without contradiction as \textit{being together}, so that they can be gathered together into \textit{one thing}, I call it a consistent multiplicity or a \textit{set}. (Ewald [10], 931-932)\]
To the present: Cantor

An 1897 letter to Hilbert is even more suggestive:

*I say of a set that it can be thought of as finished . . . if it is possible without contradiction (as can be done with finite sets) to think of all its elements as existing together . . . or (in other words) if it is possible to imagine the set as actually existing with the totality of its elements.* (Ewald [10], 927)

Of course, Cantor considered all of the transfinite sets, such as the natural numbers and the real numbers, to be actual infinities. Most of our present concern is with those, but we will address the potentiality of the iterative hierarchy.

At least with hindsight, it is clear what Aristotle, and the ancient, medieval, and early modern mathematicians did not have: infinitely large, completed collections and infinitely large geometric figures. But what did they have instead? Just what is a potential infinity?
Concerning the contemporary scene, Karl-George Niebergall [29] writes:

To gain a better understanding of [the phrases “x is potentially infinite” and “T makes an assumption of the potentially infinite”] I regard as a goal in itself. One simply has to admit that they play an important role both in philosophy and in investigations on the foundations of mathematics. My feeling is, however, that a clear meaning has never been given to them. (p. 231)
Further skepticism

He later adds

*Personally, I simply have no ordinary understanding of these phrases, and I do not find much help in the existing literature on them. It seems that even examples are missing ...* [T]hose philosophers who are interested in the theme of the potentially infinite are usually drawn to it because they regard it as desirable to avoid assumptions of infinity (i.e., of the actual infinity), yet do not want to be restricted to a mere finitist position. An assumption of merely the potentially infinite seems to be a way out of this quandary: ... *it allows you to have your cake and eat it too (p. 256-7).*
Niebergall [28], §§2.5,3; [29], §6 argues that, on some straightforward attempts at definition, the potentially infinite just collapses into the actually infinite (or the finite). So, what we cannot have, he claims, is a notion of the potentially infinite that is to be distinguished from both the finite and the actually infinite. If we follow contemporary practice and allow the actually infinite a place, then there is no room for the merely potentially infinite. Everything is either finite or infinite—nothing can fit between those:
Further skepticism

given the modern conception of sets in classical logic, I have never understood “potentially infinite” in a way which would give the potential infinite any use—a place of its own different from the finite and the infinite. It seems to me that merely in intuitionistic frameworks the predicate “potentially infinite” may find a reasonable rational reconstruction. Thus, in dealing with this discourse I feel as if I were a (Quinean) linguist in the process of developing a radical translation from a foreign language into his mother tongue. ([28], 167, n. 41)
In [29], 257, however, he concedes: “It is granted that one could try to define ‘x is potentially infinite’ by employing a modal vocabulary.” That is the plan here. We hope to articulate a serviceable notion of the potentially infinite, one that can live, side by side, with the actually infinite. This plays a role in understanding the notions of mathematical “construction”, in indefinite extensibility, and in the debate over absolute generality. It will also help to articulate the thesis that the iterative hierarchy is itself potential.
Niebergall goes on to claim that “at least since Quine’s criticism there is reasonable doubt as to the general understandability of the modal idiom”. He adds a note that when it comes to mathematics, “talk of possibility and necessity becomes dispensable”, since “a mathematical sentence is regarded as necessary if true” (p. 258). It is, of course, incumbent on us to say something about the modality involved, and to distinguish it from the sense in which every mathematical sentence is “necessary if true”.
Our goal

Our over-arching aim in this paper is to analyze the notion of potential infinity and assess its scientific merits. This aim leads to a number of more specific questions.

Perhaps the most pressing one is whether the conception of potential infinity can be explicated in a way that is both interesting and substantially different from the now-dominant conception of actual infinity. One might suspect that, when metaphors and loose talk give way to precise definitions, the apparent differences will evaporate.

As we will explain, however, a number of differences still remain. Some of the most interesting and surprising differences concern consequences that one’s conception of infinity has for higher-order logic.
Another important question concerns the relation between potential infinity and mathematical intuitionism. More than anyone else, intuitionists have continued to take potential infinity seriously even after the Cantorian revolution. In part as a result of this, it is now commonly thought that the appropriate logic for potential infinity is intuitionistic.

We will analyze the question and find that potential infinity is not inextricably tied to intuitionistic logic. There are interesting explications of potential infinity that underwrite classical logic, while still differing in important ways from actual infinity.

However, we will also find that on some more stringent explications, potential infinity does indeed lead to intuitionistic logic. We take this clarification of the relation between potential infinity and intuitionistic logic to be one of the main achievements of the paper.
Recall Aristotle’s analysis of the potential infinite:

For generally the infinite is as follows: there is always another
and another to be taken. And the thing taken will always be
finite, but always different (Physics, 206a27-29).

A nice example is Aristotle’s claim that matter is infinitely divisible. Consider a stick. However many times one has divided the stick, it is always possible to divide it again. Now, it is fairly natural to explicate Aristotle’s temporal vocabulary in a modal way. This yields the following analysis of the infinite divisibility of a stick $s$:

$$\Box \forall x (Pxs \rightarrow \Diamond \exists y Pyx)$$

where $Pxy$ means that $x$ is a proper part of $y$. 

$$\text{(1)}$$

Modal explication
Modal explication

If, by contrast, the divisions of the stick formed an actual infinity, the following would hold:

$$\forall x (Pxs \rightarrow \exists y Pyx)$$  \hspace{1cm} (2)

According to Aristotle, it is not even possible to complete infinitely many divisions of the stick, that is:

$$\neg \Box \forall x (Pxs \rightarrow \exists y Pyx)$$  \hspace{1cm} (3)

By endorsing both (1) and (3), Aristotle is asserting that the divisions of the stick are merely potentially infinite, or incompletable, as we will also put it.
Consider now the natural numbers, which according to Aristotle are also merely potentially infinite. We can represent this view as the conjunction of the following theses:

\[
\Box \forall m \exists n \text{SUCCESSOR}(m, n) \quad (4)
\]

\[
\neg \Diamond \forall m \exists n \text{SUCCESSOR}(m, n) \quad (5)
\]

Thus, provided we are willing to use the resources of modal logic, there is no problem whatsoever in distinguishing the merely potential infinite from the actual infinite.
This use of modal logic introduces a problem of translation, however. Other than in informal glosses, the language of mathematics is today relentlessly non-modal. In particular, when the question of the appropriate logic of potential infinity arises, it tends to do so in the ordinary, non-modal language of arithmetic. So we will need some bridge between the modal language in which we analyze potential infinity and the ordinary non-modal language.
Modal explication

The most famous such bridge is provided by the Gödel translation $G$ of the language of intuitionistic logic. The translation is given by the following clauses:

$$
\begin{align*}
\phi & \mapsto \Box \phi, \quad \text{for } \phi \text{ atomic} \\
\neg \phi & \mapsto \Box \neg \phi^G \\
\phi \lor \psi & \mapsto \phi^G \lor \psi^G \\
\phi \land \psi & \mapsto \phi^G \land \psi^G \\
\phi \rightarrow \psi & \mapsto \Box (\phi^G \rightarrow \psi^G) \\
\forall x \phi & \mapsto \Box \forall x \phi^G \\
\exists x \phi & \mapsto \exists x \phi^G
\end{align*}
$$
Modal explication

It is well known that the Gödel translation validates intuitionistic logic when the logic of the modal language is S4:

**Theorem 1:** Let $\vdash_{\text{int}}$ be intuitionistic deducibility in the given language, but if the language is plural or higher-order, we remove any comprehension axioms (taking those to be non-logical). Let $\vdash_{S4}$ be the corresponding deducibility relation in S4 and classical logic. Then we have:

$$\phi_1, \ldots, \phi_n \vdash_{\text{int}} \psi \text{ if and only if } \phi_1^G, \ldots, \phi_n^G \vdash_{S4} \psi^G.$$
One might therefore think that the Gödel translation provides the desired bridge, and that we have accordingly established that the logic of potential infinity is intuitionistic.

This thought turns out to be badly flawed, however. The Gödel translation is hopeless in connection with our modal explication of potential infinity. Consider the principle that every number has a successor:

\[ \forall m \exists n \text{SUCCESSOR}(m, n) \]

This is an axiom of Peano and Heyting arithmetic. But its Gödel translation is

\[ \Box \forall m \exists n \text{SUCCESSOR}(m, n) \]

which requires that every world that contains one number, contains all of them. But this is precisely what a potentialist denies.
Modal explication

Something similar happens with respect to the Kripke modal semantics for intuitionistic logic (which also invokes the modal logic S4, insisting that domains do not decrease along the accessibility relation). Consider, again, the principle that every number has a successor:

$$\forall m \exists n \text{Successor}(m, n),$$

an axiom of Heyting arithmetic. For that to be true at the base node in a Kripke model, every number in that node has to have a successor in that node. Given that zero exists at the base node of a model of Heyting arithmetic, every natural number exists at the base node. If the structure does not sanction the law of excluded middle, it must have other nodes with non-standard (i.e., infinitely large) numbers.
In sum, the Gödel translation and the Kripke semantics both capture intuitionistic logic, but neither makes much sense of the notion of potential infinity.

The heart of potentialism is rather that the existential quantifier of ordinary non-modal arithmetic has an implicit potential character. When a potentialist says that a number has a successor, she really means that it potentially has a successor—that is, that it is possible to construct or define a successor.

To remain as neutral as we can on how the successor is introduced, we will talk schematically about generating a successor.
This suggests that the right translation of $\exists$ is $\Diamond \exists$, rather than the homophonic translation used above, or the clause in the Kripke semantics.

And since the universal quantifier can hardly be less inclusive in its range than the existential, this also suggests that $\forall$ be translated as $\Box \forall$, that is, as a statement that whatever objects we go on to generate, everything generated will be thus-and-so.

Let us call this the *potentialist translation*. And let us call the modal operator-quantifier hybrids $\Box \forall$ and $\Diamond \exists$ *modalized quantifiers*. 
Modal explication

We can now raise the question of which principles in the non-modal language are validated via this bridge. The answer will obviously depend on the modal logic that we plug in on the modal side. In the next section, we will defend a certain choice, namely a system known as S4.2. In this system, we can prove the highly desirable result that the modalized quantifiers behave logically like the ordinary quantifiers, in a sense to be made formally precise.
This yields a conditional answer to the question of the logic of potential infinity. If the background modal logic is *classical* S4.2, then the potentialist translation validates classical logic, while if the background modal logic is *intuitionistic* S4.2, the same goes for the logic that is validated.

This shows that potential infinity must (or at least can) be separated from intuitionistic logic.
Some theorists may object that potential infinity no longer looks very different from actual infinity, in light of the classical version of the “mirroring theorem” we just alluded to.

There is one fairly conservative response available. One can insist that, despite the shared adherence to classical logic, the outlined explication of potential infinity is nevertheless importantly different from actual infinity. This difference is manifested not only in the (implicit) presence of modal operators, which a skeptic might challenge, but also in the higher-order logic that is validated. Some details will come later.
Philosophical interlude

There is also a more radical response. One might insist that if we are serious about the merely potential infinity of, say, the natural numbers, it is not enough to insist that every number be generated after finitely many steps. One might additionally require that every arithmetical truth somehow is made true after finitely many steps in the process.

It is not all that clear what this means, but perhaps we can illustrate the idea. The strings $\Box \forall m$ and $\Diamond \exists m$ express universal generalizations that at least seem to concern all the numbers, including ones not yet generated. A strict potentialist cannot allow such a generalization to be true merely in virtue of the space of possible worlds, considered in its entirety. Rather, the generalization must be made true after finitely many steps in the process. In short, a strict potentialist will insist that every question that has an answer, gets this answer in a finitary manner. We should never have to wait until the end of time.
Philosophical interlude

The distinction we have in mind can be illustrated with the so-called weak Brouwerian counterexamples. Can the potentialist say, now, that it either is or is not the case that every even number greater than 2, whenever it is generated, is the sum of two primes? Or whether there will be, at some point, a string 55555555 in the decimal expansion of $\pi$.

A classicist will say “yes” to both questions, as they are just instances of excluded middle. But from the point of view of a potentialist, both questions assume that there are determinate facts about what is, and is not, possible—facts about the structure of possibilities, so to speak. A strict potentialist will demur at those instances of excluded middle. For her, if the Goldbach conjecture (or the thesis about the decimal expansion of $\pi$), or its negation, is true, it must somehow be made true after a finite period of generation.
Philosophical interlude

We must accordingly distinguish between four main orientations towards the infinite. Actualism about the infinite unreservedly accept actual infinities and finds no use for modal notions in mathematics. Perhaps the actualist will say that the existence of actual infinities—such as the natural numbers and the real numbers—is necessary. These numbers exist in all worlds.

A second character asserts the (necessity of the) possible existence of an actual infinity:

$$\Box \Diamond \forall m \exists n \text{Successor}(m, n).$$

We will not cover this option further, nor will we give it a name, but modal actualism or GH might be appropriate. See [12].
Philosophical interlude

In contrast to these views, the potentialist insists that the objects with which mathematics is concerned are generated successively, such that at any one stage, there are never more than finitely many objects, but that we always (i.e., necessarily) have the ability to go on and generate more such objects.

As hinted above, there are two versions of this view. The liberal potentialist holds that there are determinate facts about what is and what is not possible. So, to repeat the above example, there is a fact of the matter concerning whether a counterexample to the Goldbach conjecture will ever be generated (and whether a given sequence will eventually occur in the decimal expansion of $\pi$).

If the name weren’t already taken, we might call this character a “modal realist”, since he is realist (in truth-value) about modality (but not about the existence of possible worlds).
Finally, the *strict potentialist* requires not only that every *object* be generated at some finite stage but also that every truth be “made true” after some finite number of steps. We will show that this loose talk about being ‘made true’ can be made formally precise and that the resulting version of strict potentialism can be satisfied. We will find that this is where distinctively intuitionist ideas enter and where intuitionistic logic is appropriate.
It is important to notice, however, that strict potentialism need not be identified with any form of anti-realism. To be sure, it is true that anti-realism provides one way to motivate strict potentialism. Indeed, an anti-realist about the modality in play would not allow any truths to obtain solely in virtue of future, unrealized possibilities. Since these possibilities aren’t (yet) real, they cannot contribute to any truths. However, one need not be an anti-realist to take an interest in strict potentialism. One’s interest may be the purely methodological one of studying the mathematics that can be established with only the weakest possible reliance on the infinite.
The modal framework

Here we sketch the modal framework used for explicating the above four positions concerning the infinite.

We invoke the contemporary heuristic of possible worlds, but we understand this as *only* heuristic, as a manner-of-speaking. The theory itself is formulated, officially, in the modal language, with the modal operators primitive. The modal language will be rock bottom, not explained or defined in terms of anything else.

The framework takes its inspiration from Linnebo [25], which develops a modal explication of the Cantorian notion that the universe of set theory is itself potential. For the time being, we remain neutral as to whether the background logic is classical or intuitionistic.
To invoke the heuristic, the idea is that some “possible worlds” have access to other possible worlds that contain objects that have been “generated” from those in the first world. From the perspective of the earlier world, the object in the second exist only potentially.

One sort of construction is geometric, following Euclid: the later world may contain, for example, a bisect of a line segment in the first. Or the later world might contain an extension of a line segment from the first world. Other sorts of constructions are arithmetic: the later world might contain more natural numbers than those of the first, say the successor of the largest natural number in the first world. Or, for a third kind of example, the later world may contain a set whose members are all in the first world. Or, to look ahead, a given sequence may have one (or more) elements in the later world than it has in the first world.
An Aristotelian would assume that every possible world is finite, in the sense that it contains only finitely many objects. That is in line with the rejection of the actually infinite. We make no such assumption here, however. Our goal is to contrast the actually infinite and the potentially infinite, so we need a framework where both can occur (to speak loosely). An actual infinity—or, to be precise, the possibility of an actual infinity—is realized at a possible world if it contains infinitely many objects.

We also assume, without much in the way of argument, that objects are not destroyed in the process of construction or generation. That is, we accept generation, but not corruption. Suppose, for example, that a given line segment is bisected. Then the resulting “world” contains the two bisects, as well as the original line segment.
The modal framework

Of course, we must say something about the modality that we invoke, which motivates a specific modal logic. To continue the heuristic, it follows from the foregoing that the domains of the possible worlds grow along the accessibility relation. So we assume:

\[ w_1 \leq w_2 \rightarrow D(w_1) \subseteq D(w_2), \]

where ‘\( w_1 \leq w_2 \)’ says that \( w_2 \) is accessible from \( w_1 \), and for each world \( w \), \( D(w) \) is the domain of \( w \).

For present purposes, we can think of a possible world as determined completely by the mathematical objects—regions, numbers, sets, etc.—it contains. So we can add that if \( D(w_1) = D(w_2) \), then \( w_1 = w_2 \); and we can strengthen the above to a biconditional. However, the above conditional is sufficient for our technical purposes.
The modal framework

As is well-known, the above conditional entails that the converse Barcan formula is valid. That is,

$$\exists x \Diamond \phi(x) \rightarrow \Diamond \exists x \phi(x)$$

(CBF)

This alone makes it doubtful that the modality in question can be “ordinary” metaphysical modality—whatever exactly that is. For it is widely held that there are objects whose existence is metaphysically contingent (Williamson [42] notwithstanding). For example, let \( \phi(x) \) say (or entail) that \( x \) does not exist. Presumably, there is someone, such as Aristotle, or a given line segment, that might not have existed. So we have \( \exists x \Diamond \neg \phi(x) \). But then it would follow via (CBF) that it is possible for there to exist something that doesn’t exist, which is absurd.
We are also rejecting the common thesis that mathematical sentences are necessary, if true, in this sense of modality anyway. We do have that once a mathematical object comes into existence—by being constructed—it continues to exist, of necessity.
The modal framework

One plausible option is that the modality employed in the explication of potential infinity is a kind of restriction of “ordinary” metaphysical modality. In terms of possible worlds, the relevant modality is the one that results from restricting the accessibility $R$ associated with metaphysical modality by imposing that additional requirement that domains only ever increase along the accessibility relation.

More explicitly, we define:

$$w_1 \leq w_2 :\iff w_1 R w_2 \& D(w_1) \subseteq D(w_2).$$
An alternative response is to sever any link between the metaphysical modality and the modality invoked in explicating potential infinity. Instead, we might regard the latter as an altogether distinct kind of modality, say the logico-mathematical modality of Putnan [32] or Hellman [12], or the interpretational modality of Fine [11] or Linnebo [25].

We remain neutral on this matter of metaphysics.
Our next issue concerns the right \textit{logic} for the modality we use to explicate potential infinity. Again, it will be useful to indulge in talk about possible worlds, writing the associated accessibility relation as $\leq$. Recall that $w \leq w'$ means that we can get from $w$ to $w'$ by generating more objects. This motivates the following principle:

\textbf{Partial ordering:} \textit{The accessibility relation $\leq$ is a partial order. That is, it is reflexive, transitive, and anti-symmetric.}

We can also require the accessibility relation to be well-founded, on the ground that all mathematical construction has to start somewhere. Nothing of substance turns on this, however.
The modal logic

At any given stage in the process of construction, we may have a choice of which objects to generate. This seems especially relevant in geometry, although it applies in set theory and real analysis. For example, given two intervals that don’t yet have bisections, we can choose to bisect one or the other of them, or perhaps to bisect both simultaneously.

Assume we are at a world $w_0$ where we can choose to generate objects, in different ways, so as to arrive at either $w_1$ or $w_2$. It makes sense to require that the licence to generate a mathematical object is never revoked as our domain expands. The option to, for example, bisect a given line segment can always be exercised at a later stage.
The modal logic

This corresponds to a requirement that the two worlds $w_1$ and $w_2$ can be extended to a common world $w_3$. This property of a partial order is called directedness and formalized as follows:

$$\forall w_1 \forall w_2 \exists w_3 (w_1 \leq w_3 \& w_2 \leq w_3)$$

We therefore adopt the following principle.

**Directedness**: *The accessibility relation $\leq$ is directed.*
The modal logic

Directedness ensures that, whenever we have a choice of mathematical objects to generate, the order in which we choose to proceed is irrelevant. Whichever object(s) we choose to generate first, the other(s) can always be generated later. Unless $\leq$ was directed, our choice whether to extend the ontology of $w_0$ to that of $w_1$ or that of $w_2$ might have an enduring effect.
NOTE: We will have occasion to consider non-directed systems, when we come to explicate mathematical constructions with some genuine indeterminacy. The key examples are Brouwerian choice sequences.
The modal logic

The mentioned properties of the accessibility relation $\leq$ allow us to identify a modal logic appropriate for studying the generation of mathematical objects. Since $\leq$ is reflexive and transitive, the modal logic S4 will be sound with respect our intended system of possible worlds.

As is well known, the directedness of $\leq$ ensures the soundness of the following principle as well:

$$\Diamond \Box p \rightarrow \Box \Diamond p.$$  \hfill (G)

The modal propositional logic that results from adding this principle to a complete axiomatization of S4 is known as S4.2.

As already discussed, we also have the Converse Barcan Formula, which means that S4.2 can be combined with an ordinary theory of quantification with no need for any complications such as a free logic or an existence predicate.
The modal logic

A note on \((G)\):

\[
\diamondsuit \Box p \rightarrow \Box \diamondsuit p. \tag{G}
\]

The underlying idea is that the only difference between a given world \(w\) and an accessible world \(w'\) is that the latter may contain some objects that have been generated from the first. So if a proposition \(p\) is made necessary in \(w'\), that must be due entirely to the existence of those generated objects.

Given directedness, those same new objects can be generated from any world that is accessible from \(w\). So the proposition \(p\) in question can be made necessary.
As noted above, two different kinds of generalization are available in the modal framework. First, there are the generalizations expressed by the ordinary quantifiers \( \forall \) and \( \exists \). Since the variables range just over the ontology of the relevant world (so to speak), this is an \textit{intraworld} form of generality.

That is, a sentence in the form \( \forall x \phi \), for example, is true at a world \( w \) just in case \( \phi \) holds of all objects in \( D(w) \), the domain of \( w \).
The modal logic

But there is also another, *transworld* form of generality available, expressed by the complex strings $\Box \forall$ and $\Diamond \exists$. These strings have the effect of generalizing not just over all entities at the relevant world, but over all entities at all (accessible) worlds.

This idea will receive a precise statement in two “mirroring” theorems that we will state shortly. Loosely speaking, this theorem says that, under some plausible assumptions, the strings $\Box \forall$ and $\Diamond \exists$ behave logically just like quantifiers ranging over all entities at all (future) worlds.

We refer to these strings $\Box \forall$ and $\Diamond \exists$ as *modalized quantifiers*, although they are strictly speaking composites of a modal operator and a quantifier proper.
The modal logic

Some definitions: First, given a non-modal formula \( \phi \) of a given mathematical language, its potentialist translation \( \phi^\Diamond \) is the formula that results from replacing each ordinary quantifier in \( \phi \) with the corresponding modalized quantifier. That is, ‘\( \forall x \)’ is replaced by ‘\( \square \forall x \)’; and ‘\( \exists x \)’ is replaced by ‘\( \Diamond \exists x \)’.

Say that a formula is fully modalized just in case all of its quantifiers are modalized.
We say that a formula $\phi$ is \textit{stable} if the necessitations of the universal closures of the following two conditionals hold:

\[
\phi \rightarrow \Box \phi
\]

\[
\neg \phi \rightarrow \Box \neg \phi
\]

Intuitively, a formula is stable just in case it never “changes its mind”, in the sense that, if the formula is true (or false) of certain objects at some world, it remains true (or false) of these objects at all “later” worlds as well.
We can now state our mirroring theorems.

First, let $\vdash$ be the relation of classical deducibility in a language $L$.

To be clear, this means classical deducibility without the use of any plural or higher-order comprehension axioms. Those are regarded here as “non-logical”. This will loom large below.

Now let $\vdash^\Diamond$ be deducibility in the modal language corresponding to $L$, by $\vdash$, S4.2, and the stability axioms for all atomic predicates of $L$. 
Theorem 2 (classical mirroring): For any formulas $\phi_1, \ldots, \phi_n, \psi$ of $L$, we have

$$\phi_1, \ldots, \phi_n \vdash \psi \quad \text{if and only if} \quad \phi_1 \boxplus, \ldots, \phi_n \boxplus \vdash \boxplus \psi \boxplus$$

See [25] for a proof.
This Mirroring Theorem tells us that, if we are interested in classical logical relations between fully modalized formulas in a modal theory that includes S4.2 and the stability axioms, we may delete all the modal operators and proceed by the ordinary non-modal logic underlying $\vdash$.

Thus, under the assumptions in question, the composite expressions $\Box \forall$ and $\Diamond \exists$ behave logically just like ordinary quantifiers, except that they generalize across all (accessible) possible worlds rather than a single world. This provides the desired bridge between actualist and potentialist theories.
It is important to notice that the potentialist translation is available only when we can assume that the accessibility relation is directed and that we thus have axiom (G). Without this axiom, the equivalence from the mirroring theorem breaks down.
There is a second mirroring theorem if the background modal logic is intuitionistic.

As usual, define a formula $\Phi$ to be *decidable* in a given intuitionistic theory if the universal closure of $(\Phi \lor \neg \Phi)$ is deducible in that theory.

Let $\vdash^I$ be the relation of *intuitionistic* deducibility in a language $L$ (again without the use of any plural or higher-order comprehension axioms). And let $\vdash^{I\diamond}$ be deducibility in the modal language corresponding to $L$, by $\vdash^I$, S4.2, the stability axioms for all atomic predicates of $L$, and the decidability of all atomic formulas of $L$. Then

*Theorem 3 (intuitionistic mirroring)*: For any formulas $\phi_1, \ldots, \phi_n, \psi$ of $L$, we have

$$\phi_1, \ldots, \phi_n \vdash^I \psi \quad \text{if and only if} \quad \phi_1^\diamond, \ldots, \phi_n^\diamond \vdash^{I\diamond} \psi^\diamond$$
We now see that one important difference among our four conceptions of the infinite arise in the background modal logic—whether it is classical or intuitionistic.

Recall that in both mirroring theorems, any plural or second-order comprehension axioms are considered to be non-logical. So a second difference among our conceptions of the infinite concerns which comprehension principles are accepted. This is significant, since virtually every mathematical theory uses higher-order axioms: induction for arithmetic, completeness for geometry, Dedekind- or Cauchy-completeness for real analysis, replacement for set theory. These are usually formulated using plural variables or second-order variables.
We begin with plural logic. The introduction and elimination rules for the plural quantifiers are not disputed by any of the views on infinity under consideration. However, interesting differences emerge concerning the plural comprehension axioms, which specify under what conditions a formula $\phi$ defines a plurality. The actualist (and modal actualist) answer is ‘always, provided that there is at least one $\phi$’. So the actualist accepts the traditional, unrestricted plural comprehension scheme:

$$\exists x \, \phi(x) \rightarrow \exists x x \forall u \, [u \prec xx \leftrightarrow \phi(u)],$$

(P-Comp)

provided only that $\phi$ does not contain ‘$xx$’ free.
Comprehension (plurals)

However, both the liberal potentialist and the strict potentialist must restrict the plural comprehension scheme. The restriction flows from their understanding of the quantifiers involved in the question whether an instance of (P-Comp) is valid. To make this understanding explicit, we apply the potentialist translation, so as to obtain:

\[ \Diamond \exists x \phi(x) \rightarrow \Diamond \exists xx \square \forall u [u < xx \leftrightarrow \phi(u)] \]  

(P-Comp\( \Diamond \))

Thus, when properly explicated, the question is for which formulas \( \phi \) that is possibly instantiated it is possible for there to be some objects \( xx \) which necessarily are all and only the \( \phi \)'s.
Comprehension (plurals)

We take pluralities to be modally rigid. That is, when \( x \) is one of some objects \( yy \), then this is necessarily so, at least on the assumption of the continued existence of \( yy \). And likewise when \( x \) is not one of \( yy \). If \( E \) is a predicate for the existence of an object or a plurality, then modal rigidity is expressed thus:

\[
\begin{align*}
x \prec yy & \rightarrow \Box(Eyy \rightarrow x \prec yy) \\
x \not\prec yy & \rightarrow \Box(Ex \land Eyy \rightarrow x \not\prec yy)
\end{align*}
\]

Modal rigidity is quite intuitive, on at least one prominent reading of the modal idiom. Consider Barack and Michelle. If Michelle were not one of some people, then these people would not be Barack and Michelle but some other people. Likewise, if Vladimir were one of some people, then these people would not be Barack and Michelle but some other people.
The rigidity of pluralities has dramatic consequences for our question of which instances of (P-Comp) are valid.

Consider for example the simple condition ‘$x = x$’ of being self-identical. Since this condition is obviously instantiated—something is self-identical in every world—the question is whether it is possible for there to be some objects $xx$ which necessarily are all and only the self-identical objects, that is, such that:

$$\Box \forall u(u \prec xx \iff u = u).$$
Comprehension (plurals)

The answer is ‘no’. By the rigidity of pluralities, the condition of being one of $xx$ is rigid and therefore not satisfied by more objects as we go to more populous possible worlds. However, the condition of being self-identical is necessarily satisfied by everything and thus must be satisfied by more objects as we go to more populous worlds. It follows that the two conditions cannot be necessarily coextensive, and hence that the corresponding instance of ($P$-Comp$\Diamond$) must be rejected.
More generally, our two potentialist views are only entitled to plural comprehension on conditions $\phi$ that it is possible to exhaust by some particular possible world. And, for both potentialists, the obvious restriction is that the formula $\phi$ can hold of at most finitely many objects.

So here we have a clear difference between both kinds of actualism, on the one hand, and the both kinds of potentialism, on the other—assuming that the mathematical theories in question are formulated in terms of plurals.
Comprehension (plurals)

Consider, for example, the question of what is the correct theory of natural numbers and pluralities thereof. Since actualists are entitled to unrestricted plural comprehension, they get a theory that is much like full classical second-order PA (the only difference being the trivial one that there is no empty class of numbers, as all pluralities must be non-empty).

By contrast, the two potentialists are committed to the view that all pluralities are finite. So their theory will be a plural variant of so-called weak second-order logic, where the second-order variables are stipulated to range over all and only finite (and rigid) collections from the first-order domain.
Comprehension (second-order)

Unlike plural logic, which generalizes into plural noun-phrase position, second-order logic (as we will henceforth use the term) generalizes into predicate position. Unlike the semantic value of a plural noun-phrase, which we have argued is modally rigid, there is no reason to expect the semantic value of a predicate to be rigid.

For example, although Socrates satisfies the predicate ‘is a philosopher’, he might not have done so. This means that the considerations that required the two kinds of potentialism to restrict plural comprehension, are not available in the case of second-order logic.
So what will our three conceptions of the infinite entail concerning second-order logic?

As before, there is complete agreement on the introduction and elimination rules for the second-order quantifiers. It is only concerning the second-order comprehension scheme

$$\exists F \forall x (Fx \leftrightarrow \phi(x))$$

that there is room for disagreement.

Both of our actualists have no reason to abandon unrestricted classical comprehension, where $\phi$ can be any formula whatsoever, provided only that it does not contain ‘$F$’ free.
The two potentialists will insist that the question be fully explicated as follows:

$$\diamond \exists F \Box \forall x (Fx \leftrightarrow \phi(x))$$

And here we get a sharp contrast between our two potentialists.
First, there is no obvious reason why the liberal potentialist should wish to restrict comprehension (beyond requiring that $\phi$ not contain ‘$F$’ free. Since the concept $F$ need not be modally rigid, it should be fine to let its application condition at any possible world be given by the condition $\phi(x)$ at that world. And since this potentialist is liberal, she has every reason to assume this condition to yield a determinate truth-condition at every world. In sum, on the question of second-order logic, there is no reason why the liberal potentialist should disagree with the actualist.

So if a given mathematical theory is formulated in a second-order language, then, at least from a mathematical point of view, there is no real difference between an actualist and a liberal potentialist.
What should the strict potentialist say about second-order logic? We believe she has good reasons to take a much stricter line than the actualist and liberal potentialist. In order to argue this, however, we first need to take a closer look at the view.

As we explained the view above, strict potentialism goes beyond liberal potentialism in requiring not only that the objects be generated in some incompletable process but also that every truth be ‘made true’ at some finite stage of this process. If we are serious about the process being incompletable, there can be no truths that obtain in virtue of the entire process—the entire space of possible worlds. Every truth must be true in virtue of some finite initial segment of the process.
Comprehension (second-order)

For atomic truths, the strict potentialist’s additional demand is unproblematic. Consider the truth that 1001 is larger than 7. This claim is wholly about the two mentioned numbers, and as soon as these numbers have been generated, we have everything in virtue of which the claim is true.

The problem arises when we consider the quantifiers, the universal quantifier in particular. In virtue of what is a generalization $\forall n \phi(n)$ true?

The classical, and perhaps most natural, answer is that it is true in virtue of every number being such as to satisfy $\phi$. But if this is our answer, then the claim can never be rendered true at some finite stage of the process of generating numbers, as the strict potentialist requires. In short, strict potentialism makes it hard to see how we can make sense of universal generalizations over all numbers.
In fact, however, there are ways to make sense of universal generalizations being made true at some finite stage. One option is the traditional intuitionistic approach, which equates mathematical truth with proof. The generalization is made true when we produce a proof of it.

Although this satisfies the strict potentialist’s requirement, it is an extreme form of anti-realism. The generative process is understood as a process of actual constructions, whereby both mathematical objects and proofs—which did not previously exist or obtain—are brought into being. So here we are entering the territory of orthodox intuitionism.
Another option is available, which avoids saddling strict potentialism with the extreme anti-realist views of orthodox intuitionism. Consider the following discussion by Hermann Weyl ([40], p. 54) of whether there is a natural number that has some decidable property $P$.

*Only the finding that has actually occurred of a determinate number with the property $P$ can give a justification for the answer "Yes," and—since I cannot run a test through all numbers—only the insight, that it lies in the essence of number to have the property [NOT] $P$, can give a justification for the answer "No"; Even for God no other ground for decision is available.*
On this view, the truth of the universal generalization—that every number is not-$P$—has nothing to do with epistemic matters, such as knowledge or proof. Even God, who is assumed to know all the facts, cannot know facts that require running through all the natural numbers—the point is that there are no such facts.

Instead, a universal generalization is “made true” by its lying “in the essence number” to have the relevant property. And presumably, more and more essential properties of number will become available as the generative process unfolds.
Admittedly, these are deep metaphysical waters, even if not exactly those of orthodox intuitionism. But the ideas in question admit of precise mathematical models.

In the case of arithmetic, at least, a good first approximation is provided by the realizability interpretation, going back to Stephen Cole Kleene [17], in 1945. The loose talk about what “lies in the essence of number” can be understood in terms of computable functions.
Comprehension (second-order)

Let \( \{e\}(n) \) be the result of applying Turing machine with index \( e \) to the input \( \bar{n} \). One can define what it is for a natural number \( e \) to be a realizer for a formula \( \phi \), written \( e \models \phi \). The idea is that \( e \) encodes information that establishes the truth of \( \phi \).

A useful metaphysical heuristic is that \( e \) functions as a “truth maker” for \( \phi \). Atomic formulas are realized by (codes of) finite computations. The most important clause is the one for the universal quantifier, where we define

\[
  e \models \forall n \phi(n) \quad \text{iff} \quad \forall n \{e\}(n) \models \phi(n)
\]

That is, \( e \) realizes the universal generalization \( \forall n \phi(n) \) just in case the Turing machine \( \{e\} \) computes a realizer for the instance \( \phi(n) \) when given any numeral \( \bar{n} \) as input.
In terms of our metaphysical heuristic: $e$ is a truth maker for $\forall n \phi(n)$ just in case $e$ specifies a function that maps any numeral $\bar{n}$ to a truth maker for the associated instance $\phi(n)$.

Let’s now regard a formula as true just in case it has a realizer or “truth maker”. Since a realizer is just a natural number, this means that any true formula is made true after finitely many steps. So our strict potentialist avoids having to wait until the end of time to make at least some universal generalizations true.
Of course, there remains the question of whether this definition yields the right truths. A natural measure of what is “right” is provided by the standard intuitionist theory of arithmetic, known as *Heyting arithmetic*, whose axioms are the same as those of first-order Peano-Dedekind arithmetic but where the logic is intuitionistic rather than classical.
Pleasingly, there is a theorem that says that Heyting arithmetic is sound with respect to the notion of truth that we have defined.

*Theorem 4.* Every theorem of Heyting arithmetic has a realizer. However, there are theorems of first-order Dedekind Peano arithmetic that do not have a realizer.
Summing up, we have described an interesting notion of arithmetical truth, which satisfies the strict potentialist's requirement that every truth be made true after some finite number of steps, and on which all the theorems of intuitionistic—but not classical—arithmetic are true. Notice also that the real locomotive of this argument is strict potentialism. There is no direct reliance on an anti-realist conception of the numbers.


Bohn, Einar Duenger, “Must there be a top level”, *Philosophical Quarterly 59* (2009), 193-201.

Cantor, G., “Mitteilugen zur Lehre vom Transfinitten 1, II”, Zeitschrift für Philosophie und philosophische Kritik 91 (1887), 81-125, 252-270; 92, 250-265; reprinted in [8], 378-439.


Niebergall, Karl-George “Is ZF finitistically reducible?”, in *One hundred years of Russell’s paradox: mathematics, logic, philosophy*,...


