Can there be natural instances of nonlinearity in the hierarchy of consistency strength?

Joel David Hamkins
Professor of Logic
Sir Peter Strawson Fellow

University of Oxford
University College, Oxford

UW Madison Logic Seminar
25 January 2021

The linearity phenomenon

It is a mystery often mentioned in the foundations of mathematics that our best and strongest mathematical theories seem to be linearly ordered and indeed well-ordered by consistency strength.

Given any two of the familiar large cardinal hypotheses, for example, generally one of them proves the consistency of the other.

Why should it be linear?

Significance of linearity

The linearity phenomenon is often seen as significant for the philosophy of mathematics.

Is the phenomenon directing us along the “one road upward” (Steel 2013), toward the final, ultimate mathematical truth?

Not actually linear

As a purely formal matter, the hierarchy of consistency strength is not well-ordered.

Actually, it is ill-founded, densely ordered, and nonlinear.

But the statements usually used to illustrate these features are often dismissed as unnatural—Gödelian trickery!

Many set theorists claim that amongst the *natural* assertions, consistency strengths are linearly ordered and indeed well ordered.

Exploring a contrary view

In this talk, I aim to explore and challenge that view.

Question

Can we find natural instances of nonlinearity and illfoundedness in the hierarchy of consistency strength?

I shall try my best.

Consistency strength

The consistency strength order on theories is defined so that $S \leq T$ just in case we can prove $\mathrm{Con}(T) \implies \mathrm{Con}(S)$ over a fixed base theory.

It is common to use PA as a base theory in the context of arithmetic and ZFC in a set-theoretic context.

But actually much weaker base theories usually suffice for most of the usual observations.

Formal instances of incomparability

Let us begin by setting aside the naturality requirement and establish nonlinearity as a purely formal matter.

These will be the often-dismissed ‘unnatural’ instances of incomparability.

Incomparable consistency strength

Theorem

There are statements σ and τ in the language of arithmetic with incomparable consistency strengths over PA.

Proof

Use the double fixed-point lemma to find two sentences:

σ asserts:

For any refutation of τ in $\mathrm{PA} + \mathrm{Con}(\mathrm{PA})$, there is a smaller refutation of σ.

τ asserts:

For any refutation of σ in $\mathrm{PA} + \mathrm{Con}(\mathrm{PA})$, there is a smaller refutation of τ.

Each asserts that it has smallest refutation, if either is refutable.

Proof, continued.

Neither σ nor τ is refutable, since then one would have smallest refutation and hence be provably true in PA.

Since σ is not refutable, there is a model of $\mathrm{PA} + \mathrm{Con}(\mathrm{PA}) + \sigma$, in which also this theory is inconsistent. So σ is refutable here in $\mathrm{PA} + \mathrm{Con}(\mathrm{PA})$, but also true. So it has smallest refutation. This is provable in PA. So this model has $\mathrm{Con}(\mathrm{PA} + \sigma)$ and $\neg\,\mathrm{Con}(\mathrm{PA} + \tau)$.

Similarly, since τ is not refutable, consider theory $\mathrm{PA} + \mathrm{Con}(\mathrm{PA}) + \tau$ analogously. We find a model with $\mathrm{Con}(\mathrm{PA} + \tau)$ but $\neg\,\mathrm{Con}(\mathrm{PA} + \sigma)$.

So σ,τ have incomparable consistency strength over PA. □

Don't need double fixed-point

Theorem

There is sentence η, such that η and $\neg\eta$ have incomparable consistency strengths over PA.

Proof

Simply let η be the Rosser sentence of $\mathrm{PA} + \mathrm{Con}(\mathrm{PA})$.

That is, η asserts that for any proof of η in $\mathrm{PA} + \mathrm{Con}(\mathrm{PA})$, there is a smaller proof of $\neg\eta$.

The usual Rosser argument shows η is neither provable nor refutable in that theory.

Proof, continued.

Since η not provable, we get a model of $\mathrm{PA} + \mathrm{Con}(\mathrm{PA}) + \neg\eta$. This model thinks η is provable with no smaller proof of $\neg\eta$. So $\mathrm{PA} \vdash \neg\eta$. So model thinks $\mathrm{Con}(\mathrm{PA} + \neg\eta) + \neg\,\mathrm{Con}(\mathrm{PA} + \eta)$.

Since η not refutable, we get model of $\mathrm{PA} + \mathrm{Con}(\mathrm{PA}) + \eta$, which thinks this theory inconsistent. Has proof of $\neg\eta$ from $\mathrm{PA} + \mathrm{Con}(\mathrm{PA})$, so by η there is such proof with no smaller proof of η. So $\mathrm{PA} \vdash \eta$ and so this is a model of $\mathrm{Con}(\mathrm{PA} + \eta) + \neg\,\mathrm{Con}(\mathrm{PA} + \neg\eta)$.

So neither $\mathrm{Con}(\mathrm{PA} + \eta)$ nor $\mathrm{Con}(\mathrm{PA} + \neg\eta)$ proves the other. So η and $\neg\eta$ have incomparable consistency strengths. □

Both η and $\neg\eta$ jump

Precisely because $\mathrm{PA}+\eta$ and $\mathrm{PA}+\neg\eta$ are incomparable in consistency strength, it follows that neither is equiconsistent with PA.

So this is a sentence η where both $\mathrm{PA}+\eta$ and $\mathrm{PA}+\neg\eta$ are strictly stronger than PA in consistency strength.

This is an instance of *double-jumping* in consistency strength.

No jumping

A sentence ρ with $\mathrm{PA} + \rho$, $\mathrm{PA} + \neg\rho$ both equiconsistent with PA. Example: Rosser sentence of PA.

Single jumping

A sentence such as $\mathrm{Con}(\mathrm{PA})$, strictly stronger than PA, but $\mathrm{PA} + \neg\mathrm{Con}(\mathrm{PA})$ equiconsistent with PA.

Double jumping

A sentence such as η above, where both $\mathrm{PA} + \eta$ and $\mathrm{PA} + \neg\eta$ are strictly stronger than PA in consistency strength.

Consistency strengths are dense

Theorem

If theories $S < T$ in consistency strength, then there is an intermediate theory $S < U < T$.

Indeed, there are incomparable theories $U \perp U'$ strictly between S and T.

Proof

Assume $S < T$ in consistency strength, both extending PA.

So $\mathrm{PA} + \mathrm{Con}(S) + \neg\mathrm{Con}(T)$ is consistent.

Let δ be the Rosser sentence of this theory.

Proof.

Since δ is not provable, get model of $\mathrm{Con}(S) + \neg\,\mathrm{Con}(T) + \neg\delta$. By $\neg\delta$, there is proof of δ with no smaller proof of $\neg\delta$. But PA can verify that, so $\mathrm{PA} \vdash \neg\delta$. So $\mathrm{Con}(S + \neg\delta) + \neg\,\mathrm{Con}(S + \delta) + \neg\,\mathrm{Con}(T)$.

Since δ not refutable, get model of $\mathrm{PA} + \mathrm{Con}(S) + \neg\,\mathrm{Con}(T) + \delta$, in which also this theory is inconsistent. So model thinks $\neg\delta$ is provable from $\mathrm{PA} + \mathrm{Con}(S) + \neg\,\mathrm{Con}(T)$. By δ, smallest such proof has no smaller proof of δ. PA proves this, so $S \vdash \delta$ here. So $\mathrm{Con}(S + \delta) + \neg\,\mathrm{Con}(S + \neg\delta) + \neg\,\mathrm{Con}(T)$.

Let $U = T \vee (S + \delta)$ and $U' = T \vee (S + \neg\delta)$.

So $\mathrm{Con}(U) = \mathrm{Con}(T) \vee \mathrm{Con}(S + \delta)$ and $\mathrm{Con}(U') = \mathrm{Con}(T) \vee \mathrm{Con}(S + \neg\delta)$.

The two models show U and U' are incomparable. Hence strictly intermediate between S and T. □

Natural instances of nonlinearity and ill-foundedness

I aim now at the hard task.

The hard task

To provide natural instances of nonlinear incomparability and ill-foundedness in the hierarchy of consistency strength, especially in the large cardinal hierarchy.

As well as I am able...

Finitely many inaccessible cardinals

Consider the assertions that there are some finite number of inaccessible cardinals.

"There are n inaccessible cardinals."

But there is a subtlety here, concerning how we describe n.

- Must we write n as $1 + 1 + \cdots + 1$?
- Consider "The number of inaccessible cardinals is at least the number of prime pairs."
- What if we want to say there are 2^{100} many inaccessible cardinals?
- Or a googol plex bang stack many?

Ultimately, we often describe a number n by specifying a computable procedure for computing it.

Nonlinearity in large cardinal assertions

Theorem

Amongst assertions of form,

"There are n inaccessible cardinals,"

where n is the output of a specific computational process, there are instances of incomparable consistency strength.

Indeed, there is a computable function f for which the assertions, "there are $f(n)$ many inaccessible cardinals," as n varies are incomparable and strongly independent with respect to consistency strength.

Strongly independent means that the consistency assertions freely generate the free countable Boolean algebra.

Universal computable function

The proof relies on an elementary version of the universal computable function.

Theorem

For any consistent $T \supseteq \mathrm{PA}$, there is a TM program e, such that for any function $f : \mathbb{N} \to \mathbb{N}$, there is model of T inside which program e on input n computes exactly $f(n)$.

Thus, the program e can compute any desired function at all, if only you run the program in the right model.

Proof.

Program e searches for a proof from T of a statement of form:

> *"the function computed by e is not precisely given by this input/output list:* $(k_0, n_0), \ldots, (k_r, n_r)$*."*

If found, then program e simply computes according to the list.

The point: you cannot refute any particular behavior for this function, since if you could, it would exhibit exactly that behavior.

Consequently, it is consistent that this program computes any desired particular values! □

Theorem

There is a computable function f for which the assertions

"There are $f(n)$ many inaccessible cardinals"

are incomparable and strongly independent in consistency strength.

Proof.

Let f be the universal computable function for theory $\mathrm{ZFC}+$"infinitely many inaccessible cardinals."

Let σ_n be the statement "there are $f(n)$ inaccessible cardinals."

If $n \neq m$, go to a model M where $f(n) < f(m)$, but both halt.

Inside that model, can find a model N where $\mathrm{ZFC} + f(n)$ many inaccessibles is consistent, but not $f(n) + 1$.

N agrees on $f(n) < f(m)$, and so $\mathrm{Con}(\mathrm{ZFC} + \sigma_n) + \neg\mathrm{Con}(\mathrm{ZFC} + \sigma_m)$.

Similar for any finite pattern of consistency. □

Pervasive natural instances of nonlinearity

The same argument works with Mahlo cardinals, measurable cardinals, and similarly with almost any of the usual large cardinal notions.

Thus, we find natural instances of incomparability for these kind of large cardinal existence assertions, throughout the large cardinal hierarchy.

A possible objection

Objection: f is not provably total.

Easily addressed.

Theorem

Amongst assertions of the form,

> *"There are as many supercompact cardinals as the running time of this specific computational process,"*

there are instances with incomparable consistency strength.

Indeed, there is a computable function f for which the supercompact running-time statements about $f(n)$ are incomparable and strongly independent with respect to consistency strength.

If these are a natural class of statements, then these are natural instances of incomparability.

More incomparability

Of course, the phenomenon is not about large cardinals.

It is about the comparative size of large natural numbers.

Theorem

Amongst the assertions of the form,

"This specific computational process halts,"

there are instances of incomparable consistency strength. Indeed, there are doubling-jumping instances.

There is a program e for which "e halts on input n" are incomparable and strongly independent in consistency strength.

Theorem

Let $A \subseteq \mathbb{N}$ be any computably enumerable nondecidable set.

1. *For any consistent theory $T \supseteq \mathrm{PA}$, there are true instances of $n \notin A$ that are independent of T.*
2. *The true instances of $n \notin A$ are not bounded in consistency strength by any consistent theory.*

Proof.

1. Otherwise you'd be able to decide A by searching for proofs.

2. By (1) there must be instances $n \notin A$ not provable in $\mathrm{PA} + \mathrm{Con}(T)$. So there is a model of $\mathrm{PA} + \mathrm{Con}(T) + n \in A$. But $n \in A$ implies $\mathrm{PA} \vdash n \in A$, so $\neg\mathrm{Con}(\mathrm{PA} + (n \notin A))$ here. So consistency strength of $n \notin A$ is not bounded by T. □

Consistency strength is inherent

Theorem

Let A be any computably enumerable m-complete decision problem.

1. *If A is* PA*-provably m-complete, with a provable reduction of the halting problem, then amongst true $n \notin A$, there are instances strictly exceeding any given consistency strength.*
2. *Even without extra assumption, amongst true $n \notin A$ there are instances with incomparable consistency strength, and double-jumping consistency strength.*
3. *Amongst consistent $n \in A$, there are instances of incomparable consistency strength and double-jumping strength.*
4. *Within both $n \notin A$, $n \in A$, there are effective enumerations that are strongly independent.*

Incomparable consistency strengths in tiling problem

Corollary

1. *Amongst assertions of the form*

   "These polygonal tiles admit a tiling of the plane"

   there are consistent instances strictly exceeding any given consistency strength.
2. *There are such assertions with incomparable consistency strength, and double-jumping strength.*
3. *There is an effective enumeration of finite tile sets $t_0, t_1, \ldots$, such that the assertions "tile set t_n admits a tiling of the plane" have incomparable and strongly independent consistency strengths.*

Are these natural instances of incomparability?

Incomparability is pervasive in mathematics

We can similarly find incomparable strongly independent consistency strength for instances of:

- The tiling problem
- Diophantine equations
- Word problem in group presentations
- Mortality problems in matrices
- And so on...

In each case, we find instances of incomparable strongly independent consistency strength.

Are these not therefore natural instances of incomparability in consistency strength?

Incomparability in any c.e. undecidable set?

I had used *m*-completeness to find incomparability in decision problems A, but I suspect a more general result is possible.

Open Questions

1. Does every computably enumerable undecidable set A admit statements $n \notin A$ of incomparable consistency strength?
2. And what of double-jumping consistency strength?
3. Can we find strongly independent families amongst statements of the form $n \notin A$, as well as $n \in A$?

Does incomparability occur in any c.e. undecidable problem?

Cautious enumeration

Let me now describe a completely different class of natural instances of incomparable consistency strength.

To begin, imagine that we believe in a certain computably enumerable theory, such as ZFC or ZFC plus large cardinals.

We aim to enumerate the axioms, but will do so *cautiously*.

We proceed with the enumeration, while at the same time looking out for some contrary indicator, a reason to doubt the actual truth of the theory, which might cause us to pause the enumeration.

Different cautious theory enumerations will pay attention to different specific indicators.

The cautious enumeration of ZFC

We enumerate the ZFC axioms as usual, proceeding as long as we find no proof in ZFC of $\neg\,\mathrm{Con}(\mathrm{ZFC})$.

If we find a proof of $\neg\,\mathrm{Con}(\mathrm{ZFC})$, we halt the enumeration.

- We don't have to see an actual contradiction, but rather just a proof that there is one.
- This is a c.e. theory, which we denote ZFC°.

The cautious enumeration ZFC° is sensible and realistic—it is what we would actually do when enumerating the theory.

Theorem

The cautious enumeration ZFC° has all the same axioms as ZFC, but with strictly lower consistency strength.

Proof.

If $\mathrm{Con}(\mathrm{ZFC})$ is consistent, we will never halt the enumeration. So ZFC° has all the ZFC axioms.

Consider a model $M \models \mathrm{ZFC} + \mathrm{Con}(\mathrm{ZFC}) + \neg\mathrm{Con}(\mathrm{ZFC} + \mathrm{Con}(\mathrm{ZFC}))$.

In M the enumeration ZFC° halts. But M thinks that ZFC proves consistency of any finite fragment. So M thinks $\mathrm{ZFC} \vdash \mathrm{Con}(\mathrm{ZFC}^\circ)$.

But M has a model $N \models \mathrm{ZFC} + \neg\mathrm{Con}(\mathrm{ZFC})$. This model agrees with M on ZFC°, so we have $\mathrm{ZFC} + \mathrm{Con}(\mathrm{ZFC}^\circ) + \neg\mathrm{Con}(\mathrm{ZFC})$.

So ZFC° has strictly weaker consistency strength than ZFC. □

Doubly cautious enumeration of ZFC

The *doubly cautious* enumeration $\mathrm{ZFC}^{\circ\circ}$ enumerates ZFC as usual, unless we find a proof in ZFC that ZFC is inconsistent or a proof in ZFC that there is such a proof of inconsistency.

In other words, we stop the enumeration when we find a proof either of $\neg\,\mathrm{Con}(\mathrm{ZFC})$ or $\neg\,\mathrm{Con}(\mathrm{ZFC}+\mathrm{Con}(\mathrm{ZFC}))$.

Theorem

The doubly cautious enumeration $\mathrm{ZFC}^{\circ\circ}$ is an alternative computable enumeration of ZFC, with strictly weaker consistency strength than even the cautious enumeration.

The proof is similar.

Ill-foundedness in consistency hierarchy

Iterating this idea, we find increasingly cautious enumerations of ZFC, strictly descending in consistency strength.

$$\cdots < \mathrm{ZFC}^{\circ\circ\circ} < \mathrm{ZFC}^{\circ\circ} < \mathrm{ZFC}^{\circ} < \mathrm{ZFC}$$

Thus, we have ill-foundedness in the consistency strength hierarchy.

All these theories are natural. They codify the sensible precautions we would actually take when actually enumerating ZFC.

So this is a natural instance of ill-foundedness in the hierarchy of consistency strength.

Cautious enumeration of ZFC plus an inaccessible

The same idea works higher up with essentially any theory.

Let I be the assertion "there is an inaccessible cardinal."

Theorem

The cautious enumeration of $\mathrm{ZFC} + I$ *has consistency strength strictly between* ZFC *and* $\mathrm{ZFC} + I$*.*

Proved in a similar manner.

The increasingly cautious enumerations constitute a natural instance of ill-foundedness:

$$\mathrm{ZFC} < \cdots < (\mathrm{ZFC} + I)^{\circ\circ\circ} < (\mathrm{ZFC} + I)^{\circ\circ} < (\mathrm{ZFC} + I)^{\circ} < \mathrm{ZFC} + I$$

Incomparability in large cardinal hierarchy

Consider the assertions:

> *"The number of inaccessible cardinals is at least the running time of $f(k)$."*

where f is the universal computable function for $\mathrm{ZFC} + I_\omega$.

Each is a cautious version of I_ω, since the number of inaccessibles is reduced only when $f(k)$ halts, which is a sensible reason to view that theory as incorrect.

But these cautious indicators are independent.

Theorem

The assertions above are incomparable and strongly independent in consistency strength. Each is strictly stronger than every I_n, but strictly weaker than I_ω.

Natural nonlinearity throughout

Of course, there are analogues of this theorem throughout the large cardinal hierarchy.

For every large cardinal notion, there are cautious enumerations of the theory that there are infinitely many, which have incomparable and strongly independent consistency strength.

The cautious theories can be seen as natural, since each is willing to weaken the large cardinal assertion in the face of contrary indicators of actual truth.

This is what we would actually do.

What is “natural”?

Let us consider philosophically the nature of “natural” examples.

What does it mean to have a “natural” example in mathematics?

- Arising in practice
- Arising “in the wild”
- Especially, arising in other mathematical subjects
- No weirdness. No strange constructions

Mathematicians commonly adopt a *know-it-when-you-see-it* attitude to naturality.

But the concept can be criticised

Concerns about naturality are often used merely to reject unfamiliar ideas or constructions.

What counts as natural changes over time. Ideas once considered unnatural are now seen as fundamental.

My view

- There is no coherent concept of what counts as natural.
- Naturality talk is too often used to reject the unfamiliar.
- “Unnatural” solutions often indicate that a question wasn’t well formulated.
- We may legitimately dismiss concerns about naturality as easily as they are raised.

Analogy with hierarchy of Turing degrees

- The Turing degrees form a complex hierarchy.
- Yet, the "natural" decision problems invariably arise in a linear, well-ordered part of the hierarchy:

$$0 < 0' < 0'' < 0''' < \cdots$$

- But computability theorists do not point to a "linearity phenomenon" in the Turing degrees, calling out for philosophical explanation.

Most computability theorists would say that to study only $0^{(n)}$ or $0^{(\alpha)}$ would be to miss the fundamental properties and essential nature of the hierarchy of Turing degrees.

Why don't set theorists have this attitude toward the consistency strengths?

Natural kinds vs. natural instances

The tiling problem is a natural kind of problem.

Does this mean that every tiling problem is natural?

Perhaps not.

Similar for diophantine equations, word problem, and so on.

Distinction between natural kinds and natural instances.

Nonstandardness objection

Objection

In my incomparable statements, the consistency-strength differences are only revealed in ω-nonstandard models.

So we don't care about these differences.

Rebuttal

Consistency strength is *inherently* about ω-nonstandard models. If we only looked at ω-standard models, they would all agree about the consistency of any given theory whatsoever, and there would be no hierarchy to speak of.

To show $S < T$ in consistency strength requires a model of $\mathrm{Con}(S) + \neg\mathrm{Con}(T)$, which must be ω-nonstandard if T is consistent.

Confirmation bias error

Throughout the large cardinal hierarchy, one generally proves that from any model of T we can construct a model of S.

In almost all cases, we proceed by forcing or inner models.

With these methods, I claim, we can never prove nonlinearity. The reason is that these methods preserve arithmetic truth.

To prove nonlinearity of consistency strength, we *must* change arithmetic truth.

Confirmation-bias moral: it shouldn't be surprising to observe only linearity, if our tools cannot observe nonlinearity.

Thank you.

Slides and articles available on http://jdh.hamkins.org.

Joel David Hamkins
Oxford University