NOTIONS OF CAUCHYNESS AND METASTABILITY

HANNES DIENER AND ROBERT LUBARSKY

Abstract. We show that several weakenings of the Cauchy condition are all equivalent under the assumption of countable choice, and investigate to what extent choice is necessary. We also show that the syntactically reminiscent notion of metastability allows similar variations, but in terms of its computational content is an empty notion.

§1. Almost Cauchyness. Apart from the last section, we work in Bishop style constructive mathematics [4]—that is mathematics using intuitionistic instead of classical logic and some appropriate set-theoretic or type theoretic foundation [1]. Unlike Bishop, however, we do not freely use the axiom of countable/dependent choice, but explicitly state every such use.

In [3] a weakened form of the usual Cauchy condition is considered. There a sequence $(x_n)_{n\geqslant 1}$ in a metric space (X, d) is called *almost Cauchy*, if for any strictly increasing $f, g : \mathbb{N} \to \mathbb{N}$

$$d(x_{f(n)}, x_{g(n)}) \to 0$$

as $n \to \infty$. (This property will be named C2 below). Unsurprisingly, and as indicated by its name, every Cauchy sequence is almost Cauchy. In the same paper mentioned above it is also shown that Ishihara's principle BD-N suffices to show the converse: that every almost Cauchy sequence is Cauchy. Thus the two conditions are equivalent not only in classical mathematics (CLASS), but also in Brouwer's intuitionism (INT) and Russian recursive mathematics á la Markov (RUSS) as in all these models BD-N holds. In fact, it was only recently that it has been shown that there are models[1] in which this principle fails [7, 10]. In this paper we will link the notion of almost Cauchyness to various other weakenings proposed by Fred Richman and investigate similarities and differences to the notion of metastability which was proposed by Terence Tao.

Without further ado we will start the mathematical part of the paper with the following convention: For two natural numbers n, m the interval $[n, m]$ will denote all natural numbers between n and m; notice that this notation does not necessitate $n \leqslant m$.

PROPOSITION 1. *Consider the following conditions for a sequence $(x_n)_{n\geqslant 1}$ in a metric space (X, d), where each condition should be read as prefaced by*

[1] As BISH is not formalised in the same spirit as normal, everyday, mathematics is formalised, we use the phrase "model of" here somewhat loosely. Of course there are strict formalisations of BISH and the structures falsifying BD-N are models of such formalisations.

"for every $\epsilon > 0$ and for all strictly increasing $f, g : \mathbb{N} \to \mathbb{N}$ there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$"

C1: $d(x_n, x_{g(n)}) < \varepsilon$
C1′: $\forall i, j \in [n, g(n)] : d(x_i, x_j) < \varepsilon$
C2: $d(x_{f(n)}, x_{g(n)}) < \varepsilon$
C2′: $\forall i, j \in [f(n), g(n)] : d(x_i, x_j) < \varepsilon$
C3: $d(x_{g(n)}, x_{g(n+1)}) < \varepsilon$
C3′: $\forall i, j \in [g(n), g(n+1)] : d(x_i, x_j) < \varepsilon$

The following implications hold.

$$\begin{array}{ccccc} C1' & \longleftrightarrow & C2' & \longleftrightarrow & C3' \\ \downarrow & & \downarrow & & \\ C1 & \longleftrightarrow & C2 & & \\ \downarrow & & & & \\ C3 & & & & \end{array}$$

Furthermore all *conditions are equivalent using countable choice.*

Proof. C1 implying C2 is a simple consequence of the triangle inequality—nevertheless, this small point is of importance in the later discussion. It is also trivial to see that Ci′ implies Ci for $i = 1, 2, 3$. With $f = \mathrm{id}$ one can also see that C1 is a special case of C2 and the same holds for C1′ and C2′. Similarly one can see that C2 implies C3 and C2′ implies C3′. Since $[f(n), g(n)] \subset [n, \max\{g(n), f(n)\}]$ for strictly increasing f and g C1′ implies C2′

To see that C3′ implies C1′ consider the intervals

$$G_n = [g^n(0), g^{n+2}(0)] \ .$$

We claim that for every n there is a k such that

$$[n, g(n)] \subset G_k \ . \tag{1}$$

To see this let $n \in \mathbb{N}$ be arbitrary. Since g is strictly increasing we can easily show by induction that $g^n(0) \geqslant n$. Therefore there exists $k \leqslant n$ such that

$$g^k(0) \leqslant n \leqslant g^{k+1}(0) \ .$$

Applying g to the second of these two inequalities we also get $g(n) \leqslant g^{k+2}(0)$, and thus $[n, g(n)] \subset G_k$.

Now consider the functions f and h defined by $f(n) = g^{2n}(0)$ and $h(n) = g^{2n+1}(0)$. By C3′ eventually

$$\forall i, j \in [f(n), f(n+1)] : d(x_i, x_j) < \varepsilon$$

and

$$\forall i, j \in [h(n), h(n+1)] : d(x_i, x_j) < \varepsilon \ .$$

Since for even n, say $n = 2k$, we have $G_n = [f(k), f(k+1)]$ and for odd n, say $n = 2\ell + 1$ we have that $G_n = [h(\ell), h(\ell + 1)]$, we can conclude that eventually

$$\forall i, j \in G_n : d(x_i, x_j) < \varepsilon$$

and thus have shown C1′.

For the rest of the proof we will assume countable choice and prove that C3 implies C3′, which in turn by transitivity will show that all conditions are equivalent. To this end let g be an arbitrary increasing function and $\varepsilon > 0$. For each n choose natural numbers i_n and j_n and a binary flagging sequence λ_n such that $g(n) \leqslant i_n < j_n \leqslant g(n+1)$ and

$$\lambda_n = 0 \Longrightarrow \forall i,j \in [g(n), g(n+1)] : d(x_i, x_j) < \varepsilon \ ,$$
$$\lambda_n = 1 \Longrightarrow d(x_{i_n}, x_{j_n}) > \frac{\varepsilon}{2} \ .$$

Notice that it might happen that $i_{n+1} = j_n$, but at least always $j_n < i_{n+2}$. Therefore, to get strictly increasing functions, we need to work with two functions f and g defined by $f(2n) = i_{2n}$, $f(2n+1) = j_{2n}$, $g(2n) = i_{2n+1}$, and $g(2n+1) = j_{2n+1}$. By C3 there is N and M such that for all $n \geqslant N$ $d(x_{f(n)}, x_{f(n+1)}) < \frac{\varepsilon}{2}$ and for all $n \geqslant M$ $d(x_{g(n)}, x_{g(n+1)}) < \frac{\varepsilon}{2}$. We claim that there cannot be $n \geqslant \max\{N, M\}$ such that $\lambda_n = 1$. For if there were such an even n we would have the contradiction

$$\frac{\varepsilon}{2} < d(x_{i_n}, x_{j_n}) = d(x_{f(n)}, x_{f(n+1)}) \ .$$

We can treat the odd case in a similar fashion, and therefore $\lambda_n = 0$ for all $n \geqslant \max\{N, M\}$, which is saying that C3′ holds. ⊣

We say that a sequence is *almost Cauchy* if it satisfies C2′ (and therefore any of the above properties).[2] Naturally, we are going to consider the following statement

(aCC) Every almost Cauchy sequence in a metric space is Cauchy.

As mentioned above, it is shown in [3] that BD-N implies that every almost Cauchy sequence is Cauchy, and therefore

$$\text{BD-N} \Longrightarrow \text{aCC} \ .$$

In the same paper it is also shown that if one drops the triangle inequality and works with so called semi-metric spaces, then this is in fact an equivalence. However, the Berger, Bridges, and Palmgren proof crucially needs semi-metric spaces instead of metric spaces, since in [8] it is shown that the statement that every almost Cauchy sequence is Cauchy implies BD-N is not provable within BISH. This is done by giving a topological model T. Notice that T also proves countable choice, so this result does not change, even if we switch to any of the other conditions of Proposition 1.

Next we can show that, conveniently, it is enough to consider the monotone, real case.

PROPOSITION 2. *(countable choice) If every decreasing/increasing sequence of reals that satisfies the almost Cauchy condition is Cauchy, then aCC holds.*

---

[2] Notice that in [3] "almost Cauchy" is defined as satisfying C2. Since in that work the authors assume countable choice this is not a conflicting definition. In the absence of choice it seems, to us, most natural to use the strongest notion.

Proof. Let x_n be a sequence satisfying C2′. First note that for every m the sequence defined by

$$r_n^{(m)} = \max_{i,j\in[m,m+n]}\{d(x_i,x_j)\}$$

is increasing. We want to show that it also satisfies the almost Cauchy condition. To this end let f and g be strictly increasing and $\varepsilon > 0$. We know, by C2′, that there is N such that for all $n \geqslant N$ and $i,j \in [m+f(n), m+g(n)]$

$$d(x_i,x_j) < \varepsilon/2 \ . \tag{2}$$

Now consider $k,\ell \in [f(n),g(n)]$. W.l.o.g. $\ell < k$. Then

$$\left|r_k^{(m)} - r_\ell^{(m)}\right| = \max_{i,j\in[m,m+k]}\{d(x_i,x_j)\} - \max_{i,j\in[m,m+\ell]}\{d(x_i,x_j)\} \ .$$

We will show that this distance is less than ε. First, by the definition of the maximum[3] we can choose $p,q \in [m,m+k]$ such that

$$d(x_p,x_q) > \max_{i,j\in[m,m+k]}\{d(x_i,x_j)\} - \varepsilon/2 \ .$$

We may assume that $p \leqslant q$. We have to distinguish three cases depending whether p and q are both to the left of $m+\ell$ or on both sides or to the right of it:

- If $q \leqslant m+\ell$, then

$$\max_{i,j\in[m,m+\ell]}\{d(x_i,x_j)\} \geqslant d(x_p,x_q)$$

and therefore

$$\begin{aligned}\max_{i,j\in[m,m+k]}\{d(x_i,x_j)\} &- \max_{i,j\in[m,m+\ell]}\{d(x_i,x_j)\} \\ &< d(x_p,x_q) + \varepsilon/2 - \max_{i,j\in[m,m+\ell]}\{d(x_i,x_j)\} \\ &\leqslant d(x_p,x_q) + \varepsilon/2 - d(x_p,x_q) < \varepsilon\end{aligned}$$

- If $p \leqslant m+\ell < q$, then

$$\max_{i,j\in[m,m+\ell]}\{d(x_i,x_j)\} \geqslant d(x_p,x_{m+\ell}) \ .$$

Furthermore if $m+\ell \leqslant q$, then $q \in [m+\ell, m+k] \subset [m+f(n), m+g(n)]$ and therefore $d(x_{m+\ell},x_q) < \varepsilon/2$ by Equation 2. Together

$$\begin{aligned}\max_{i,j\in[m,m+k]}\{d(x_i,x_j)\} &- \max_{i,j\in[m,m+\ell]}\{d(x_i,x_j)\} \\ &< d(x_p,x_q) + \varepsilon/2 - \max_{i,j\in[m,m+\ell]}\{d(x_i,x_j)\} \\ &\leqslant d(x_p,x_{m+\ell}) + d(x_{m+\ell},x_q) + \varepsilon/2 - \max_{i,j\in[m,m+\ell]}\{d(x_i,x_j)\} \\ &\leqslant d(x_p,x_{m+\ell}) + d(x_{m+\ell},x_q) + \varepsilon/2 - d(x_p,x_{m+\ell}) \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon\end{aligned}$$

[3] Notice that in general, even given just two numbers x,y we cannot decide whether $x = \max\{x,y\}$ or $y = \max\{x,y\}$, since that would imply the non-constructive *lesser limited principle of omniscience*. We can, however, given real numbers $x_1,\dots,x_n$ and $\varepsilon > 0$ find i such that $x_i > \max\{x_1,\dots,x_n\} - \varepsilon$.

- If $m+\ell \leqslant p$, then $p,q \in [m+\ell, m+k] \subset [m+f(n), m+g(n)]$ and therefore $d(x_p,x_q) < \varepsilon/2$ by Equation 2. Thus we have

$$\max_{i,j\in[m,m+k]}\{d(x_i,x_j)\} \quad < \quad d(x_p,x_q) \ + \ \varepsilon/2 \quad < \quad \varepsilon/2 \ + \ \varepsilon/2 \quad = \quad \varepsilon \quad .$$

And in particular

$$\max_{i,j\in[m,m+k]}\{d(x_i,x_j)\} - \max_{i,j\in[m,m+\ell]}\{d(x_i,x_j)\} \leqslant \max_{i,j\in[m,m+k]}\{d(x_i,x_j)\} < \varepsilon$$

That is in all cases for $n \geqslant N$ and $k,\ell \in [f(n), g(n)]$

$$\left| r_k^{(m)} - r_\ell^{(m)} \right| < \varepsilon \ ,$$

which means that the sequence $\left(r_n^{(m)}\right)_{n\geqslant 1}$ satisfies the almost Cauchy condition. Thus, by our assumption, it is Cauchy and converges to a limit, say y_m.

Since by definition $r_{n+1}^{(m)} \geqslant r_n^{(m+1)}$ we also have that in the limit $y_m \geqslant y_{m+1}$ ([4, Proposition 2.3.4.f]), so $(y_m)_{m\geqslant 1}$ is decreasing. We want to show that it also satisfies the almost Cauchy condition C2′. So let f and g be strictly increasing and $\varepsilon > 0$. Since $(r_k^{(i)})_{k\geqslant 1}$ converges to y_i for every n we can use countable choice to fix a function $h : \mathbb{N} \to \mathbb{N}$ such that

$$\forall i \in [f(n), g(n)] : |y_i - r_k^{(i)}| < \varepsilon/4$$

for all $k \geqslant h(n)$. Since x_n satisfies C2′ there exists N such that for all $n \geqslant N$ we have

$$\forall i', j' \in [\min\{f(n), g(n)\}, \max\{f(n), g(n)\} + h(n)] : d(x_{i'}, x_{j'}) < \varepsilon/4 \ .$$

Then, in particular, for all $i, j \in [f(n), g(n)]$ we have that

$$\begin{aligned} &\left| \max_{\ell,\ell'\in[i,i+h(n)]} d(x_\ell, x_{\ell'}) - \max_{p,p'\in[j,j+h(n)]} d(x_p, x_{p'}) \right| \\ &\qquad \leqslant \left| \max_{\ell,\ell'\in[i,i+h(n)]} d(x_\ell, x_{\ell'}) \right| + \left| \max_{p,p'\in[j,j+h(n)]} d(x_p, x_{p'}) \right| \\ &\qquad\qquad \leqslant \varepsilon/4 + \varepsilon/4 = \varepsilon/2 \ , \end{aligned}$$

since

$$[i, i+h(n)] \subset [\min\{f(n), g(n)\}, \max\{f(n), g(n)\} + h(n)]$$

and

$$[j, j+h(n)] \subset [\min\{f(n), g(n)\}, \max\{f(n), g(n)\} + h(n)] \ .$$

Combining all of this we get that for all $n \geqslant N$ and $i, j \in [f(n), g(n)]$

$$\begin{aligned} |y_i - y_j| &\leqslant |y_i - r_{h(n)}^{(i)}| + |y_j - r_{h(n)}^{(j)}| + |r_{h(n)}^{(i)} - r_{h(n)}^{(j)}| \\ &\leqslant \varepsilon/4 + \varepsilon/4 + \varepsilon/2 \ . \end{aligned}$$

So $(y_m)_{m\geqslant 1}$ is a Cauchy sequence converging to a limit $z \geqslant 0$. We want to show that $z = 0$. So assume[4] $z > 0$. That means that $y_m \geqslant z > 0$ for all $m \in \mathbb{N}$, since

[4]We remind the reader that even constructively equality is stable.

y_m is decreasing. Therefore, using countable choice, we can fix $g : \mathbb{N} \to \mathbb{N}$ such that

$$r^{(m)}_{g(m)} > {}^{z}\!/_{2} \ .$$

But if we apply property C2′ of $(x_n)_{n\geqslant 1}$ to $f = \mathrm{id}, \mathrm{id} + g$, and $\varepsilon = {}^{z}\!/_{2}$ we get that for $i, j \in [m, m + g(m)]$

$$d(x_i, x_j) < {}^{z}\!/_{2}$$

eventually, and therefore

$$r^{(m)}_{g(m)} = \max_{i,j\in[m,m+g(m)]} \{d(x_i, x_j)\} < {}^{z}\!/_{2}$$

eventually. This is a contradiction and thus $z = 0$

So $z = 0$ and since $d(x_i, x_j) \leqslant y_m$ for all $i, j \geqslant m$, we have shown that $(x_n)_{n\geqslant 1}$ is Cauchy. ⊣

§2. Metastability. In a program suggested by Terence Tao [13], it is proposed to recover the "finite" (constructive) content of theorems by replacing them with logically (using classical logic) equivalent ones that can be proven by finite methods. Since often there is no way to establish the Cauchy condition it is suggested to be replaced with the following notion of metastability. A sequence $(x_n)_{n\geqslant 1}$ in a metric space (X, d) is called *metastable* iff

$$\forall \epsilon > 0, f : \exists m : \forall i, j \in [m, f(m)] : d(x_i, x_j) < \varepsilon \ .$$

Notice that this is almost the same definition as C1′, and, in fact, one can easily show that an almost Cauchy sequence is metastable. However—as we will see—metastability contains almost no constructive content.

As noted in [2] every non-decreasing sequence of reals bounded by $B \in \mathbf{R}$ is metastable since it is impossible that $d(x_m, x_{f(m)}) > \frac{\varepsilon}{2}$ for all $1 \leqslant m \leqslant \frac{2B}{\varepsilon}$. How about the converse: is every non-decreasing metastable sequence bounded? There is no hope in finding a constructive proof since we will see that it is equivalent to the non-constructive *limited principle of omniscience*

(LPO) For every binary sequence $(a_n)_{n\geqslant 1}$ we can decide whether

$$\forall n \in \mathbb{N} : a_n = 0 \vee \exists n \in \mathbb{N} : a_n = 1.$$

Under the assumption of countable choice LPO is equivalent to deciding for all real numbers whether $x < 0 \vee x = 0 \vee 0 < x$. Countable choice is needed to given a real number x construct a sequence of rationals converging to x. LPO is also equivalent to even stronger statements:

Proposition 3. *(countable choice) LPO is equivalent to either of the following*

1. *The Bolzano Weierstraß theorem: every sequence of reals in* $[0, 1]$ *has a convergent subsequence.*
2. *For every binary sequence* $(a_n)_{n\geqslant 1}$

$$\exists N : \forall n \geqslant N : a_n = 0 \vee \exists k_n \in \mathbb{N}^{\mathbb{N}} : a_{k_n} = 1.$$

PROOF. The equivalence of LPO with the Bolzano Weierstraß theorem can be found in [11].

2 obviously implies LPO. Conversely we can show 2 by applying LPO countably many times: using LPO (and unique choice) construct a binary sequence b_n such that

$$b_k = 0 \Longrightarrow \exists n \geqslant k : a_n = 1$$
$$b_k = 1 \Longrightarrow \forall n \geqslant k : a_n = 0$$

Now, using LPO again, either $\exists N : b_N = 1$ or $\forall k : b_k = 0$. In the first case $\forall n \geqslant N : a_n = 0$. In the second case we can use unique choice[5] to find $k_n \in \mathbb{N}^{\mathbb{N}}$ such that $\forall n \in \mathbb{N} : a_{k_n} = 1$. ⊣

PROPOSITION 4. *(countable choice) LPO is equivalent to the statement that every metastable, non-decreasing sequence of rationals is bounded.*

PROOF. Assume $(x_n)_{n\geqslant 1}$ is non-decreasing and metastable. For every k we can fix, using LPO countably many times, a binary sequence $(\lambda^{(k)})_{n\geqslant 1}$ such that

$$\lambda_n^{(k)} = 0 \Longrightarrow x_n \leqslant k$$
$$\lambda_n^{(k)} = 1 \Longrightarrow x_n > k \ .$$

Then for every k, using LPO on $(\lambda_n^{(k)})_{n\geqslant 1}$, we can decide whether k is an upper bound of $(x_n)_{n\geqslant 1}$ or not. So we can fix another binary sequence η_k such that

$$\eta_k = 1 \Longrightarrow k \text{ is an upper bound}$$
$$\eta_k = 0 \Longrightarrow \exists \ell : x_\ell > k \ .$$

Using LPO yet again, we can thus either find an upper bound or, using dependent choice, we can fix a function $f : \mathbb{N} \to \mathbb{N}$ such that $x_{f(n+1)} > x_{f(n)} + 1$ for all $n \in \mathbb{N}$. Since x_n is non-decreasing f is increasing. Furthermore

$$d(x_{f(n+1)}, x_{f(n)}) > 1 \ ;$$

a contradiction to the metastability. Hence $(x_n)_{n\geqslant 1}$ is bounded.

Conversely, let $(a_n)_{n\geqslant 1}$ be a binary sequence that has, w.l.o.g., at most one 1. Now consider

$$x_n = \sum_{i=1}^{n} i a_i \ . \tag{3}$$

It is easy to see that x_n is metastable: if $f : \mathbb{N} \to \mathbb{N}$ is increasing, then either $a_i = 0$ for all $i \in [1, f(1)]$ or $a_i = 0$ for all $i \in [f(2), f(f(2))]$. In both cases x_i is constant on an interval of the form $[m, f(m)]$.

Now if x_n is bounded, there is $N \in \mathbb{N}$ with $x_n < N$. If there was $i > N$ with $a_i = 1$, then $x_i = i > N$ which is a contradiction. Hence $a_i = 0$ for all $i > N$, that is we only need to check finitely many entries to see if $(a_n)_{n\geqslant 1}$ consists of 0s or whether there is a term equalling 1. ⊣

[5] To use unique choice we need to always pick the *smallest* $k_{n+1} > k_n$ such that $a_{k_{n+1}} = 1$.

Since the construction of the sequence in the proof above (see Equation 3) relies on the terms being potentially very large one might still hope that there is maybe a chance that every *bounded*, metastable sequence converges. However, also this statement is equivalent to LPO.

Proposition 5. *(countable choice) LPO is equivalent to the statement that every bounded, metastable sequence of rationals converges.*

Proof. Assume that LPO holds and that $(x_n)_{n\geqslant 1}$ is a bounded and metastable sequence of rationals. Since LPO implies the Bolzano Weierstraß theorem (see Proposition 4) there exists $x \in \mathbf{R}$ and $k_n \in \mathbb{N}^{\mathbb{N}}$ such that x_{k_n} converges to x. Now let $\varepsilon > 0$ be arbitrary. For every $n \in \mathbb{N}$ we can use LPO to decide whether

$$|x - x_n| < \varepsilon \vee |x - x_n| \geqslant \varepsilon \ .$$

So, using (unique) countable choice we can fix a binary sequence $(\lambda_n)_{n\geqslant 1}$ such that

$$\lambda_n = 0 \Longrightarrow |x - x_n| < \varepsilon$$
$$\lambda_n = 1 \Longrightarrow |x - x_n| \geqslant \varepsilon \ .$$

By Proposition 4 either there exists N such that $\lambda_n = 0$ for all $n \geqslant N$ or there exists a strictly increasing $\ell_n \in \mathbb{N}^{\mathbb{N}}$ such that $\lambda_{\ell_n} = 1$ for all $n \in \mathbb{N}$. We will show that the second alternative is ruled out by the metastability: fix M such that $|x_{k_n} - x| < \frac{\varepsilon}{2}$ for $n \geqslant M$ and hence

$$|x_{k_n} - x_{\ell_n}| \geqslant \frac{\varepsilon}{2} \text{ for } n \geqslant M \ . \tag{4}$$

Now define $f : \mathbb{N} \to \mathbb{N}$ by $f(n) = \max\{k_{n+M}, \ell_{n+M}\}$. Then f is increasing. Since $(x_n)_{n\geqslant 1}$ is metastable there exists m such that for all $i, j \in [m, f(m)]$ we have $|x_i - x_j| < \frac{\varepsilon}{2}$. Since $k_{m+M}, \ell_{m+M} \in [m + M, f(m)]$ we get the desired contradiction to 4.

Conversely let $(a_n)_{n\geqslant 1}$ be a binary sequence with at most one term equalling 1. We will show that $(a_n)_{n\geqslant 1}$ is metastable. So let $f : \mathbb{N} \to \mathbb{N}$ an increasing function. Now either there exists $i \in [1, f(1)]$ such that $a_i = 1$ or for all $i \in [1, f(1)]$ we have $a_i = 0$. In the first case, since $(a_n)_{n\geqslant 1}$ has at most one 1, for all $i \in [f(1) + 1, f(f(1) + 1)]$ we have $a_i = 0$. In both cases there exists m such that, regardless of $\varepsilon > 0$, we have

$$\forall i, j \in [m, f(m)] : |a_i - a_j| = 0 < \varepsilon \ ;$$

that is $(a_n)_{n\geqslant 1}$ is metastable. Now if this sequence converges it must converge to 0. So there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $a_N = 0$. So we only need to check finitely many indices $n \in \mathbb{N}$ for $a_n = 1$, and hence LPO holds. ⊣

It is only natural to ask how variants of metastability along the lines considered in the first section interact. To this end let us consider the properties $MS^{(\prime)}1 - 3$ of a sequence, which are the same as $C^{(\prime)}1 - 3$ in Proposition 1, only that they are read as being prefaced by "for every $\epsilon > 0$ and for all strictly increasing $f, g : \mathbb{N} \to \mathbb{N}$ there exists $n \in \mathbb{N}$." With this notation metastability, as defined above, is $MS1'$. Not surprisingly we can reuse large parts of the proof of Proposition 1 in the next one.

PROPOSITION 6. *The following implications hold between conditions $MS^{(\prime)}1-3$ for a sequence in a metric space.*

$$MS1' \leftrightarrow MS2' \leftrightarrow MS3'$$
$$\downarrow$$
$$MS2$$
$$\swarrow \quad \searrow$$
$$MS1 \qquad MS3$$

The proof is identical to the one of Proposition 1, apart from the proof that $MS1$ implies $MS2$ which does not translate and which leaves us therefore with fewer implications.

§3. Choice is Necessary. In recent years, there has been an increasing sensitivity to the use of countable choice in constructive mathematics. In Bishop's words, "meaningful distinctions deserve to be preserved" and some researchers have argued [12] that the distinctions which are removed by the use of countable choice are, indeed, meaningful.

So the elephant-in-the-room question raised by Section 1 is, whether the use of countable choice in Proposition 1 was really necessary. We will show in the next proposition that this is the case—at least to prove the equivalence between the weakest (C3) and the strongest (C3′) notion. This means that in the absence of choice the middle (C1) must be inequivalent to at least one of the other two levels, but it is not clear to which, and whether it is to both of them.

THEOREM 7. *C3 does not imply C3′.*

PROOF. The counter-example will involve a sequence x_n of reals. We will also make use of particular natural numbers $a < b$, with several counter-examples i, j coming from the interval $[a, b]$. These a and b will be non-standard, and also a non-standard distance apart. The sequence x_n will be 0 outside the interval $[a, b]$; within that interval, the sequence will increase by $2/(b - a)$ each step for the first half of that interval, up to a value of 1, and then decrease that same amount for each step in the second half, back down to 0. How does this help satisfy C3? For g's which take on no values in $[a, b]$, there is nothing to do, as then $d(x_{g(n)}, x_{g(n+1)})$ is always 0. For other g's, in the end we will see that we need concern ourselves only with standard ϵ. For values within $[a, b]$, whenever $g(n+1) - g(n)$ is standard, $d(x_{g(n)}, x_{g(n+1)})$ is infinitesimal and hence less than ϵ. Of course, from x_n one can easily define a and b, and so their midpoint $(a + b)/2$, which would ruin C3. To avoid this, we will need to fuzz x_n up, so that the earlier cases mentioned are the only ones that happen.

In order to accomplish all of this we will need to exercise some care in the choice not only of a and b but also in the model in which they are embedded. It is easiest to work with an ultrapower of the universe V. (For the model theory about to be used, see standard references, such as [5, Sec. 4.3].) Where $\mathfrak{c}$ is (the size of) the continuum, take an ultrapower M using a $\mathfrak{c}$-regular ultrafilter.

Then M is $\mathfrak{c}^+$ saturated over V ([5, Cororllary 4.3.14]). In the following, we will identify a set in V with its image in M. In particular, g refers both to a function (from $\mathbb{N}$ to $\mathbb{N}$) in V and to its image in M.

The point of the saturation is that the model realizes any type of size $\mathfrak{c}$. The type of interest to us is in a triple a, b, and k of (symbols standing for) natural numbers. Start by including the formulas $b > a, b-a > 0, b-a > 1, \ldots$, as well as $2^k \leq a$ and $b \leq 2^{k+1}$. This much is easily seen to be consistent, by compactness.

Toward realizing the first option listed above, consider an axiom which says "g takes on no values in (a, b)"; more formally,

$$\phi_g = \exists n : g(n) \leq a \wedge b \leq g(n+1) \ ,$$

for some $g : \mathbb{N} \to \mathbb{N}$ in V.

Of course, ϕ_g might not be consistent (with the rest of the type); consider for example the identity function. For those g's, we will go toward the second option from above. For any $g \in V$, and standard natural number β (for "bound"), let $\psi_{g,\beta}$ be

$$\forall k : (\text{if } g(k) \text{ or } g(k+1) \text{ is in the interval } (a,b), \text{ then } g(k+1) - g(k) < \beta'') \ .$$

Notice that there are only $\mathfrak{c}$-many formulas of the form ϕ_g and $\psi_{g,\beta}$. Let the type Ty be a maximal consistent extension of the starting formulas by ϕ_g's and $\psi_{g,\beta}$'s. By the $\mathfrak{c}^+$-saturation of M, Ty is realized in M. We would like to show that, for all increasing $h \in V$, either $\phi_h \in Ty$ or, for some β, $\psi_{h,\beta} \in Ty$.

If h takes on no values in (a, b), then ϕ_h is true, and hence consistent with Ty, and so by maximality is in Ty. Else consider the non-empty set

$$I_h = \{ \, h(k+1) - h(k) \mid h(k) \text{ or } h(k+1) \text{ is in } (a,b) \, \} \ .$$

If every member of I_h were standard, then, since I_h is definable in M, it has a standard bound, say β. Immediately, $\psi_{h,\beta}$ is true, and so is consistent with Ty, and therefore is in Ty. The other possibility is that I_h contains a non-standard element. There are several cases here.

The simplest case is that, for some k with $h(k+1)-h(k)$ non-standard, $a \leq h(k)$ and $h(k+1) \leq b$. In that case, a and b could be re-interpreted to be $h(k)$ and $h(k+1)$ respectively. That would still satisfy Ty, and make ϕ_h true, and again we would be done by maximality. If that does not happen, then, whenever k generates a non-standard element of I_h, either $h(k) < a$ or $h(k+1) > b$ (so I_h has size at most 2). We will show what to do when both of those possibilities occur (for different k's, of course). This will call for a two-step procedure. If only one of those possibilities occurs, then only one of those steps need be done.

Toward this end, we have $h(k) < a$, and also $a < h(k+1) < b$, else I_h would be empty. The first sub-case is that $h(k+1) - a$ is non-standard. Then, similarly to the above, we could re-interpret b as $h(k+1)$ (and leave a fixed), and get ϕ_h consistent with Ty. The other sub-case is that $h(k+1) - a$ is standard. Then we could interpret a as $h(k+1)$ (and leave b fixed). Re-interpreting Ty with this new choice of a, the new I_h has size 1. Now one considers the other choice of k, with $h(k) < b < h(k+1)$, and argues similarly.

So we conclude that Ty is complete in this sense. Returning to the model construction, we are going to work over a two-node Kripke model. To the bottom node associate V, and the top M. Take the full model F over that structure.

(For the definition of a full model, see [6].) Consider the sequence x_n described at the beginning of this work: x_n is 0 outside of $[a, b]$, increases starting at a by $2/(b-a)$ at each step up to a value of 1 at the midpoint, then decreases by the same amount back to 0 at b. It would be no trouble to show C3 for this sequence for $g \in V$. The problem is, the full model contains a lot more functions g than just those from V. In particular, from x_n, a and b are easily definable, which would kill C3 holding. So we need to hide things better. This can be done by working within a topological model (built over the full model), and then taking a sub-model of it.

Working in F, let the space T consist of all sequences y_n such that $|x_n - y_n| < 1/n$ (starting the indexing from 1, obviously), except for n in the interval $[a, b]$, in which $|x_n - y_n| < 2/(b-a)$. A basic open set is given by restricting each component y_n to an open interval. Let G_n be the generic. In passing, we mention that, by standard arguments, any $g : \mathbb{N} \to \mathbb{N}$ in the topological model is in its ground model, which in this case is the full model F. We need more than that: we need a model in which any such g, at least at $\bot$, is in V.

To this end, we build essentially $L[G]$. At $\top$, this would be unambiguous. That is, at $\top$ we have a topological model over M, which models IZF_{Ref}, the version with Reflection. It was shown in [9] that such a theory can define its version of L and show it to be a model of IZF_{Ref}. It is not immediately clear, though, that this construction is consistent with what we need to do at $\bot$. So we describe the situation at $\bot$, and bring $\top$ along for the ride, and show what we need to for both.

The definition of $L_\alpha[G]$, inductively on α an ordinal of M, is straightforward, and is the same as in classical set theory. For $\alpha \in V$ an ordinal, its image in the full model, for which we temporarily use the notation α_f (f for "full"), works as follows: $\bot \Vdash$ "$x \in \alpha_f$" iff for some $\beta < \alpha$, $\bot \Vdash$ "$x = \beta_f$; and $\top \Vdash$ "$x \in \alpha_f$" iff, in M, identifying α with its image under the elementary embedding into M, for some $\beta < \alpha$, $\bot \Vdash$ "$x = \beta_f$. Since α_f is in the full model, which is the ground model for the topological model, it is also in the topological model. So within the topological model, the set $L_{\alpha_f}[G]$ can be defined by induction. We do not know whether the topological model can separate ordinals of the form α_f from any others, or even whether the full model can do so, so the final step is done in V resp. M: at bottom, $L[G]$ is defined in V to be the union over the ordinals α of $L_{\alpha_f}[G]$, whereas at $\top$ that union is taken in M.

What remains to be shown? For one, IZF, which we postpone to the end. By the presence of G, it should be clear that C3′ fails, for $g(n) = 2^n$: even if the generic differs from x_n, it is only by an infinitesimal amount at each component. All that remains is that C3 holds for G. At $\top$, G is a Cauchy sequence, so that is taken care of. We need check C3 for G only at $\bot$.

At $\bot$, we need concern ourselves with only standard ϵ. For any such $\epsilon > 1/n$, n standard, let N be n. For $g \in V$, the whole set-up all along the way is to make C3 true for that g. So we will be done if we can show that any h is in V: if $\bot \Vdash h : \mathbb{N} \to \mathbb{N}$ then, for some $g \in V$, $\bot \Vdash h = g$.

By hypothesis, $\bot$ forces a standard value for h on each standard input. So that is the obvious choice for g: let g be such that $g(n)$ is the value forced by $\bot$ for $h(n)$. In our model, let X be $\{ n \mid h(n) \neq g(n) \}$. This could have only non-standard elements. We will show that it is decidable: for any n, either $T \models n \in X$ or $T \models n \notin X$. Then we will show that any decidable set is either empty or has a standard member. That then suffices.

For decidability: any value for $h(n)$ has to be forced by T, by standard arguments, as follows. If not, let O be a maximal open set forcing a value for $h(n)$. Pick a point on the boundary of O. What value could a neighborhood of it force? If there is no such neighborhood, then h is not total, so need not be considered. If it forces a different value than O does, then, by connectedness, consider the overlap: $h(n)$ is then no longer single-valued. So whatever value $h(n)$ has is forced by T. Now compare that to $g(n)$.

As for a non-empty decidable set X having standard members: If it has a member at all, consider the definition ϕ of X over $L_\alpha[G]$. With regard to the parameters in ϕ, one can unpack them by their definitions, ultimately reducing the parameters used to finitely many standard ordinals and G. We will in the course of this argument consider alternate interpretations of a and b. Of course, when doing so, there is no longer any reason to believe that C3 still holds, or that C3′ does not. This is of no matter for showing our current goal. The construction of T, and of $L[G]$, still makes sense, for any choice of $a < b$. Now consider the space T_1 based on the pair $a-1, b-1$. Notice that $T \cap T_1$ is a non-empty open subset of both T and T_1. So it forces the same facts about X that T does, and that T_1 does. So the interpretation of X stays the same when we shift a and b down by 1. Iterate this procedure until the lower number is some standard value, say $\overline{a}$, larger than all of the natural number parameters used in ϕ. Then hold $\overline{a}$ fixed, and reduce the upper number by one. By similar arguments, again X remains unchanged. Iterate until this upper number is standard, say $\overline{b}$. Call the space based on $\overline{a}$ and $\overline{b}$ U. In M, X is still interpreted the same way, so, in M, $U \models$ "X has a member." By elementarity, the same holds in V. Any such member there has to be standard: $V \models k \in X$. So X has a standard member.

Finally, we sketch briefly why IZF holds. For the axioms of Empty Set and Infinity, $\emptyset$ is definable over $L_0[G]$, and ω over $L_\omega[G]$. Pair and Union hold easily. Extensionality is valid because that is how equality is defined. $\in$-Induction holds, even though M is ill-founded, because $\in$-Induction holds in M. Reflection holds because it holds in V: If V_α is a Σ_n-elementary substructure of V, then the initial segment of $L[G]$ up to α is itself a Σ_n-elementary substructure of the whole thing. From this, Separation follows easily. For Power Set, given $x \in L[G]$, since the whole construction took place in V, the ordinals at which new subsets of x appear are bounded, and at the next level they can all be collected into one set. $\dashv$

REFERENCES

[1] Peter Aczel and Michael Rathjen, *Notes on constructive set theory*, ***Technical Report 40***, Institut Mittag-Leffler, The Royal Swedish Academy of Sciences, 2001.

[2] Jeremy Avigad, Edward T Dean, and Jason Rute, *A metastable dominated convergence theorem*, ***Journal of Logic and Analysis***, vol. 4 (2012).

[3] Josef Berger, Douglas Bridges, and Erik Palmgren, *Double sequences, almost cauchyness and BD-N*, ***Logic Journal of IGPL***, vol. 20 (2012), no. 1, pp. 349–354.

[4] Errett Bishop and Douglas Bridges, ***Constructive analysis***, Springer-Verlag, 1985.

[5] C.C. Chang and H.J. Keisler, ***Model theory: Third edition***, Dover Books on Mathematics, Dover Publications, 2013.

[6] Matthew Hendtlass and Robert Lubarsky, *Separating fragments of WLEM, LPO, and MP*, ***The Journal of Symbolic Logic***, (to appear).

[7] Peter Lietz, *From constructive mathematics to computable analysis via the realizability interpretation*, ***Ph.D. thesis***, TU Darmstadt, 2004.

[8] Robert Lubarsky and Hannes Diener, *Principles weaker than BD-N*, ***The Journal of Symbolic Logic***, vol. 78 (2014), no. 3, pp. 873–885.

[9] Robert S. Lubarsky, Intuitionistic L, pp. 555–571, Birkhäuser Boston, 1993, pp. 555–571.

[10] Robert S. Lubarsky, *On the failure of BD-N and BD, and an application to the anti-Specker property*, ***The Journal of Symbolic Logic***, vol. 78 (2013), no. 1, pp. 39–56.

[11] Mark Mandelkern, *Limited omniscience and the Bolzano-Weierstrass principle*, ***Bull. London Math. Soc.***, vol. 20 (1988), pp. 319–320.

[12] Fred Richman, *Constructive mathematics without choice*, ***Reuniting the antipodes: Constructive and nonstandard views of the continuum*** (Peter Schuster, Ulrich Berger, and Horst Osswald, editors), Synthese Library, Kluwer A. P., 2001.

[13] Terence Tao, *Soft analysis, hard analysis, and the finite convergence principle*, http://terrytao.wordpress.com/2007/05/23/soft-analysis-hard-analysis-and-the-finite-convergence-principle/, May 2007.