A metastable dominated convergence theorem

JEREMY AVIGAD

EDWARD T. DEAN

JASON RUTE

Abstract: The dominated convergence theorem implies that if (f_n) is a sequence of functions on a probability space taking values in the interval $[0, 1]$, and (f_n) converges pointwise a.e., then $(\int f_n)$ converges to the integral of the pointwise limit. Tao [26] has proved a quantitative version of this theorem: given a uniform bound on the rates of metastable convergence in the hypothesis, there is a bound on the rate of metastable convergence in the conclusion that is independent of the sequence (f_n) and the underlying space. We prove a slight strengthening of Tao's theorem which, moreover, provides an explicit description of the second bound in terms of the first. Specifically, we show that when the first bound is given by a continuous functional, the bound in the conclusion can be computed by a recursion along the tree of unsecured sequences. We also establish a quantitative version of Egorov's theorem, and introduce a new mode of convergence related to these notions.

1 Introduction

If (a_n) is a nondecreasing sequence of real numbers in the interval $[0, 1]$, then (a_n) converges, and hence is Cauchy. Say that $r(\varepsilon)$ is a *bound on the rate of convergence* of (a_n) if for every $\varepsilon > 0$, $|a_n - a_{n'}| < \varepsilon$ whenever n and n' are greater than or equal to $r(\varepsilon)$. In general, one cannot compute a bound on the rate of convergence from the sequence itself: such a bound is not even continuous in the data, since the sequence (a_n) can start out looking like a constant sequence of 0's and then increase to 1 unpredictably.

But suppose that instead of a bound on the rate of convergence, we fix a function $F : \mathbb{N} \to \mathbb{N}$ and ask for an m such that $|a_n - a_{n'}| < \varepsilon$ for every n and n' in the interval $[m, F(m)]$. Since the sequence (a_n) cannot increase by ε more than $\lfloor 1/\varepsilon \rfloor$ times,

at least one element of the sequence $0, F(0), F(F(0)), \ldots, F^{\lfloor 1/\varepsilon \rfloor}(0)$ has the desired property. Hence there is always such a value of m less than or equal to $F^{\lfloor 1/\varepsilon \rfloor}(0)$.

Now notice that not only is this bound on m easily computable from F and a rational $\varepsilon > 0$, but it is, moreover, entirely independent of the sequence (a_n). What has happened is that we have replaced the assertion

$$\forall \varepsilon > 0 \; \exists m \; \forall n, n' \geq m \; |a_n - a_{n'}| < \varepsilon$$

by a "metastable" version,

$$\forall \varepsilon > 0, F \; \exists m \; \forall n, n' \in [m, F(m)] \; |a_n - a_{n'}| < \varepsilon.$$

The two statements are logically equivalent: an m as in the first statement is sufficient for any F in the second, and, conversely, if the first statement were false for some $\varepsilon > 0$ then for every m we could define $F(m)$ to return a value large enough so that $[m, F(m)]$ includes a rogue pair n, n'. But whereas one cannot compute a bound on the m in the first statement from ε and (a_n), one can easily compute a bound on the second m that depends only on ε and F.

If (a_n) is any sequence, say that $M(F)$ is a *bound on the ε-metastable convergence of* (a_n) if the following holds:

> For every function $F : \mathbb{N} \to \mathbb{N}$ there is an $m \leq M(F)$ such that for every $n, n' \in [m, F(m)]$, $|a_n - a_{n'}| < \varepsilon$.

Then what we have observed amounts to the following:

- There is a bound on the ε-metastable convergence of (a_n) if and only if there is an m such that $|a_n - a_{n'}| < \varepsilon$ for all $n, n' \geq m$. Hence, a sequence (a_n) is Cauchy if and only if there is a bound on the ε-metastable convergence of (a_n) for every $\varepsilon > 0$.

- For every $\varepsilon > 0$ the function $M(F) = F^{\lfloor 1/\varepsilon \rfloor}(0)$ is a bound on the ε-metastable convergence of any nondecreasing sequence (a_n) of elements of the real interval $[0, 1]$.

Thus there is a sense in which the second statement provides a quantitative, uniform version of the original convergence theorem.

This transformation is an instance of Kreisel's "no-counterexample" interpretation [18, 20], which is, in turn, a special case of Gödel's *Dialectica* interpretation [2, 8, 14]. The particular example above is discussed by Kreisel [19, page 49]. Variations on this idea have played a role in the Green-Tao proof [9] that there are arbitrarily long arithmetic progressions in the primes, and in Tao's proof [26] of the convergence of

certain diagonal averages in ergodic theory. In these instances the Kreiselian trick takes the form of an "energy incrementation argument"; see also [25] and [27, Sections 1.3–1.4]. The Birkhoff and von Neumann ergodic theorems and generalizations have also been analyzed in these terms [3, 15, 16, 17].

Here we are concerned with measure-theoretic facts such as the dominated convergence theorem, which relate one mode of convergence to another. Inspired by Tao [26], our goal will be to show that from a suitable metastable bound on the first type of convergence, one can obtain a suitable metastable bound on the second; and that, moreover, the passage from the first to the second is uniform in the remaining data.

For example, if (f_n) is a sequence of measurable functions on a measure space $\mathcal{X} = (X, \mathcal{B}, \mu)$, then (f_n) is said to converge *almost uniformly* if for every $\lambda > 0$, there is a set A with measure at most λ such that $(f_n(x))$ converges uniformly for $x \notin A$. This is equivalent to saying that for every $\lambda > 0$ and $\varepsilon > 0$ there is an m such that $\mu(\{x \ : \ \exists n, n' \geq m \ |f_n(x) - f_{n'}(x)| \geq \varepsilon\}) < \lambda$, since for a fixed $\lambda' > 0$ we can choose a sequence (ε_i) decreasing to 0 and then, for each ε_i, apply this last statement with $\lambda = \lambda'/2^{i+1}$. Thus the fact that f_n converges almost uniformly can be expressed as follows:

(AU) $\forall \lambda > 0, \varepsilon > 0 \ \exists m \ \mu(\{x \ : \ \exists n, n' \geq m \ |f_n(x) - f_{n'}(x)| \geq \varepsilon\}) < \lambda.$

By manipulations similar to the ones above, (AU) has the following metastable equivalent:

(AU*) $\forall \lambda > 0, \varepsilon > 0, F \ \exists m \ \mu(\{x \ : \ \exists n, n' \in [m, F(m)] \ |f_n(x) - f_{n'}(x)| \geq \varepsilon\}) < \lambda.$

As above, say that $M(F)$ is a *bound on the λ-uniform ε-metastable convergence of (f_n)* if the following holds:

> For every F, there is an $m \leq M(F)$ such that
> $$\mu(\{x \ : \ \exists n, n' \in [m, F(m)] \ |f_n(x) - f_{n'}(x)| \geq \varepsilon\}) < \lambda.$$

In other words, fixing λ and ε, $M(F)$ provides a bound on a value of m asserted to exist by (AU*).

Egorov's theorem asserts that if $\mathcal{X}$ is a probability space and (f_n) converges pointwise almost everywhere, then it converges almost uniformly. In Section 3, we obtain the following quantitative version. Say that $M(F)$ is a *λ-uniform bound for the ε-metastable pointwise convergence of (f_n)* if the following holds:

> For every $F : \mathbb{N} \to \mathbb{N}$,
> $$\mu(\{x \ : \ \forall m \leq M(F) \ \exists n, n' \in [m, F(m)] \ |f_n(x) - f_{n'}(x)| \geq \varepsilon\}) < \lambda.$$

In other words, for every F, $M(F)$ provides a uniform ε-metastable bound for the convergence of each sequence $(f_n(x))$ outside a set of measure at most λ. Compare this to the previous definition: if $M(F)$ is a bound on the λ-uniform ε-metastable convergence of (f_n), then $M(F)$ provides a bound on a *single m* that works outside a set of measure at most λ. With this terminology in place, we can state our quantitative version of Egorov's theorem: given $\varepsilon > 0$, $\lambda > \lambda' > 0$, and a λ'-uniform bound $M_1(F)$ on the ε-metastable pointwise convergence of (f_n), there is a bound $M_2(F)$ on the λ-uniform ε-metastable convergence of (f_n); and moreover $M_2(F)$ depends only on ε, λ, λ', and $M_1(F)$, and not on the underlying probability space or the sequence (f_n). In fact, we provide an explicit description of $M_2(F)$ in terms of this data, and explicit bounds on the complexity of M_2 when M_1 is a computable functional that can be defined using Gödel's schema of primitive recursion in the finite types. The proof relies on a combinatorial lemma, presented in Section 2, whose proof can be viewed as an energy incrementation argument that is iterated along a well-founded tree.

It is easy to show that if (f_n) is a sequence of functions taking values in $[0, 1]$ and (f_n) converges almost uniformly, then the sequence $(\int f_n)$ converges. Thus the dominated convergence theorem follows easily from Egorov's theorem in the case where $\mathcal{X}$ is a probability space and the sequence (f_n) is dominated by a constant function. In a similar way, we show in Section 3 that our quantitative version of Egorov's theorem implies a quantitative version of the dominated convergence theorem, a mild strengthening of Theorem A.2 of Tao [26], again with an explicit description of the computation of one metastable bound from the other.

The notion of a λ-uniform bound on the ε-metastable pointwise convergence of a sequence gives rise to a new mode of convergence that sits properly between pointwise convergence and almost uniform convergence. In Section 4, we explore the relationships between these notions.

2 A combinatorial fact

This section is devoted to establishing a key combinatorial fact that underlies our quantitative convergence theorems. As a warmup, consider the following:

Proposition 2.1 *Let (A_n) be a sequence of measurable subsets of a probability space $\mathcal{X} = (X, \mathcal{B}, \mu)$, and let $\lambda > 0$. Then the following are equivalent:*

(1)   *There is an M such that $\mu(\bigcup_{n \geq M} A_n) < \lambda$.*

(2)   *There is an M such that for every function $F(m)$,*

$$\mu\left(\bigcap_{m \leq M} \bigcup_{n \in [m,F(m)]} A_n\right) < \lambda.$$

(3)   *There is a $\lambda' < \lambda$ such that for every F there is an M such that*

$$\mu\left(\bigcap_{m \leq M} \bigcup_{n \in [m,F(m)]} A_n\right) < \lambda'.$$

Proof   (1) clearly implies (2) because

$$\bigcap_{m \leq M} \bigcup_{n \in [m,F(m)]} A_n \subseteq \bigcap_{m \leq M} \bigcup_{n \geq m} A_n = \bigcup_{n \geq M} A_n,$$

and (2) clearly implies (3). To show (3) implies (1), fix $\lambda > \lambda' > 0$ and for each m, let $F(m)$ be large enough so that

$$\mu\left(\bigcup_{n \geq m} A_n \setminus \bigcup_{n \in [m,F(m)]} A_n\right) < (\lambda - \lambda')/2^{m+1}.$$

By hypothesis, for this F, there is an M such that $\mu(\bigcap_{m \leq M} \bigcup_{n \in [m,F(m)]} A_n) < \lambda'$. Then

$$\bigcup_{n \geq M} A_n = \bigcap_{m \leq M} \bigcup_{n \geq m} A_n \subseteq$$

$$\left(\bigcap_{m \leq M} \bigcup_{n \in [m,F(m)]} A_n\right) \cup \bigcup_{m \leq M} \left(\bigcup_{n \geq m} A_n \setminus \bigcup_{n \in [m,F(m)]} A_n\right),$$

whose measure is at most $\lambda' + \sum_{m \leq M}(\lambda - \lambda')/2^{m+1} < \lambda$. Hence $\mu(\bigcup_{n \geq M} A_n) < \lambda$, as required. $\square$

In particular, if (3) holds, there is an n such that $\mu(A_n) < \lambda$. Now suppose we are given a functional $M(F)$ witnessing (3). The main result of this section, Theorem 2.2, shows that there is a bound on n that depends only on $M(F)$, λ, and λ'. In particular, the bound is independent of $\mathcal{X}$ and the sequence (A_n).

Theorem 2.2   *For every functional $M(F)$ and $\lambda > \lambda' > 0$, there is a value M' with the following property. Suppose (A_n) is a sequence of measurable subsets of a probability*

space $\mathcal{X}$ with the property that for every function F,

$$\mu\left(\bigcap_{m\leq M(F)}\bigcup_{n\in[m,F(m)]} A_n\right) < \lambda'.$$

Then there is an $n \leq M'$ such that $\mu(A_n) < \lambda$.

A functional M is said to be *continuous* if the value of $M(F)$ depends on only finitely many values of F. Say that two functions F and F' *agree up to k* if $F(j) = F'(j)$ for every $j \leq k$. If M is continuous, a functional $k(F)$ with the property that $M(F) = M(F')$ whenever F and F' agree up to $k(F)$ is said to be a *modulus of continuity* for M.

The next lemma shows that, without loss of generality, we can assume the functional M in the hypothesis of Theorem 2.2 is continuous, because one can always replace it by a suitable continuous version, $\overline{M}$.

Lemma 2.3 *Given any functional M, there is a continuous functional $\overline{M}$ with the following property: for every F, there is an F' such that $\overline{M}(F) = M(F')$ and F and F' agree up to $\overline{M}(F)$.*

Proof Given M, define

$$\overline{M}(F) = \min\{M(F') \ : \ F \text{ and } F' \text{ agree up to } M(F')\}.$$

The last set is nonempty since it contains $M(F)$ itself. Clearly $\overline{M}(F)$ satisfies the stated condition, so we only need to show that $\overline{M}$ is continuous.

In fact, we claim that $\overline{M}$ is its own modulus of continuity. To see this, suppose F and F'' agree up to $\overline{M}(F)$. We need to show $\overline{M}(F) = \overline{M}(F'')$. By the definition of $\overline{M}$, there is an F' such that $\overline{M}(F) = M(F')$ and F and F' agree up to $M(F')$. But then F'' and F agree up to $M(F')$, and so $\overline{M}(F'') \leq M(F') = \overline{M}(F)$.

Since F and F'' agree up to $\overline{M}(F)$, *a fortiori*, they agree up to $\overline{M}(F'')$. But now the symmetric argument shows that $\overline{M}(F) \leq \overline{M}(F'')$. So $\overline{M}(F) = \overline{M}(F'')$. $\qquad\square$

The condition on $\overline{M}$ imposed by Lemma 2.3 ensures that any sequence (A_n) of a measure space $\mathcal{X}$ satisfying

$$\forall F \, \mu\left(\bigcap_{m\leq M(F)}\bigcup_{n\in[m,F(m)]} A_n\right) < \lambda'$$

also satisfies

$$\forall F \; \mu \left(\bigcap_{m \leq \overline{M}(F)} \bigcup_{n \in [m, F(m)]} A_n \right) < \lambda'$$

and so it suffices to prove Theorem 2.2 for $\overline{M}$ in place of M. By similar machinations, we could arrange that $\overline{M}(F) \leq \overline{M}(G)$ whenever F is pointwise less than or equal to G, and that $\overline{M}$ is determined by the values it takes on nondecreasing F's. However, we will not need these additional conveniences below.

Notice that the passage from M to $\overline{M}$ is noneffective; in general it will not be possible to "compute" $\overline{M}(F)$ from descriptions of M and F. We will show, however, that in the case where M is continuous, the M' in the conclusion of Theorem 2.2 *can* be computed from a suitable description of M.

To explain our algorithm, we need to establish some background involving computation on well-founded trees. If σ is a finite sequence of natural numbers, we index the elements starting with 0 so that $\sigma = (\sigma_0, \ldots, \sigma_{\text{length}(\sigma)-1})$, and write $\sigma\hat{\ }n$ to denote the sequence extending σ with an additional element n. If τ is another finite sequence of natural numbers, write $\sigma \subseteq \tau$ to indicate that σ is an initial segment of τ. By a *tree on* $\mathbb{N}$, we mean a set T of finite sequences of natural numbers that is closed under initial segments. Think of the empty sequence, (), as denoting the root, and the elements $\sigma\hat{\ }n$ as being the children of σ in the tree.

Identify functions F from $\mathbb{N}$ to $\mathbb{N}$ with infinite sequences, and write $\sigma \subset F$ if σ is an initial segment of F. A tree T on $\mathbb{N}$ is said to be *well-founded* if it has no infinite branch, which is to say, for every function F there is a $\sigma \subset F$ such that σ is not in the tree. One can always carry out a proof by induction on a well-founded tree: if P_σ is any property that holds everywhere outside a tree T and, moreover, P_σ holds whenever $P_{\sigma\hat{\ }n}$ holds for every n, then P_σ holds for every σ; otherwise, one could successively extend a counterexample σ to build an infinite branch F that never leaves the tree. By the same token, one can define a function on finite sequences of natural numbers by a schema of recursion:

$$G(\sigma) = \begin{cases} H(\sigma) & \text{if } \sigma \text{ is not in } T \\ K(\sigma, \lambda n. \; G(\sigma\hat{\ }n)) & \text{otherwise} \end{cases}$$

where $\lambda n. \; G(\sigma\hat{\ }n)$ denotes the function which maps n to $G(\sigma\hat{\ }n)$. Using induction on T, one can show that G is well-defined. Moreover, if T and the functions H and K are computable, so is G. For example, the computation of G on the empty string requires recursive calls to $G((n))$, for various n; these, in turn, require recursive calls to $G((n, n'))$, for various n', and so on. The well-foundedness of T guarantees that every branch of the computation terminates.

Now suppose $M(F)$ is a continuous functional. Say that a finite sequence σ is *unsecured* if there are F_1, F_2 extending σ such that $M(F_1) \neq M(F_2)$. In words, σ is unsecured if it does not provide sufficient information about a function F to determine the value of M. Let $T = \{\sigma \,:\, \sigma \text{ is unsecured}\}$. Then it is not hard to see that T is a tree, and the continuity of M implies it is well-founded.

Suppose moreover that $k(F)$ is a modulus of continuity for M. For any finite sequence σ of natural numbers, use $\hat{\sigma}$ to denote the function

$$\hat{\sigma}(n) = \begin{cases} \sigma_n & \text{if } n < \text{length}(\sigma) \\ 0 & \text{otherwise.} \end{cases}$$

One can check that the set $T' = \{\sigma \,:\, \forall \tau \subseteq \sigma \; k(\hat{\tau}) \geq \text{length}(\tau)\}$ is again a well-founded tree that includes T. In the next proof, given a continuous functional M, we will define a function $N(\sigma)$ by recursion on any well-founded tree that includes the tree of sequences that are unsecured for M. When this tree is given by a modulus of continuity, $k(F)$, as above, this amounts to the principle of *bar recursion*, due to Spector [24] (see also [2, 5, 14]).

We now turn to the proof of Theorem 2.2.

Proof By Lemma 2.3, we can assume without loss of generality that M is continuous. Fix $\lambda > \lambda' > 0$, and let T be any well-founded tree that includes all the sequences that are unsecured for M. We will define a function $N(\sigma)$ by recursion on T, and simultaneously show, by induction on T, that $N(\sigma)$ satisfies the following property, P_σ, for every σ: whenever $\mathcal{X}$ and (A_n) satisfy

$$(Q_\sigma) \qquad\qquad \forall F \supset \sigma \; \mu \left(\bigcap_{m \leq M(F)} \bigcup_{n \in [m, F(m)]} A_n \right) < \lambda'$$

and

$$(R_\sigma) \qquad \forall m < \text{length}(\sigma) \; \mu \left(\bigcup_{n \in [m, N(\sigma)]} A_n \setminus \bigcup_{n \in [m, \sigma_m]} A_n \right) \leq (\lambda - \lambda')/2^{m+1},$$

there is an $n \leq N(\sigma)$ such that $\mu(A_n) < \lambda$. In that case, $N(())$ is the desired bound, since $Q_{()}$ is the desired hypothesis, and $R_{()}$ is vacuously true.

In the base case, suppose σ is not in T, and hence secured for M. Define $N(\sigma) = M(\hat{\sigma})$. To see that $N(\sigma)$ satisfies P_σ, suppose $\mathcal{X}$ and (A_n) satisfy Q_σ and R_σ. Define $\tilde{\sigma}$ to be the function

$$\tilde{\sigma}(n) = \begin{cases} \sigma_n & \text{if } n < \text{length}(\sigma) \\ N(\sigma) & \text{otherwise.} \end{cases}$$

Since σ is secured and $\tilde{\sigma} \supset \sigma$, $M(\tilde{\sigma}) = M(\hat{\sigma}) = N(\sigma)$, and Q_σ implies

$$\mu \left(\bigcap_{m \leq N(\sigma)} \bigcup_{n \in [m, \tilde{\sigma}(m)]} A_n \right) < \lambda'.$$

Similarly, R_σ implies

$$\forall m \leq N(\sigma) \, \mu \left(\bigcup_{n \in [m, N(\sigma)]} A_n \setminus \bigcup_{n \in [m, \tilde{\sigma}(m)]} A_n \right) \leq (\lambda - \lambda')/2^{m+1},$$

since for $m \geq \text{length}(\sigma)$, $\tilde{\sigma}(m) = N(\sigma)$. We use a calculation similar to that of Proposition 2.1, with $N(\sigma)$ now playing the role of infinity.

$$A_{N(\sigma)} = \bigcap_{m \leq N(\sigma)} \bigcup_{n \in [m, N(\sigma)]} A_n \subseteq$$

$$\left(\bigcap_{m \leq N(\sigma)} \bigcup_{n \in [m, \tilde{\sigma}(m)]} A_n \right) \cup \bigcup_{m \leq N(\sigma)} \left(\bigcup_{n \in [m, N(\sigma)]} A_n \setminus \bigcup_{n \in [m, \tilde{\sigma}(m)]} A_n \right).$$

As before, the measure of this set is at most $\lambda' + \sum_{m \leq M} (\lambda - \lambda')/2^{m+1} < \lambda$, and so $N(\sigma)$ itself satisfies the conclusion of P_σ.

In the inductive case where σ is not in T, we can assume that we have already defined $N(\sigma\hat{\ }n)$ for every n so that $P_{\sigma\hat{\ }n}$ is satisfied. Define the sequence n_i by setting $n_0 = 0$ and $n_{i+1} = N(\sigma\hat{\ }n_i)$, set $\bar{m} = \text{length}(\sigma)$, and set $N(\sigma) = \max_{i \leq \lceil 2^{\bar{m}+1}/(\lambda-\lambda') \rceil} n_i$.

To show that $N(\sigma)$ satisfies P_σ, fix $\mathcal{X}$ and (A_n) satisfying Q_σ and R_σ. We need to show that there is an $n \leq N(\sigma)$ satisfying $\mu(A_n) < \lambda$. By the definition of $N(\sigma)$, this is the same as showing that for some $i \leq \lceil 2^{\bar{m}+1}/(\lambda - \lambda') \rceil$, there is an $n \leq n_i$ with this property.

Start by trying $i = 1$. Suppose the conclusion fails, that is, there is no $n \leq n_1$ satisfying $\mu(A_n) < \lambda$. Since $n_1 = N(\sigma\hat{\ }n_0)$ satisfies $P_{\sigma\hat{\ }n_0}$, this implies that either $Q_{\sigma\hat{\ }n_0}$ or $R_{\sigma\hat{\ }n_0}$ fails. But we are assuming Q_σ, and that implies $Q_{\sigma\hat{\ }n_0}$, so $R_{\sigma\hat{\ }n_0}$ fails. This means that there is an $m < \text{length}(\sigma\hat{\ }n_0) = \text{length}(\sigma) + 1$ such that

$$\mu \left(\bigcup_{n \in [m, N(\sigma\hat{\ }n_0)]} A_n \setminus \bigcup_{n \in [m, (\sigma\hat{\ }n_0)_m]} A_n \right) > (\lambda - \lambda')/2^{m+1}.$$

But our assumption of R_σ implies that this does not hold for $m < \text{length}(\sigma)$, since $N(\sigma\hat{\ }n_0) = n_1 \leq N(\sigma)$. So the only possibility is that it holds for $m = \bar{m} = \text{length}(\sigma)$;

in other words, we have

$$\mu \left(\bigcup_{n \in [\bar{m}, n_1]} A_n \setminus \bigcup_{n \in [\bar{m}, n_0]} A_n \right) > (\lambda - \lambda')/2^{\bar{m}+1}.$$

Now repeat this argument for $i = 2, 3, \ldots, \lceil 2^{\bar{m}+1}/(\lambda - \lambda') \rceil$. If the conclusion fails each time, then for each i we have

$$\mu \left(\bigcup_{n \in [\bar{m}, n_i]} A_n \setminus \bigcup_{n \in [\bar{m}, n_{i-1}]} A_n \right) > (\lambda - \lambda')/2^{\bar{m}+1}.$$

This implies $\mu(\bigcup_{n \in [\bar{m}, N(\sigma)]} A_n) > 1$, a contradiction.   □

The argument just employed is reminiscent of "energy incrementation" arguments used by Green and Tao [9, 25, 26], wherein the success of an algorithm is guaranteed by the fact that a bounded nondecreasing sequence can increase by more than a fixed $\varepsilon > 0$ only finitely many times. What is novel here is that we employ a transfinite iteration of that strategy.

Notice that the value of M' in the theorem depends on the values of λ, λ', and the *functional M*. It is therefore somewhat difficult to make sense of the question as to whether the bound computed in the proof is, in some sense, asymptotically sharp. Given M, λ, λ', one *can* effectively determine whether or not a putative value of M' works; so given any bound, one can also compute the least value of M' that satisfies the conclusion. So at issue is not whether we can compute the precise bound, but, rather, come up with a perspicuous characterization of the rate of growth.

One can easily use recursion along fairly simple trees to define functions that grow astronomically fast. Nonetheless, there are some things we can say about the complexity of M' in terms of M. It is well known that Gödel's system T of primitive recursive functionals of finite type can be stratified into levels T_n. At the bottom level, T_0, primitive recursion is restricted in such a way that the only *functions* from natural numbers to natural numbers that are definable in the system are primitive recursive. The functionals of T_0 are said to be *primitive recursive functionals in the sense of Kleene*, in contrast to the functionals of T, which are are said to be *primitive recursive functionals in the sense of Gödel* (see [13, 2, 15]). The results of Howard [12] show the following:

Theorem 2.4 *In the previous theorem, if M is definable in Gödel's T_n for some $n \geq 0$, then, as a function of λ and λ', M' is definable in T_{n+1}.*

(See also [22, Section 10], which relates Howard's results explicitly to the fragments T_n.) Theorem 2.4 implies that if M is a primitive recursive functional in the sense of Kleene, then M' is in T_1 (which is to say, roughly Ackermannian). The results of Kreuzer [21] provide even more information:

Theorem 2.5 *In Theorem 2.2, if M is definable in the calculus $\mathsf{G}_\infty\mathsf{A}^\omega$ (see, for example, [14, Section 3]), then M' is primitive recursive.*

It would be interesting to know whether these results can be improved. Alternatively, one can consider Theorem 2.2 for particular functionals $M(F)$. One can show, for example, that with $M(F) = F(0) + n$, the smallest value of M' that works is roughly $n/(\lambda - \lambda')$. Using the algorithm given in the proof of Theorem 2.2 yields the bound $M' = n \cdot \lceil 2/(\lambda - \lambda') \rceil$, but this can be improved to $n \cdot \lceil 1/(\lambda - \lambda') \rceil$ by tinkering with the values $(\lambda - \lambda')/2^{m+1}$ in the right hand side of condition R_σ. An explicit construction gives a lower bound of $n \cdot \left(\left\lceil \frac{1-\lambda'}{\lambda-\lambda'} \right\rceil - 1 \right)$.

However, even for simple functionals like $M(F) = F(F(0)) + n$, the combinatorial details quickly become knotty. In this particular case our algorithm gives an $M' = m_{\lceil 2/(\lambda-\lambda')\rceil}$, where $m_0 = n$ and $m_{i+1} = n \cdot \lceil 2^{m_i+1}/(\lambda - \lambda') \rceil$. This is an iterated exponential in n, where the depth of the stack depends on $\lambda - \lambda'$; but we do not know whether such a rate of growth is necessary.

3 Metastable convergence theorems

We can now prove our metastable version of Egorov's theorem.

Theorem 3.1 *For every $\varepsilon > 0$, $\lambda > \lambda' > 0$, and functional $M_1(F)$, there is a functional $M_2(F)$ with the following property: for any probability space $\mathcal{X} = (X, \mathcal{B}, \mu)$ and sequence (f_n) of measurable functions, if $M_1(F)$ is a λ'-uniform bound on the ε-metastable pointwise convergence of (f_n), then $M_2(F)$ is a bound on the λ-uniform ε-metastable convergence of (f_n). In other words, if for every F_1*

$$\mu(\{x \ : \ \exists m \leq M_1(F_1) \, \forall n, n' \in [m, F_1(m)] \, |f_n(x) - f_{n'}(x)| < \varepsilon\}) > 1 - \lambda',$$

then for every F_2 there is an $m \leq M_2(F_2)$ such that

$$\mu(\{x \ : \ \forall n, n' \in [m, F_2(m)] \, |f_n(x) - f_{n'}(x)| < \varepsilon\}) > 1 - \lambda.$$

Proof Fix $\varepsilon > 0$, $\lambda > \lambda' > 0$, and M_1. Given F_2, define

$$M(F) = M_1\left(\lambda m. \max_{n \in [m, F(m)]} F_2(n)\right),$$

and let $M_2(F_2)$ be the value M' given by Theorem 2.2. Let

$$A_n = \left\{x \; : \; \exists k, k' \in [n, F_2(n)] \, |f_k(x) - f_{k'}(x)| \geq \varepsilon\right\}.$$

We wish to show $\mu(A_n) < \lambda$ for some $m \leq M_2(F_2)$. By the definition of $M_2(F_2)$, it is enough to show that for every function $F(m)$,

$$\mu\left(\bigcap_{m \leq M(F)} \bigcup_{n \in [m, F(m)]} A_n\right) < \lambda'.$$

For each $m \leq M(F)$, we have

$$\bigcup_{n \in [m, F(m)]} A_n = \bigcup_{n \in [m, F(m)]} \bigcup_{k, k' \in [n, F_2(n)]} \{x \; : \; |f_k(x) - f_{k'}(x)| \geq \varepsilon\}$$

$$\subseteq \bigcup_{k, k' \in [m, \max_{n \in [m, F(m)]} F_2(n)]} \{x \; : \; |f_k(x) - f_{k'}(x)| \geq \varepsilon\}$$

Taking $F_1(m) = \max_{n \in [m, F(m)]} F_2(n)$ in the hypothesis of the theorem gives the desired conclusion. □

This straightforwardly yields our quantitative version of the dominated convergence theorem.

Theorem 3.2 *For every $\varepsilon > 0$, $\lambda > \lambda' > 0$, and $M_1(F)$, there is an $M_2(F)$ such that, for any probability space $\mathcal{X}$ and sequence (f_n) of nonnegative measurable functions dominated by the constant function 1, if $M_1(F)$ is a λ'-uniform bound on the ε-metastable pointwise convergence of (f_n), then $M_2(F)$ is a bound on the $(\varepsilon + \lambda)$-metastable convergence of $(\int f_n)$. In other words, if for every F*

$$\mu(\{x \; : \; \exists m \leq M_1(F) \, \forall n, n' \in [m, F(m)] \, |f_n(x) - f_{n'}(x)| < \varepsilon\}) > 1 - \lambda',$$

then for every F there is an $m \leq M_2(F)$ such that

$$\forall n, n' \in [m, F(m)] \; \left|\int f_n - \int f_{n'}\right| < \varepsilon + \lambda.$$

Proof From the hypotheses, Theorem 3.1 yields an $M_2(F)$ that is a bound on the λ-uniform ε-metastable convergence of (f_n). Thus, for all F, there is $m \leq M_2(F)$ such that

$$\mu\left(\{x \mid \forall n, n' \in [m, F(m)] \, |f_n(x) - f_{n'}(x)| < \varepsilon\}\right) > 1 - \lambda.$$

Call the set just indicated A. From our choice of λ and the definition of A, it follows that for all $n, n' \in [m, F(m)]$,

$$
\begin{aligned}
\left| \int f_n - \int f_{n'} \right| &\leq \int |f_n - f_{n'}| \\
&= \int_A |f_n - f_{n'}| + \int_{X \setminus A} |f_n - f_{n'}| \\
&< \varepsilon + \lambda.
\end{aligned}
$$

That is, $M_2(F)$ provides a bound on the $(\varepsilon + \lambda)$-metastable convergence of $(\int f_n)$ as desired. $\square$

Theorem 3.2 strengthens Tao's Theorem A.2 [26] in three ways. First, we formulate convergence in terms of the Cauchy criterion, rather than referring to a fixed limit, as Tao does. This is more natural in the context of metastability, and our result implies Tao's, since one can always consider a sequence $f_0, f, f_1, f, f_2, f, \ldots$ in which a fixed limit f has been interleaved. Second, Tao used the stronger hypothesis that $M_1(F)$ provides a bound that works almost everywhere, rather than outside a set of measure at most λ'. Finally, and most importantly, our proof of Theorem 2.2 provides an explicit description of the bound, $M_2(F)$.

Tao also stated his theorem for the convergence of nets indexed by the directed set $\mathbb{N} \times \mathbb{N}$, as was needed in his application. But as he himself noted, the extension to arbitrary countable nets is straightforward. Given any countable net $(f_i)_{i \in I}$, one can define an increasing cofinal sequence $(a_i)_{i \in \mathbb{N}}$ of elements of the directed set I. To adapt Theorem 2.2, for example, suppose we are given a net $(A_i)_{i \in I}$ of measurable subsets of a probability space $\mathcal{X}$ with the property that for every function F,

$$
\mu \left(\bigcap_{m \leq M(F)} \bigcup_{i \in [a_m, a_{F(m)}]} A_i \right) < \lambda',
$$

where the notation $[a, b]$ denotes $\{ i \ : \ a \leq i \leq b \}$. Define the sequence $(A'_n)_{n \in \mathbb{N}}$ by $A'_n = \bigcup_{i \in [a_n, a_{n+1}]} A_i$. Then (A'_n) satisfies the requirements of Theorem 2.2, and hence there is an $i \leq a_{M'}$ such that $\mu(A_i) < \lambda$.

Notice that the expression $\int |f_n - f_{n'}|$ in the proof of Theorem 3.2 is the L^1 norm of $f_n - f_{n'}$. In fact, the same argument shows the following:

Theorem 3.3 *For every $\varepsilon > 0$, $\lambda > \lambda' > 0$, and $M_1(F)$, there is an $M_2(F)$ such that, for any probability space $\mathcal{X}$ and sequence (f_n) of nonnegative measurable functions*

dominated by the constant function 1, if $M_1(F)$ is a λ'-uniform bound on the ε-metastable pointwise convergence of (f_n), then for every F_2 there is an $m \leq M_2(F_2)$ such that for every $p \geq 1$,

$$\forall n, n' \in [m, F(m)] \ \|f_n - f_{n'}\|_p^p < \varepsilon^p + \lambda.$$

We have considered a metastable version of the dominated convergence theorem where $\mathcal{X}$ is a probability space and the sequence (f_n) is uniformly dominated by the constant function 1. The dominated convergence theorem itself is usually stated more generally where $\mathcal{X}$ is an arbitrary measure space, and the sequence (f_n) is dominated by an arbitrary integrable function g. The general case can be reduced to the one we have considered, taking into account that given an integrable function g and any δ_1, δ_2 greater than 0, there is a set A with finite measure such that $\int_{X \setminus A} g < \delta_1$, and a K sufficiently large so that $\int_A (g - \min(g, K)) < \delta_2$. The bound M_2 in the conclusion, however, now depends on bounds on K and the size of A, for certain δ_1 and δ_2 depending on ε.

4   A new mode of convergence

Recall that a sequence (f_n) of measurable functions converges pointwise a.e. if for almost every x,

(AE) $\forall \varepsilon > 0 \ \exists m \ \forall n, n' \geq m \ |f_n(x) - f_{n'}(x)| < \varepsilon,$

and we noted in Section 1 it converges almost uniformly if

(AU) $\forall \lambda > 0, \varepsilon > 0 \ \exists m \ \mu(\{x \ : \ \exists n, n' \geq m \ |f_n(x) - f_{n'}(x)| \geq \varepsilon\}) < \lambda.$

Each of these has an equivalent expression in terms of metastable convergence. Our formulation of Egorov's theorem provides yet another mode of convergence, which we will call *almost uniform metastable pointwise convergence*:

(AUM) $\forall \lambda > 0, \varepsilon > 0, F \ \exists M$

$\mu(\{x \ : \ \forall m \leq M \ \exists n, n' \in [m, F(m)] \ |f_n(x) - f_{n'}(x)| \geq \varepsilon\}) < \lambda.$

In other words, M, as a function of F, provides a bound on the ε-metastable convergence of the sequences $(f_n(x))$ that is uniform in x, and valid outside a set of measure at most λ.

Recall that if $\mathcal{X}$ is a probability space, or if the sequence (f_n) is dominated by an L^p function, then a.e. convergence and almost uniform convergence coincide. More generally, we have the following relationships between these three modes of convergence:

Proposition 4.1 *Let (f_n) be a sequence of measurable functions on a measure space $\mathcal{X} = (X, \mathcal{B}, \mu)$.*

(1) $\text{AU} \to \text{AUM} \to \text{AE}$. *(Hence, if $\mathcal{X}$ is a probability space or the sequence (f_n) is dominated, the three notions coincide.)*

(2) *If $\mu(\{x : |f_n(x) - f_{n'}(x)| \geq \varepsilon\}) < \infty$ for all $\varepsilon > 0$, n, and n', then AE implies AUM. (In particular, the conclusion holds if for some $p \geq 1$, $f_n \in L^p$ for every n.)*

(3) *In general, the implications in (1) do not reverse.*

Proof For (1), note that AU is equivalent to its metastable version, AU*, which clearly implies AUM. Similarly, AUM implies, in particular, that for almost every x the sequence $(f_n(x))$ is metastably convergent, and hence convergent.

For (2), prove the contrapositive. Suppose AUM fails. Then there are ε, λ, F such that for all M,

$$\mu\left(\bigcap_{m \leq M} \bigcup_{n,n' \in [m, F(m)]} \{x : |f_n(x) - f_{n'}(x)| \geq \varepsilon\}\right) \geq \lambda.$$

By the assumption that each $\{x : |f_n(x) - f_{n'}(x)| \geq \varepsilon\}$ has finite measure, we can take the limit as $M \to \infty$ to get

$$\mu\left(\bigcap_{m} \bigcup_{n \in [m, F(m)]} \{x : |f_n(x) - f_{n'}(x)| \geq \varepsilon\}\right) \geq \lambda.$$

Further, removing F gives,

$$\mu\left(\bigcap_{m} \bigcup_{n \geq m} \{x : |f_n(x) - f_{n'}(x)| \geq \varepsilon\}\right) \geq \lambda.$$

Hence, (f_n) is not a.e. Cauchy.

For (3), $f_n = \chi_{[n,n+1]}$ converges AUM by part (2), but it is easily shown that f_n does not converge AU. Finally, $g_n = (-1)^n \chi_{[n,\infty)}$ converges only AE. $\qquad\square$

There is also a non-Cauchy version of AUM, which refers to a limit function f:

(AUM′) $\exists f \; \forall \lambda > 0, \varepsilon > 0, F \; \exists M$

$$\mu(\{x : \forall m \leq M \; \exists n \in [m, F(m)] \; |f_n(x) - f(x)| \geq \varepsilon\}) < \lambda.$$

It is easy to see that AUM′ implies AUM, but the converse need not hold; for example, $h_n = \chi_{[n,\infty)}$ converges AUM, but not AUM′. Moreover, the analogue of Proposition 4.1 holds when AUM is replaced by AUM′. Thus we have the implications

$$\text{AU} \rightarrow \text{AUM}' \rightarrow \text{AUM} \rightarrow \text{AE},$$

none of which can be reversed in general.

5   Final comments

As noted in Section 2, it would be interesting to know the extent to which the bounds we obtain are sharp. For example, can one show that there are functionals M that are primitive recursive in the sense of Kleene for which the M' in Theorem 2.2 is not primitive recursive?

When Tao [26] presented his quantitative version of the dominated convergence theorem, he observed that the bound M' can be computed in principle.

> In practice, though, it seems remarkably hard to do; the proof of the Lebesgue dominated convergence theorem, if inspected carefully, relies implicitly on the infinite pigeonhole principle, which is notoriously hard to finitize.

He went on to note that since the Lebesgue dominated convergence theorem is equivalent, in the sense of reverse mathematics, to the arithmetic comprehension axiom (ACA) (see Yu [28]), the dependence of M' on the parameters is likely to be "fantastically poor." The dependence we have obtained is, indeed, rather poor, but it is at least explicit and comprehensible.

In fact, the axiomatic strength of the dominated convergence theorem is sensitive to the way in which it is formulated. Elsewhere [1] we have shown that the formulation of the dominated convergence theorem that corresponds to Tao's quantitative version is strictly weaker than (ACA). It is possible, however, that the quantitative version, which quantifies over continuous functionals, is axiomatically stronger than the original. We suspect that each of Theorems 2.2, 3.1, and 3.2 is equivalent to (ACA). Hirst [10] has shown that the kind of transfinite induction used in the proof of Theorem 2.2 is equivalent to (ACA), so at issue is whether the full strength of this principle is needed. The situation is reminiscent of Gaspar and Kohlenbach [7], which provides a sense in which a quantitative version of the infinitary pigeonhole principle is axiomatically stronger than the non-quantitative version.

The results here can be viewed as instances of "proof mining," which aims to extract quantitative and computationally meaningful information from nonconstructive results; see [14] and [3, 15, 16, 17]. Specifically, the instance of arithmetic comprehension used in the implication from (3) to (1) of Proposition 2.1 can be cast as a choice principle, and Spector [24] has shown how that can be eliminated in favor of bar recursion; see also [14, Section 11.3]. Paulo Oliva has, moreover, shown [23] that for a fixed space $\mathcal{X}$ and sequence of sets (A_n), one can straightforwardly compute a witness to the implication using a particular form of bar recursion (the "explicitly controlled product of selection functions" of [5]; see also Definition 18 of [6]). One can then obtain a bound that is uniform in $\mathcal{X}$ and (A_n) using hereditary majorizability in the sense of either Howard [11] or Bezem [4].

References

[1] **J Avigad**, **E Dean**, **J Rute**, *Algorithmic randomness, reverse mathematics, and the dominated convergence theorem*, submitted.

[2] **J Avigad**, **S Feferman**, *Gödel's functional ("Dialectica") interpretation*, in: *Handbook of Proof Theory* (ed. S Buss), North-Holland, Amsterdam, 1998, 337–405; doi: 10.1016/S0049-237X(98)80020-7.

[3] **J Avigad**, **P Gerhardy**, **H Towsner**, *Local stability of ergodic averages*, Transactions of the American Mathematical Society 362 (2010), 261–288; doi: 10.1090/S0002-9947-09-04814-4.

[4] **M Bezem**, *Strongly majorizable functionals of finite type: a model for bar recursion containing discontinuous functionals*, Journal of Symbolic Logic 50 (1985), 652–660; doi: 10.2307/2274319.

[5] **M Escardó**, **P Oliva**, *Computational interpretations of analysis via products of selection functions*, in: *Programs, proofs, processes* (eds. F Ferreira, B Löwe, E Mayordomo, L Gomes) Springer, Berlin, 2010, 141–150.

[6] **M Escardó**, **P Oliva**, *Sequential games and optimal strategies*, Proceedings of the Royal Society, Series A 467 (2011), 1519–1545; doi: 10.1098/rspa.2010.0471.

[7] **J Gaspar**, **U Kohlenbach**, *On Tao's "finitary" infinite pigeonhole principle*, Journal of Symbolic Logic 75 (2010), 355–371; doi: 10.2178/jsl/1264433926.

[8] **K Gödel**, *Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes*, Dialectica 12 (1958), 280–287; reprinted with English translation in: *Kurt Gödel: Collected Works* (eds. S Feferman et al.), volume 2, Oxford University Press, New York, 1990, 241–251.

[9] **B Green**, **T Tao**, *The primes contain arbitrarily long arithmetic progressions*, Annals of Mathematics 167 (2008), 481–547; doi: 10.4007/annals.2008.167.481.

[10] **J L Hirst**, *Ordinal inequalities, transfinite induction, and reverse mathematics*, The Journal of Symbolic Logic 64 (1999), 769–774; doi: 10.2307/2586499.

[11] **W A Howard**, *Hereditarily majorizable functionals of finite type*, in: *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis* (ed. A S Troelstra), Springer, Berlin, 1973, 454–461.

[12] **W A Howard**, *Ordinal analysis of simple cases of bar recursion*, Journal of Symbolic Logic 46 (1981), 17–30; doi: 10.2307/2273252.

[13] **S C Kleene**, *Recursive functionals and quantifiers of finite types, I*, Transactions of the American Mathematical Society 91 (1959), 1–52.

[14] **U Kohlenbach**, *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*, Springer Monographs in Mathematics, Springer, Berlin, 2008.

[15] **U Kohlenbach**, *On quantitative versions of theorems due to F. E. Browder and R. Wittmann*, Advances in Mathematics, 226 (2011), 2764–2795; doi: 10.1016/j.aim.2010.10.002.

[16] **U Kohlenbach**, *A uniform quantitative form of sequential weak compactness and Baillon's nonlinear ergodic theorem*, Communications in Contemporary Mathematics, to appear.

[17] **U Kohlenbach**, **L Leuştean**, *A quantitative mean ergodic theorem for uniformly convex Banach spaces*, Ergodic Theory and Dynamical Systems 29 (2009), 1907–1915; doi: 10.1017/S0143385708001004.

[18] **G Kreisel**, *On the interpretation of non-finitist proofs, part I*, Journal of Symbolic Logic 16 (1951), 241–267.

[19] **G Kreisel**, *On the interpretation of non-finitist proofs, part II: Interpretation of number theory, applications*, Journal of Symbolic Logic 17 (1952), 43–58.

[20] **G Kreisel**, *Interpretation of analysis by means of constructive functionals of finite type*, in *Constructivity in Mathematics: Proceedings of the Colloquium (Amsterdam, 1957)* (ed. A. Heyting), Studies in Logic and the Foundations of Mathematics, North-Holland Publishing Co., Amsterdam, 1959, 101–128.

[21] **A P Kreuzer**, *Primitive recursion and the chain antichain principle*, to appear in the *Notre Dame Journal of Formal Logic*.

[22] **A P Kreuzer**, **U Kohlenbach**, *Term extraction and Ramsey's theorem for pairs*, preprint.

[23] **P Oliva**, *Alternative construction for metastable convergence*, personal communication, September 1, 2011.

[24] **C Spector**, *Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics*, in: *Recursive Function Theory: Proc. Symposia in Pure Mathematics* (ed. J C E Dekker), American Mathematical Society, Providence, Rhode Island, 1962, 1–27.

[25] **T Tao**, *A quantitative ergodic theory proof of Szemerédi's theorem*, Electronic Journal of Combinatorics 13 (2006), Research Paper 99, 49 pp.

[26] **T Tao**, *Norm convergence of multiple ergodic averages for commuting transformations*, Ergodic Theory and Dynamical Systems 28 (2008), 657–688; doi: 10.1017/S0143385708000011.

[27] **T Tao**, *Structure and Randomness: pages from year one of a mathematical blog*, American Mathematical Society, Rhode Island, 2008.

[28] **X Yu**, *Lebesgue convergence theorems and reverse mathematics*, Mathematical Logic Quarterly 40 (1994), 1–13.

Department of Philosophy, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Department of Philosophy, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA

avigad@cmu.edu, edean@andrew.cmu.edu, jrute@andrew.cmu.edu

http://www.andrew.cmu.edu/user/avigad,
http://www.andrew.cmu.edu/user/edean/, http://www.math.cmu.edu/~jrute/