

Proof Mining in Topological Dynamics

Philipp Gerhardy

Abstract A famous theorem by van der Waerden states the following: Given any finite coloring of the integers, one color contains arbitrarily long arithmetic progressions. Equivalently, for every q, k , there is an $N = N(q, k)$ such that for every q -coloring of an interval of length N one color contains a progression of length k . An obvious question is what is the growth rate of $N(q, k)$. Some proofs, like van der Waerden's combinatorial argument, answer this question directly, while the topological proof by Furstenberg and Weiss does not. We present an analysis of (Girard's variant of) Furstenberg and Weiss's proof based on monotone functional interpretation, both yielding bounds and providing a general illustration of proof mining in topological dynamics. The bounds do not improve previous results by Girard, but only—as is also revealed by the analysis—because the combinatorial proof and the topological dynamics proof in principle are identical.

1 Introduction

“Proof mining” is the activity of extracting additional information from proofs in mathematics and computer science. The two main types of additional information that may be extracted are quantitative and qualitative information. An example of the former is extracting a rate of convergence from a proof that a certain iteration sequence in a compact metric space converges. An example of the latter is establishing that the convergence is uniform in the starting point of the iteration or that the result not only holds for compact metric spaces, but already for bounded ones. Naturally, when the proof to be analyzed is constructive, we expect quantitative information—realizers or bounds—to be present explicitly in the proof, and qualitative strengthenings of the theorem may follow trivially from the exact structure of the realizers. Thus, the main focus in proof mining is on extracting such information even from ineffective proofs, that is, proofs using full classical logic.

Received December 4, 2007; accepted April 23, 2008; printed September 25, 2008
2000 Mathematics Subject Classification: Primary, 03, 03F

Keywords: mathematical logic, proof mining, van der Waerden, combinatorics, topological dynamics

© 2008 by University of Notre Dame 10.1215/00294527-2008-022

To achieve these aims in proof mining one focuses on two things: on the one hand, developing general metatheorems that classify theorems and proofs from which additional information can be extracted and, on the other hand, carrying out case studies by analyzing actual proofs. These activities are complimentary. New, stronger metatheorems may open up new classes of proofs to analysis, and results that were obtained ad hoc in case studies may inspire new proof mining techniques and lead to new, stronger metatheorems.

The main tool in proof mining are so-called proof interpretations. The idea of proof interpretations is to transform a given proof into an enriched proof from which the desired additional information can be read off. There are various kinds of proof interpretations spanning from, for example, cut elimination, which significantly changes the structure of the proof during transformation, and functional interpretations, which preserve the structure but at the cost of introducing functionals of higher types into the interpretation. We will discuss this in more detail in the context of a concrete proof analysis carried out in Section 5.

In this paper, we discuss and analyze two topological proofs of van der Waerden's theorem, more precisely, proofs of the Multiple Birkhoff Recurrence theorem, from which van der Waerden's theorem follows by an easy argument. The topological proofs are by Furstenberg and Weiss [2] and Girard [5], where the latter proof is a slightly modified variant of the former. The original combinatorial proof of van der Waerden's theorem (by van der Waerden himself [14]), does provide some bounds, that is, a number $N = N(q, k)$ so that for every q -coloring of $[-N, N]$ some color contains an arithmetic progression of length k . However, the bounds are of Ackermann growth and thus of poor complexity, while the best known lower bounds are elementary. The proof by Furstenberg and Weiss does not directly yield such a bound and the hope is that an analysis of this "different" proof allows to extract different, better bounds for van der Waerden's theorem.

In [5], Girard modified the proof by Furstenberg and Weiss to facilitate the extraction of bounds for van der Waerden's theorem from the topological proof. Girard's analysis is based on cut elimination and he essentially obtains the same bounds as van der Waerden, but the analysis does not treat the full proof, only its specialization to the setting relevant to van der Waerden's theorem. The skeptical proof analyst might suspect that the poor bounds are at least influenced by the well-known poor complexity behavior of cut elimination and by the fact that cut elimination significantly changes the structure of a given proof with cuts.

In this paper, we present a different analysis of Girard's proof, based on so-called monotone functional interpretation. This method is faithful to the structure of the proof, but again yields the same bounds as van der Waerden. The analysis also reveals that this is not due to the extraction of bounds itself, but rather because the proofs are "identical": except for the different settings of finite combinatorics and topological dynamics all the steps in the proofs by van der Waerden and Girard's variant of the proof of Furstenberg and Weiss are exactly the same. We also suggest that Girard's modified proof can be motivated by analyzing the Furstenberg and Weiss proof, that is, by retaining a bit more information in the lemmas of the original proof by Furstenberg and Weiss.

Combined, the proof analyses provide good examples of carrying out proof mining in topological dynamics. Furthermore, the author plans to explore whether the analogies between combinatorics and topological dynamics can be exploited

for other purposes, for example, to derive a new topological proof of the Multiple Birkhoff Recurrence theorem from Shelah's improved combinatorial proof of van der Waerden's theorem (see [11]). The author hopes to pursue this in a future paper.

We now give an overview of the paper. In Section 2, we review van der Waerden's theorem and the connections to multiple recurrence in topological dynamics. In Section 3, we recall the topological proof of Multiple Birkhoff Recurrence by Furstenberg and Weiss, and Section 4 discusses how Girard's refined proof can be motivated by "extracting" (or rather, retaining) some additional information in some of the lemmas in the Furstenberg and Weiss proof. Finally, Sections 5 and 6 describe the analysis of Girard's proof in detail and discuss the relationship of the extracted bounds to van der Waerden's original combinatorial proof.

2 Van der Waerden's Theorem and Topological Dynamics

Let us start by formally stating van der Waerden's theorem.

Definition 2.1 An arithmetic progression of length k is a sequence of the form $a, a + b, a + 2b, \dots, a + (k - 1)b$ for integers $a, b > 0$.

Van der Waerden's Theorem Let $\mathbb{Z} = C_1 \cup \dots \cup C_q$ be a (finite) coloring of the integers. Then one of the colors contains arbitrarily long arithmetic progressions.

This is the common formulation of van der Waerden's theorem, but it does not suggest much constructive information, except perhaps which color contains the progressions. Thus, instead of establishing the existence of arbitrarily long arithmetic progressions, we may ask how large an interval we need to consider to ensure that for any q -coloring of that interval some color contains a progression of length k , where the size depends on the number of colors q and the length k . This is the version that was actually proven in van der Waerden's original paper [14]. This (equivalent) finite version can be stated as follows.

Van der Waerden's Theorem (finite version) For any $q, k > 0$, there exists an $N = N(q, k) > 0$ such that for any q -coloring $C_1 \cup \dots \cup C_q$ of $[-N, N] \subseteq \mathbb{Z}$, one of the colors contains an arithmetic progression of length k .

With this formulation the obvious question is to determine either exact values of the function $N(q, k)$ or at least information on the growth rate of the function. Except for the simplest cases, exact values are difficult to determine. The study of upper and lower bounds for van der Waerden's theorem, though those are simpler to find than exact values, is a dynamic field of research and keeping track of the newest exact values and bounds is beyond the scope of this paper. In [9], the following currently known best lower bound is presented.

Theorem 2.2 For all $q \geq 2$, $N(q, k) > \frac{q^k}{eq^k}(1 + o(1))$.

The best upper bound known is due to Gowers [6].

Theorem 2.3 Let $f(q, k) = q^{2^{k+9}}$, then $N(q, k) \leq 2^{2^{f(q,k)}}$.

The huge gap between upper and lower bounds suggests that there still is potential for future research in this field. In van der Waerden's original paper [14], the following

upper bounds were obtained:

$$\begin{aligned} N(q, 2) &= q + 1, \\ N(q, k + 1) &= n_q, \quad \text{where} \\ n_0 &= 1, \\ n_{l+1} &= N(q^{n_l}, k) + \lceil \frac{N(q^{n_l} - 1, k)}{k - 1} \rceil + n_l - 1. \end{aligned}$$

One may simplify the last clause to $n_{l+1} = 2N(q^{n_l}, k)$. As these bounds are defined by double recursion, they essentially are of Ackermann growth.

In [11], Shelah presents a different combinatorial proof of van der Waerden's theorem (in fact, of the Hales-Jewett theorem, which is a generalization of van der Waerden's theorem) and obtains the following (at that time, first) primitive recursive upper bounds:

$$\begin{aligned} f(1, q) &= q + 1, \\ f(l + 1, q) &= q^{f(l, q)^{2l}} + 1, \\ HJ(1, q) &= 1, \\ HJ(n + 1, q) &= HJ(n, q) \cdot f(HJ(n, q), q^{(n+1)^{HJ(n, q)}}), \\ N(q, k) &= (k - 1)HJ(k, q), \end{aligned}$$

where f is an auxiliary function and $HJ(n, q)$ denotes a bound for the Hales-Jewett theorem. In Hales-Jewett the number n denotes the size of an alphabet, and this size roughly compares to the length of an arithmetic progression in van der Waerden's theorem. Note that Shelah's bound for $HJ(n, q)$ only depends on values of HJ where n is strictly smaller and q is the same. Thus no double recursion is necessary, whereas in van der Waerden's bounds decreasing the other parameter comes at the price of drastically increasing q , the number of colors.

Gowers's bounds are obtained from his proof of Szemerédi's theorem (first proved in [12]), which is a density version of van der Waerden's theorem: Instead of finite colorings of the integers, one looks at subsets of the integers of positive density (in some measure-theoretic sense) and finds arithmetic progressions in such sets. Gowers uses a score of analytic techniques to obtain his result, but a more detailed summary is beyond the scope of this paper.

Surprisingly, van der Waerden's purely combinatorial theorem has a counterpart in topological dynamics, where the notion of arithmetic progressions is represented by some suitable notion of recurrence. This result was obtained by Furstenberg and Weiss [2].¹

Definition 2.4 A system (X, G) is a dynamical system if (X, d) is a compact metric space and G is a group of commuting homeomorphisms of X . If G is generated by a single homeomorphism T , we may write (X, T) for the corresponding dynamical system.

Remark There exist several equivalent ways of formulating compactness. We generally consider a space to be compact if it is totally bounded and complete.

Definition 2.5 Let (X, d) be a compact metric space and T a homeomorphism of X . A point $x \in X$ is a recurrent point for T if, for every $\varepsilon > 0$, there exists an $n > 0$ such that $d(T^n x, x) \leq \varepsilon$.

The Multiple Birkhoff Recurrence theorem states that for a tuple T_1, \dots, T_l of commuting homeomorphism, we can find a point that is recurrent simultaneously for all T_i s.

Multiple Birkhoff Recurrence Theorem ([2]) Let (X, d) be a compact metric space and T_1, \dots, T_l commuting homeomorphisms of X . Then there exists a point $x \in X$ such that for every $\varepsilon > 0$ there is an $n > 0$ satisfying $d(T_i^n x, x) \leq \varepsilon$ simultaneously for $i = 1, \dots, l$.

It is easy to see that van der Waerden’s theorem follows from the Multiple Birkhoff Recurrence theorem: A q -coloring of the integers can be represented by a function $x : \mathbb{Z} \rightarrow \{1, \dots, q\}$. Define $(Tx)(n) = x(n + 1)$, let $X = \{\overline{T^n x} \mid n \in \mathbb{Z}\}$, and define a metric $d(x, y) = 2^{-k}$ for the biggest k such that $x(l) = y(l)$ and $x(-l) = y(-l)$ for all $l \leq k$. Then (X, d) is a compact metric space and T a homeomorphism of X . Now let $T_i = T^i$, that is, T iterated i times; obviously, the homeomorphisms T_i commute. Applying the Multiple Birkhoff Recurrence theorem to those T_i , we find a point that is recurrent for all T_1, \dots, T_l .

Observe that if $d(x, y) < 1$ in our metric space, we have that $x(0) = y(0)$ for the corresponding colorings. Thus we merely need to find a point $z \in X$ such that $d(T_i^n z, z) < 1$ for some $n > 0$ and all $i = 1, \dots, l$. The existence of such a point z can be derived from the existence of a multiply recurrent point. Either the point z directly is a translate of x , that is, $z = T^m x$, or, if z lies in the boundary of X , then there is a $T^m x$ close enough to z to ensure the result (by the continuity of the T_i). This then yields the existence of some $m, n > 0$ such that $x(m) = x(m+n) = \dots = x(m+ln)$, that is, an arithmetic progression of length $l + 1$. For more details on this correspondence, see, for example, [2], [1], and [10].

It is worth noting already here that to prove van der Waerden’s theorem from a multiple recurrence result in topological dynamics, it is not strictly necessary to obtain a multiply recurrent point. This is also mentioned in the paper by Furstenberg and Weiss ([2], p. 66). It suffices to find for each $\varepsilon > 0$ some point z such that all $T_i^n z$ are ε -close to z , instead of one point that works for all $\varepsilon > 0$.

3 The Proof by Furstenberg and Weiss

We first give an informal sketch of the proof (by Furstenberg and Weiss) of the Multiple Birkhoff Recurrence theorem. We find a recurrent point for commuting homeomorphisms T_1, \dots, T_l by induction on l . The case $l = 1$ is essentially the Poincaré recurrence theorem. For the induction step, assume we know how to find a recurrent point for l homeomorphisms. Given $l + 1$ homeomorphisms T_1, \dots, T_{l+1} , we form l new homeomorphisms $S_i = T_i T_{l+1}^{-1}$. Applying the induction hypothesis to the S_i , we can find, for each $\varepsilon > 0$, some points $x, y \in X$ and an $n \in \mathbb{N}$ such that $d(x, T_i^n y) < \varepsilon$ simultaneously for $i = 1, \dots, l + 1$. Using the concept of minimality for dynamical systems (to be explained below), this can be strengthened to the following: for every $x \in X$ there is a $y \in X$ and an $n \in \mathbb{N}$ such that $d(x, T_i^n y) < \varepsilon$ simultaneously. We then form a sequence of points z_n —choose a z_0 , for z_0 obtain the corresponding y and let this be z_1 , for z_1 obtain the corresponding z_2 and so on—satisfying $d(T^{n_j + \dots + n_{i+1}} z_j, z_i) < \varepsilon/2$ for any $i < j$. By compactness of X , there is a pair $i < j$ such that $d(z_i, z_j) < \varepsilon/2$ and the induction step is complete.

Most steps of the proof are rather straightforward (once stated in an appropriate form) and from a computational point of view, the greatest challenges are a number

of applications of compactness in the proof. Before we discuss these in detail, let us briefly recall each of the main lemmas in Furstenberg and Weiss's proof.

Lemma 3.1 *Let (X, d) be a compact metric space and let $T : X \rightarrow X$ be a self-mapping of X . Then for any $\varepsilon > 0$ there exists an $x \in X$ and an $n > 0$ such that $d(T^n x, x) < \varepsilon$.*

Proof Take any $x_0 \in X$ and consider the sequence $x_i := T^i x_0$ for $i = 0, 1, \dots$. By compactness, more precisely, by total boundedness, there is an m such that some x_i, x_j are $\varepsilon/2$ -close for $0 \leq i < j \leq m$. Hence, $d(T^{j-i} x_i, x_i) < \varepsilon$. \square

This lemma constitutes the induction base in the proof of the Multiple Birkhoff Recurrence theorem. The next three lemmas constitute the induction step. In the following, (X, d) will always be a compact space and T_i, S_i denote commuting homeomorphisms of X .

Lemma 3.2 *Assume that for any $\delta > 0$ and any T_1, \dots, T_l there exists a $z \in X$ and an $n > 0$ such that simultaneously $d(T_i^n z, z) < \delta$ for all $i = 0, \dots, l$. Then for any S_1, \dots, S_{l+1} and any $\varepsilon > 0$ there exist $x, y \in X$ and an $m > 0$ such that simultaneously $d(S_i^m x, y) < \varepsilon$ for all $i = 0, \dots, l + 1$.*

Proof From S_1, \dots, S_{l+1} form l new homeomorphisms by defining $T_i := S_i S_{l+1}^{-1}$ for $i = 1, \dots, l$. Applying the premise to those T_i and ε , we obtain a point $z \in X$ and an $n > 0$ such that simultaneously $d(S_i^n S_{l+1}^{-n} z, z) < \varepsilon$ for $i = 0, \dots, l$. Define $x = S_{l+1}^{-n} z$ and $y = z$, then $d(S_i^n x, y) < \varepsilon$ simultaneously for all $i = 0, \dots, l + 1$. \square

The next lemma uses the notions of minimality and homogeneity for dynamical systems. The next two lemmas are stated and proven for a general compact metric space and a *single* homeomorphism. Given (X, d) and T_1, \dots, T_l , the space we have in mind is the l -fold product of our original space (X, d) and the homeomorphism T is $T = T_1 \times \dots \times T_l$. A point x satisfying $d(x, T^n x) < \varepsilon$ for this product space and product homeomorphism is then a point in the original space satisfying $d(x, T_i^n x) < \varepsilon$ simultaneously for $i = 1, \dots, l$.

Definition 3.3 A dynamical system (X, G) is minimal if no proper subset of X is left invariant by all the transformations of G .

Definition 3.4 A closed (nonempty) subset $A \subseteq X$ is homogeneous with regard to a dynamical system (X, G) if there is a group G' of homeomorphisms commuting with G such that G' leaves A invariant and (A, G') is minimal.

Every dynamical system (X, T) has a minimal subsystem (using Zorn's Lemma), so, without loss of generality, we assume (X, T) is minimal. Then for the l -fold product of (X, T) (with the new $T = T_1 \times \dots \times T_l$) the diagonal subspace (i.e., the set of l -tuples (x, x, \dots, x)) is a homogeneous set where the group G is the group generated by (l -fold products) $(T_1 \times \dots \times T_1), \dots, (T_l \times \dots \times T_l)$. For the details of this argument see [2].

Lemma 3.5 *Let a dynamical system (X, T) be given. Let $A \subseteq X$ be closed and homogeneous (for a group G). If for any $\delta > 0$ there exist $u, v \in A$ and $n > 0$ such that $d(T^n u, v) < \delta$, then for every $x \in A$ there exists a $y \in A$ and an $m > 0$ such that $d(x, T^m y) < \varepsilon$.*

Proof Let G be the group commuting with T and let $\varepsilon > 0$ be given. By minimality (and compactness) of A there is a finite subset $G_0 \subseteq G$ such that $\min_{g \in G_0} d(gz, z') < \varepsilon/2$

for every $z, z' \in A$. Since the elements of G_0 are (uniformly) continuous, we can choose a $\delta > 0$ such that $d(z, z') < \delta$ implies $d(gz, gz') < \varepsilon/2$ for all $g \in G_0$.

Now, let $u, v \in A$ and $n > 0$ be given by the premise with $d(T^n u, v) < \delta$. Let $x \in A$ be given. Then by assumption $d(g_i v, x) < \varepsilon/2$ for some $g_i \in G_0$. Using the continuity of G_0 , we also have $d(T^n g_i u, g_i v) < \varepsilon/2$, so $d(T^n g_i u, x) < \varepsilon$. \square

This property of homogeneous subsets is used in the next lemma.

Lemma 3.6 *Let $A \subseteq X$ be a closed subset such that for all $\delta > 0$ and $u \in A$ there exists $v \in A$ and $n > 0$ with $d(T^n v, u) < \delta$. Then for every $\varepsilon > 0$ there exists $x \in A$ and $m > 0$ such that $d(T^m x, x) < \varepsilon$.*

Proof Let $\varepsilon > 0$ be given, take any $z_0 \in A$, and define $\varepsilon_1 = \varepsilon/4$. By the premise, there is $z_1 \in A$ and an $n_1 > 0$ such that $d(T^{n_1} z_1, z_0) < \varepsilon_1$. Next, choose $\varepsilon_2 > 0$ such that $\varepsilon_2 < \varepsilon_1/2$ and $d(z, z_1) < \varepsilon_2$ implies $d(T^{n_1} z, T^{n_1} z_1) < \varepsilon_1/2$ and thus also $d(T^{n_1} z, z_0) < \varepsilon/2$. This step uses the continuity of T . We then choose z_2, n_2 for $\varepsilon/2$. Continuing in this fashion, we construct a sequence z_i which satisfies $i < j \rightarrow d(T^{n_j + \dots + n_{i+1}} z_j, z_i) < \varepsilon/2$. By the total boundedness of X , there is a pair $i < j$ such that $d(z_i, z_j) < \varepsilon/2$ and thus for that j and $n = n_j + \dots + n_{i+1}$ we have $d(T^n z_j, z_j) < \varepsilon$. \square

In the final argument to get a recurrent point from points $z \in X$ obtained in the previous lemma, we consider $F(x) = \inf_{n \in \mathbb{N}} d(T^n x, x)$. By the previous lemma, $F(x)$ is not bounded from below (by some $\delta > 0$). Using homogeneity (of a set A as in Lemma 3.5) and the upper semicontinuity of F , one shows that $F(x)$ attains its infimum, namely, 0, at points of continuity for F (on the homogeneous set A), which then are recurrent points. For the full details of this argument see [2] and [1].

Most steps in the proof are direct appeals to continuity or the triangle inequality, so let us return to the three applications of compactness (in the induction step):

1 In Lemma 3.5, compactness is used to obtain a finite subcover from a given cover of the space X . Given a group G , an open neighborhood $V \subseteq X$, and a cover of X by neighborhoods $g^{-1}V$ for $g \in G$, we want to find a finite set of elements $\{g_1, \dots, g_k\} (= G_0) \subseteq G$ that covers X . To make the lemma computationally explicit, it is not enough to know that a finite cover exists, but we also need to know which elements comprise the finite cover, as we need to know their continuity behavior. This computational challenge can be avoided by slightly reformulating the theorem and the accompanying lemmas, as in Girard's modified topological proof [5]. We will discuss this aspect in Section 4.

2 In Lemma 3.6, compactness is used in the form of total boundedness; that is, given a sequence of points in X and an $\varepsilon > 0$, we want to know how many points we need to consider to ensure that two elements are ε -close. This can be made explicit by asking for a modulus of total boundedness, that is, a function that given an $\varepsilon > 0$ provides a $k > 0$ such that for any k elements $x_i \in X$, two elements x_i, x_j are ε -close for some $0 \leq i < j \leq k$. The argument in the induction base, that is, Lemma 3.1, is the same.

3 A final application of compactness occurs when we obtain the multiply recurrent point from the fact that for every $\varepsilon > 0$ there exists a $z \in X$ and an $n > 0$ such that $T_i^n z$ and z are ε -close for all $i = 1, \dots, l + 1$. From a logical point of view, this constitutes a uniformity result: Instead of ‘for every $\varepsilon > 0$ there exists a $z \in X$ ’, we obtain ‘there exists a $z \in X$ such that for all $\varepsilon > 0$ ’. However, as discussed above, for the proof of van der Waerden’s theorem the weaker, “nonuniform” version suffices as we only need some point z for a particular ε , namely, $\varepsilon = 1$. Thus, this application of compactness can be omitted here.

4 Elimination of Minimality and Girard’s Proof

In the proof by Furstenberg and Weiss, the appeal to minimality is crucial: without it Lemma 3.2 cannot be strengthened to Lemma 3.5, and this strengthening is necessary to form the sequence of points z_n in Lemma 3.6. In [5], Girard gives a variant of Furstenberg and Weiss’s topological proof which avoids the appeal to minimality. This is possible because the proof of Lemma 3.2, as well as the proof of Lemma 3.1, proves more than is stated in the paper by Furstenberg and Weiss. Retaining this additional information makes the appeal to minimality in the induction step unnecessary.

Let us start the discussion with Lemma 3.1. As it is stated in [2], the lemma claims the existence of *some* $x \in X$ and *some* $n \in \mathbb{N}$ such that $d(x, T^n x) < \varepsilon$. This is because compactness guarantees that for every $\varepsilon > 0$ there is an $m \in \mathbb{N}$ such that among any m points in X two are ε close. Applying this to the set $\{x_0, Tx_0, T^2x_0, \dots, T^{m-1}x_0\}$, we obtain—for any choice of $x_0 \in X$!—two points, for example, $T^i x_0, T^j x_0$ for $0 \leq i < j < m$ are ε -close. Let $x = T^i x_0$ and $n = j - i$ and the lemma follows.

However, since the value m is uniform in the choice of $x_0 \in X$, we have actually proved the following.

Notation Let G^T be a finitely generated group, that is, generated by the set $T = \{T_1, \dots, T_l\}$, then $G_M^T \subseteq G^T$ (for $M > 0$) is the subset of group elements which can be written as a word of length $< M$ over the generators. If there is no confusion about the generating set, we drop the superscript T .

Lemma 4.1 *Let (X, d) be a compact metric space, let $T : X \rightarrow X$, and let G be the group generated by T . Let $\varepsilon > 0$. Then there exist $N, M \in \mathbb{N}$ such that for each $x \in X$ $\min_{0 < n \leq N} \min_{g \in G_M} d(T^n gx, gx) < \varepsilon$.*

If one treats Lemma 3.2 under the same aspects, we obtain a similar stronger version. Namely, we obtain that there exist (explicit expressions for) $N, M \in \mathbb{N}$ such that for each $z \in X$ $\min_{0 < n \leq N} \min_{g, h \in G_M} d(T^n gz, hz) < \varepsilon$. In other words, instead of proving the existence of *some* $x, y \in X$, we obtain a finite set G_M such that for any $z \in X$ we have a finite list of candidates for explicit expressions gz, hz for $x, y \in X$ with $g, h \in G_M$.

This can be used in (a stronger version of) Lemma 3.6. With this additional information, we can form the (finite) sequence z_1, \dots, z_k (with k large enough to appeal to compactness for a given $\varepsilon > 0$) satisfying $d(T^{n_j + \dots + n_{i+1}} z_j, z_i) < \varepsilon$ for any $0 \leq i < j \leq k$ without appealing to minimality first. Already the example of three points z_0, z_1, z_2 is illustrative.

Using the strengthened version of Lemma 3.2, we obtain N_1, M_1 for a given $\varepsilon_1 = \varepsilon/2$. Let $\varepsilon_2 > 0$ be small enough so that $d(x, y) < \varepsilon_2$ implies $d(T^n gx, T^n gy) < \varepsilon_1/2$ for all $n \leq N_1$ and $g \in G_{N_1}$. Next, obtain N_2, M_2 for ε_2 .

Choose $y \in X$ arbitrary. By the strengthened Lemma 3.2, there exists $g_2, h_2 \in G_{M_2}$ and $n_2 \leq N_2$ such that $d(T^{n_2} g_2 y, h_2 y) < \varepsilon_2$. By choice of ε_2 this implies $d(T^{n_1+n_2} g_1 g_2 y, T^{n_1} g_1 h_2 y) < \varepsilon_2$. Moreover, there exist $g_1, h_1 \in G_{M_1}$ and $n_1 \leq N_1$ such that $d(T^{n_1} g_1 h_2 y, h_1 h_2 y) \varepsilon_1$. Let $z_2 = g_1 g_2 y, z_1 = h_1 g_2 y$, and $z_0 = h_2 h_1 x$; then those three points satisfy the above requirement. This can obviously be extended to arbitrarily long finite sequences z_n . Note that this depends crucially on the uniformity of the bounds on N_i, M_i in the starting point $y \in X$.

In conclusion, retaining more information (here: uniform bounds) in Lemma 3.1 and Lemma 3.2 allows one to eliminate Lemma 3.5 and the appeal to minimality therein and prove (the conclusion of) Lemma 3.6 directly from (the conclusion of) Lemma 3.2. Interpreting the notion of minimality is a far greater computational challenge than keeping track of the groups G_M .

We have already sketched the idea behind Girard’s modification of the proof by Furstenberg and Weiss. For completeness sake, we briefly state the lemmas as they appear in [5], which we will analyze in detail in Section 5. In Girard’s proof the structure is slightly different.

Lemma 4.2 *Let (X, d) be a compact metric space, let T be a homeomorphism of X , and let G be commutative group of homeomorphisms of X such that T commutes with G . Assume $A \subseteq X$ is closed and left invariant by G . If*

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \exists S_1, \dots, S_M \in G \forall x \in A (\min_{0 \leq n \leq N} \min_{0 \leq i, j \leq M} d(T^n S_i x, S_j x) < \varepsilon),$$

then

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \exists S_1, \dots, S_M \in G \forall x \in A (\min_{0 \leq n \leq N} \min_{0 \leq i, j \leq M} d(T^n S_i x, S_j x) < \varepsilon).$$

This lemma corresponds to the statement that the conclusion of Lemma 3.2, with the existence of x, y made explicit as discussed above, implies the conclusion of Lemma 3.6. From this lemma, the following version of the Multiple Birkhoff Recurrence theorem is proved by induction on l .

Theorem 4.3 *Let T_1, \dots, T_l be homeomorphisms of a compact space (X, d) and let the T_i commute with each other. Let G be the commutative group generated by T_1, \dots, T_l , then*

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \exists S_1, \dots, S_M \in G \\ \forall x \in X \exists n \leq N \exists i \leq M (d(T_1^n S_i x, S_i x) < \varepsilon \wedge \dots \wedge d(T_l^n S_i x, S_i x) < \varepsilon).$$

We will adapt Girard’s idea of modifying the Furstenberg and Weiss proof but will keep the structure of the original proof; thus we merely skip Lemma 3.5 and the appeal to minimality therein. Also, instead of making the S_1, \dots, S_M explicit, we will compute a bound on M such that the group G_M contains S_1, \dots, S_M .

5 Proof Analysis

In [5], Girard analyzes his modified topological proof of the Multiple Birkhoff Recurrence theorem using cut elimination. However, Girard does not analyze the full proof but rather the specialization of the proof to the concrete compact metric space

of sequences representing (shifts of) a given coloring of the integers. From a sufficiently formal (now first-order) proof, Girard then obtains a purely combinatorial proof by eliminating those cuts which “prevent the proof from being purely combinatorial” ([5], p. 244). The bounds obtained from this combinatorial first-order proof are the same bounds van der Waerden obtains from his combinatorial proof.

The analysis of Girard’s modified proof in this section treats the general case of the Multiple Birkhoff Recurrence theorem; that is, the extracted bounds are valid for arbitrary compact metric spaces and arbitrary homeomorphisms. Only later, the generic bounds are applied to the particular compact metric spaces and homeomorphisms that van der Waerden’s theorem gives rise to, thus producing bounds in this concrete case. Furthermore, the proof analysis given below leaves the structure of the proof intact, whereas cut elimination changes the structure and blows up the size of the proof. The individual steps of this cut-free proof may be simpler and more elementary, but due to the increased size and lost structure the proof as a whole is more difficult to understand.

The proof analysis presented in this section is based on Gödel’s functional (also: *Dialectica*) interpretation, more precisely, on the monotone variant due to Kohlenbach. The general idea of analyzing proofs using functional interpretation is to give a computational interpretation of the axioms, constants, and derivation rules of a given formal system by functionals of higher type. Then a sufficiently formal proof can inductively be transformed into an enriched proof, from which some desired additional information can be read off. Note that this does not significantly change the structure of the proof. Of course, there are natural limits to what can be achieved by any kind of proof interpretation: the halting problem can be expressed by a simple $\forall\exists\forall$ -statement,² where the existential quantifier expresses after how many steps a given Turing machine halts, if it halts at all. Obviously, there can be no computational realizer for that existential quantifier.

With monotone functional interpretations, one looks for bounds on quantifiers rather than exact realizers. This simplifies the interpretation of axioms and rules and may also allow to interpret certain nonconstructive principles that allow for computable bounds but do not have an exact computational interpretation at all, for example, weak König’s lemma, which can be considered a kind of choice axiom.

Using functional interpretations, one can prove general metatheorems that classify theorems and proofs from which additional information may be extracted. These metatheorems do not only apply to Peano arithmetic in all finite types, and thus also to metric spaces that can be given an effective representation in such formal systems; they may also be extended to treat abstract metric and normed spaces, where the spaces are represented by a new type X and the defining algebraic axioms, but where no inherent computability structure is given. For the most recent metatheorems, see [8] and [4]. For a comprehensive treatment of functional interpretations and other proof interpretations, see [7] and [13].

Now, general logical metatheorems are concerned with formal systems and formal proofs and mechanical procedures to transform a given proof into an enriched proof. In practice, it is rarely necessary to formalize a proof completely in order to carry out proof mining, and the proof analysis is guided by heuristics (derived from the underlying proof interpretation) rather than blind mechanical procedures.

The first step is usually to ensure that a given proof is in principle formalizable in some formal system to which some metatheorem applies. This is done by identifying the mathematical principles, general or particular, to some special structure like compact metric spaces, which are actually used in the proof. This in itself allows one to check with existing metatheorems so as to predict the kind of additional information one can extract from the proof, which, for example, allows one to establish uniformities without carrying out any further proof analysis.

Next, the main statements and lemmas of the theorem and proof to be analyzed are put into a suitable form. This is done by asking for additional information, that is, asking for a modulus of continuity when a function is assumed to be continuous or unwinding common mathematical abbreviations such as $\lim_{n \rightarrow \infty} x_n = 0$, which actually means $\forall \varepsilon > 0 \exists n \forall m \geq n (|x_m| < \varepsilon)$. We give examples of this below.

Then one goes through the steps of the (informal) proof and carries the necessary computational information from step to step. Where the computational meaning of a step is not obvious, one formalizes that step in greater and greater detail, until either the computational meaning becomes apparent or can be obtained by directly applying the proof interpretation of choice.

We now turn to the actual analysis of the Furstenberg-Weiss proof of the Multiple Birkhoff Recurrence theorem. As discussed in Section 4, the main building blocks of the proof are the triangle inequality, the continuity behavior of various homeomorphisms, respectively, finite subsets of groups of homeomorphisms and appeals to the total boundedness of the space. The triangle inequality has no computational content. For continuity, we ask for a modulus of continuity, and since the space (X, d) is compact this is equivalent to a uniform modulus of continuity.

Definition 5.1 A function ω_T is a modulus of uniform continuity for a mapping $T : X \rightarrow X$, if $d(x, y) < \omega_T(\varepsilon) \Rightarrow d(Tx, Ty) < \varepsilon$ for every $\varepsilon > 0$ and every $x, y \in X$.

Likewise, total boundedness is represented by a suitable modulus.

Definition 5.2 A function γ is a modulus of total boundedness if, for every $\varepsilon > 0$ and for any $x_0, \dots, x_{\gamma(\varepsilon)}$, there exist some $0 \leq i < j \leq \gamma(\varepsilon)$ such that $d(x_i, x_j) < \varepsilon$.

Also note that to obtain computable bounds, we eventually represent such moduli by number-theoretic functions $\gamma : \mathbb{N} \rightarrow \mathbb{N}$, where ε and δ are represented by $k, l \in \mathbb{N}$, that is, $2^{-k}, 2^{-l}$. Here we stick to the usual mathematical notation for readability.

In these variants of the lemmas, we generally fix a group G to be the group generated by a set T_1, \dots, T_l of homeomorphisms of X . The “inputs” (i.e., parameters of the theorem) are a common modulus ω_T for the set of homeomorphisms and a modulus γ of total boundedness as well as, of course, an $\varepsilon > 0$. The outputs are an upper bounds $N, M > 0$ such that $d(T^n gx, gx) < \varepsilon$ for some $0 \leq n \leq N$ and $g \in G_M^T$.

Lemma 5.3 *Let (X, d) be a totally bounded metric space with modulus γ of total boundedness, let $T : X \rightarrow X$, and let G be the group generated by T . Let $\varepsilon > 0$ be given and let $N = M = \gamma(\varepsilon/2)$. Then for each $x \in X$ $\min_{0 < n \leq N} \min_{g \in G_M} d(T^n gx, gx) < \varepsilon$.*

Proof This is the same proof as before, only now the total boundedness of X has been made explicit by the modulus γ . □

In the following, (X, d) will be a totally bounded metric space with modulus γ and T_i, S_i will refer to commuting homeomorphisms of X .

Lemma 5.4 *Suppose ϕ_N^l and ϕ_M^l satisfy the following: for any $\delta > 0$ and any T_1, \dots, T_l of X with common modulus of uniform continuity ω_T , if $N = \phi_N^l(\delta, \gamma, \omega_T)$ and $M = \phi_M^l(\delta, \gamma, \omega_T)$, then for any $x \in X$ simultaneously*

$$\min_{0 < n \leq N} \min_{g \in G_M} d(T_i^n g x, g x) < \delta \text{ for } i = 1, \dots, l.$$

Let $\varepsilon > 0$ and any S_1, \dots, S_{l+1} with common modulus of continuity ω_S be given and let $N' = \phi_N^l(\varepsilon, \gamma, \omega_S^2)$ and $M' = 2 \phi_M^l(\varepsilon, \gamma, \omega_S^2) + N'$, where $\omega_S^2 = \omega_S \circ \omega_S$. Then for any $x \in X$ simultaneously

$$\min_{0 < n' \leq N'} \min_{g', h' \in G_{M'}} d(S_i^{n'} g' x, h' x) < \varepsilon$$

for $i = 1, \dots, l + 1$.

Proof Let $\varepsilon > 0$ be given and fix $x \in X$. From the homeomorphisms S_i form $T_j = S_j S_{l+1}^{-1}$ for $j = 1, \dots, l$ and apply ϕ_N^l, ϕ_M^l to ε, γ and ω_S^2 as above. By assumption on ϕ_N^l, ϕ_M^l , for some $g \in G_M^T = G_{2M}^S$ and some $n \leq N$ simultaneously $d(S_j^n S_{l+1}^{-n} g x, g x) < \varepsilon$ for $j = 1, \dots, l$. Define $g' = S_{l+1}^{-n} g$ and $h' = g$, then also simultaneously $d(S_i^n g' x, h' x) < \varepsilon$ for $i = 1, \dots, l + 1$. \square

As discussed in Section 4, we don't need an equivalent of Lemma 3.5 and instead directly "glue together" Lemma 3.2 and 3.6.

Lemma 5.5 *Suppose ψ_N^l, ψ_M^l satisfy the following: for any $\delta > 0$ and any T_1, \dots, T_l with common modulus of uniformity ω_T , if $N = \psi_N^l(\delta, \gamma, \omega_T)$ and $M = \psi_M^l(\delta, \gamma, \omega_T)$, then simultaneously*

$$\min_{n \leq N} \min_{g, h \in G_M} d(T_i^n g x, h x) < \delta \text{ for } i = 1, \dots, l.$$

Let $\varepsilon > 0$ and homeomorphisms T_1, \dots, T_l be given and define $\chi_N(i) = \psi_N^l(\varepsilon_i, \gamma, \omega_T)$ and $\chi_M(i) = \psi_M^l(\varepsilon_i, \gamma, \omega_T)$, where $\varepsilon_1 = \varepsilon/4$ and $\varepsilon_{k+1} = \omega_T^{\chi_N(k)+k\chi_M(k)}(\varepsilon_k/2)$. Let $N' = \chi_N(\gamma(\varepsilon/2)) \cdot \gamma(\varepsilon/2)$ and $M' = \chi_M(\gamma(\varepsilon/2)) \cdot \gamma(\varepsilon/2)$. Then for any $x \in X$ simultaneously

$$\min_{n' \leq N'} \min_{g \in G_{M'}} d(T_i^n g x, g x) < \varepsilon \text{ for } i = 1, \dots, l.$$

Proof Set $\varepsilon_1 = \varepsilon/4$. Given ε_k , we want to form ε_{k+1} satisfying (1) $\varepsilon_{k+1} < \varepsilon_k/2$, and (2) for all $n < \chi_N(k)$ and all $v = g_1 g_2 \dots g_k, k' \leq k$ and $g_i \in G_{\chi_M(i)}$ we have that $d(x, x') < \varepsilon_{k+1}$ implies that $d(T_i^n v x, T_i^{n'} v x') < \varepsilon_k/2$. With the mild assumption that $\omega_T(\varepsilon) \leq \varepsilon$, one readily sees that the above definition of the sequence ε_n satisfies these two properties.

Now, for any given k , we define z_0, \dots, z_k as follows:

- (i) $y_k = x$,
- (ii) $y_{k-1} = h_k y_k$ for g_k, h_k and $n_k \leq \chi_N(k)$ such that $d(T^{n_k} g_k x, h_k x) < \varepsilon_k$ (which exist by the premise),
- (iii) $z_i = g_i \dots g_1 y_i$ for $i = 1, \dots, k$.

Then using the properties of the ε_k and the corresponding g_k, h_k , we get that $d(T^{n_k} z_k, z_{k-1}) < \varepsilon_k/2$. Furthermore, using the properties of the ε_k and n_k , this yields $d(T^{n_j + \dots + n_{i+1}} z_j, z_i) < \varepsilon/2$ for all $0 \leq i < j \leq k$. By total boundedness, some $0 \leq i < j \leq \gamma(\varepsilon/2)$ satisfy $d(z_i, z_j) < \varepsilon/2$, and the result follows. \square

Multiple Birkhoff Recurrence Theorem (effective version) Let (X, d) be a metric space with modulus of total boundedness γ , let T_1, \dots, T_l be commuting homeomorphisms of X with common modulus of uniform continuity ω_T , and let G be the group generated by T_1, \dots, T_l . Then for every $\varepsilon > 0$ there exist $N, M > 0$ (to be defined below) such that for every $x \in X$ simultaneously $\min_{0 < n \leq N} \min_{g \in G_M} d(T_i^n gx, gx) < \varepsilon$ for $i = 1, \dots, l$. Define

1. $N^1(\varepsilon, \gamma, \omega) = M^1(\varepsilon, \gamma, \omega) = \gamma(\varepsilon/2)$;
2. $\varphi_N^{k+1}(i) = N^k(\varepsilon_i^{k+1}, \gamma, \omega^2)$;
3. $\varphi_M^{k+1}(i) = 2M^k(\varepsilon_i^{k+1}, \gamma, \omega^2) + N^k(\varepsilon_i^{k+1}, \gamma, \omega^2)$;
4. $\varepsilon_1^k = \varepsilon/4$ and $\varepsilon_{i+1}^k = \omega^{\varphi_N^k(i)+i \cdot \varphi_M^k(i)}(\varepsilon_i/2)$;
5. $N^{k+1}(\varepsilon, \gamma, \omega) = \varphi_N^{k+1}(\gamma(\varepsilon/2)) \cdot \gamma(\varepsilon/2)$;
6. $M^{k+1}(\varepsilon, \gamma, \omega) = \varphi_M^{k+1}(\gamma(\varepsilon/2)) \cdot \gamma(\varepsilon/2)$.

Then $N = N^l(\varepsilon, \gamma, \omega)$ and $M = M^l(\varepsilon, \gamma, \omega)$.

Proof This follows from the previous lemmas by combining the functionals obtained there. □

6 Van der Waerden's Theorem (Again)

With the bounds on the Multiple Birkhoff Recurrence theorem obtained in Section 5, we may obtain bounds on (the finite version of) van der Waerden's theorem. Given a q -coloring of the integers and a desired length k of the arithmetic progression, we simply write down a modulus γ for the corresponding metric space and a common modulus ω_T of uniform continuity for the necessary homeomorphisms T_1, \dots, T_{k-1} . These moduli will be expressed in terms of q and k and will be number-theoretic functions; that is, $\gamma, \omega_T : \mathbb{N} \rightarrow \mathbb{N}$. Instead of mapping an $\varepsilon > 0$ to a $\delta > 0$, they will map a $k \in \mathbb{N}$ representing a $\varepsilon = 2^{-k}$ to an $l \in \mathbb{N}$ representing a $\delta = 2^{-l}$. We then apply the bounds for the Multiple Birkhoff Recurrence theorem to these moduli and $\varepsilon = 1$ ($= 2^0$, so $k = 0$). Given the resulting N, M , we know that for any coloring x there are $0 < n \leq N$ and $m \leq M$ such that $x(m) = x(m+n) = \dots = x(m+(l-1)n)$. Thus $M + (k-1)N$ is an upper bound on the length of an interval, so that any q -coloring of that interval contains an arithmetic progression of length k .

Let us recall the definition of the space (X, d) derived from a given q -coloring of the integers. A q -coloring can be represented by a function $x : \mathbb{Z} \rightarrow \{1, \dots, q\}$. Given the shift T such that $(Tx)(n) = x(n+1)$ the space³ $X = \{T^n x | n \in \mathbb{Z}\}$. The metric on X is $d(x, y) = 2^{-k}$ for the largest k such that $x(l) = y(l)$ and $x(-l) = y(-l)$ for all $l \leq k$.

It is then clear that for a given q -coloring and a 2^{-k} the set of all finite sequences of length $2k+1$ with values in $\{1, \dots, q\}$ is a 2^{-k} -dense subset of X . Thus, $\gamma(k) = q^{2k+1}$ is a modulus of total boundedness for X . It is also clear that the 1-shift T is uniformly continuous with modulus $\omega_T(k) = k+1$. For the mappings T_1, \dots, T_l with $T_i = T^i$, that is, T iterated i times, we have the common modulus of uniformity $\omega_T(k) = k+l$.

Applying the bounds for the Multiple Birkhoff Recurrence theorem to these moduli, we easily see that—modulo some ad hoc optimizations of the bounds—the bounds obtained in this way are identical to the bounds obtained by van der Waerden

and identical to the bounds obtained by Girard from his modified Furstenberg and Weiss proof. This is not very surprising as, upon closer inspection, the proofs by van der Waerden and by Girard are identical with regard to their algorithmic content; that is, step by step the two proofs employ the same constructions and operations to prove the result. The only difference is that one proof is formulated in the context of combinatorics, whereas the other is formulated in the context of topological dynamics.

Let us again recall the analogy between van der Waerden's theorem and the Multiple Birkhoff Recurrence theorem. On the combinatorial side, we deal with colorings of the natural numbers and arithmetic progressions, on the topological side with homeomorphisms of a compact metric space and ε -recurrent points. The length of the progression corresponds to the number of homeomorphisms, whereas the number of colors corresponds to the number of ε -neighborhoods necessary to cover the entire space. In both settings, the simple case is where you only have a progression of length 2, respectively, only one homeomorphism, as then we only need to exhaust the number of colors, respectively, ε -neighborhoods.

To inductively prove the general combinatorial result, we solve the problem for long progressions and few colors by transforming it into a problem with shorter progressions and more colors until we reach the trivial case. The greater set of colors results from a coloring of blocks of natural numbers derived from the coloring of the natural numbers. The length of the blocks must be large enough to allow one to find long enough progressions of smaller blocks within each block, and the depth of blocks within blocks must be large enough to allow a sequence of starting points and endpoints for progressions across blocks on different levels to exhaust the number of colors at the outset.

Topologically, we find a recurrent point for many homeomorphisms and big ε by finding a point for fewer homeomorphisms and smaller ε . As in the combinatorial setting, we find a recurrent point for fewer homeomorphisms (shorter progressions) for a sequence of smaller and smaller ε (larger and larger number of colors) until the length of that sequence (and a sequence of points derived from that) exhausts the ε -neighborhoods for the original ε (original number of colors). Even the "growth" of the ε happens in a similar way: given one $\varepsilon_k > 0$ in a sequence, the next $\varepsilon_{k+1} > 0$ must be small enough to ensure that "shifts" $T_i^{n_k}$ (by continuity of the T_i) do not move points that are ε_{k+1} -close more than ε_k apart. This has an analogue in the combinatorial setting in making blocks large enough to contain progressions of smaller blocks.

Thus, the bounds are identical, because the proofs in principle are identical! While we obtain no new bounds, we do obtain a deeper insight into the relationship between the combinatorial and topological proofs of van der Waerden's theorem and the Multiple Birkhoff Recurrence theorem, respectively. These connections may be useful when further exploring the computational content of topological and analytical proofs of combinatorial theorems. The reverse is also conceivable, as not only bounds, but also proof ideas may be transferred from one setting to the other. Thus one may perhaps obtain a "new," different topological proof of the Multiple Birkhoff Recurrence theorem from Shelah's ingenious, improved combinatorial proof of van der Waerden's theorem. The author hopes to explore this in a forthcoming paper.

Notes

1. Chronologically, this result predates Gowers's result and coincides with Furstenberg's ergodic theoretic proof of Szemerédi's theorem [3], so the surprise is genuine.
2. $\forall\exists\forall$ describes the quantifier prefix before a decidable statement. Here, the statement is, "for every Turing machine x , every input i , there is a y such that for all z either x applied to i stops after y stops or it doesn't stop for any number z of steps"; checking whether a Turing machine stops after a given number of steps is, of course, decidable.
3. In Section 2, we defined X to be the completion of this space, but the completeness is not needed, so here (X, d) is only totally bounded.

References

- [1] Furstenberg, H., *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, Princeton, 1981. [Zbl 0459.28023](#). [MR 603625](#). [435](#), [437](#)
- [2] Furstenberg, H., and B. Weiss, "Topological dynamics and combinatorial number theory," *Journal d'Analyse Mathématique*, vol. 34 (1978), pp. 61–85. [Zbl 0425.54023](#). [MR 531271](#). [432](#), [434](#), [435](#), [436](#), [437](#), [438](#)
- [3] Furstenberg, H., "Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions," *Journal d'Analyse Mathématique*, vol. 31 (1977), pp. 204–56. [Zbl 0347.28016](#). [MR 0498471](#). [445](#)
- [4] Gerhardy, P., and U. Kohlenbach, "General logical metatheorems for functional analysis," *Transactions of the American Mathematical Society*, vol. 360 (2008), pp. 2615–60. [Zbl 1130.03036](#). [MR 2373327](#). [440](#)
- [5] Girard, J.-Y., *Proof Theory and Logical Complexity*, vol. 1 of *Studies in Proof Theory. Monographs*, Bibliopolis, Naples, 1987. [Zbl 0635.03052](#). [MR 903244](#). [432](#), [437](#), [438](#), [439](#), [440](#)
- [6] Gowers, W. T., "A new proof of Szemerédi's theorem," *Geometric and Functional Analysis*, vol. 11 (2001), pp. 465–588. [Zbl 1028.11005](#). [MR 1844079](#). [433](#)
- [7] Kohlenbach, U., "Applied Proof Theory: Proof Interpretations and Their Use in Mathematics," to appear in Springer Monographs in Mathematics, 2008. [440](#)
- [8] Kohlenbach, U., "Some logical metatheorems with applications in functional analysis," *Transactions of the American Mathematical Society*, vol. 357 (2005), pp. 89–128. [Zbl 1079.03046](#). [MR 2098088](#). [440](#)
- [9] Landman, B. M., and A. Robertson, *Ramsey Theory on the Integers*, vol. 24 of *Student Mathematical Library*, American Mathematical Society, Providence, 2004. [Zbl 1035.05096](#). [MR 2020361](#). [433](#)
- [10] McCutcheon, R., *Elemental Methods in Ergodic Ramsey Theory*, vol. 1722 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin, 1999. [Zbl 0945.05063](#). [MR 1738544](#). [435](#)
- [11] Shelah, S., "Primitive recursive bounds for Van der Waerden numbers," *Journal of the American Mathematical Society*, vol. 1 (1988), pp. 683–97. [Zbl 0649.05010](#). [MR 929498](#). [433](#), [434](#)

- [12] Szemerédi, E., “On sets of integers containing no k elements in arithmetic progression. Collection of articles in memory of Jurii Vladimirovič Linnik,” *Acta Arithmetica*, vol. 27 (1975), pp. 199–245. [Zbl 0303.10056](#). [MR 0369312](#). [434](#)
- [13] Troelstra, A. S., editor, *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, vol. 344 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin, 1973. [Zbl 0275.02025](#). [MR 0325352](#). [440](#)
- [14] van der Waerden, B. L., “Beweis einer baudeischen Vermutung,” *Nieuw Archief voor Wiskunde*, vol. 15 (1927), pp. 212–16. [Zbl 53.0073.12](#). [432](#), [433](#)

Acknowledgments

The author wishes to thank the reviewer for very helpful comments on improving the presentation in this paper.

Department of Mathematics
The University of Oslo
Blindern
N-0316 Oslo
NORWAY
philipp.gerhardy@gmail.com