2010

Ergodic and Combinatorial Proofs of van der Waerden's Theorem

Matthew Samuel Rothlisberger
Claremont McKenna College

CLAREMONT McKENNA COLLEGE

ERGODIC AND COMBINATORIAL PROOFS OF VAN DER WAERDEN'S THEOREM

Submitted to

Professor Ami Radunskaya
And
Professor Robert J. Valenza

And

Dean Gregory Hess

By

Matthew Rothlisberger

For

Senior Thesis

November 29, 2010

Contents

1 Introduction

As an undergraduate student I found that I enjoyed many pure math courses. Specifically number theory. When I first saw van der Waerden's theorem I was very intruiged. Van der Waerden's theorem can be stated in many ways, one of which is:

For any natural numbers l and k, there exists a number $n(k, l)$ such that when you split the set of integers $\{1, 2, 3, ..., n(k, l)\}$ into k classes one of these classes will contain an arithmetic sequence of length l.

This theorem was proved in 1928 by a young van der Waerden [5]. The idea of this theorem was first conjectured by Baudet who said, "that if l is a given natural number and if the set of natural numbers is divided over two classes, then at least one of these classes contains an arithmetic progression of length l." It is unknown as to why Baudet conjectured this [10]. I will be covering two proofs of this theorem, one of which is ergodic while the other is a combinatorial inductive proof.

The ergodic proof follows from the Multiple Birkhoff Recurrence theorem which was proved just before van der Waerden gave his proof [5]. Surprisingly van der Waerden gave a proof that did not rely on the Multiple Birkhoff Recurrence theorem. We now see that the result that van der Waerden proved is a result that is hinted at in many different fields. It has now been generalized by Szemeredi's Theorem, which was conjectured in 1936 [3], and finally proved in 1975 [9].

I was very curious as to how this theorem could be true. From my perspective it states a result that isn't very intuitive to me whatsoever. Luckily there was quite a bit of literature with proofs of this theorem. I began with the ergodic proof of the theorem. As I worked on it I found the going very difficult, so I found the inductive combinatorial proof. After completing the combinatorial proof I returned back to the ergodic proof and found the concepts it describes much less foreign. What is interesting about these two proofs is that they are new proofs of an old theorem.

The first proof, being inductive, focuses on building the $n(k, l)$ numbers. Then once we have a method for building these numbers we can show that we can find a $n(k, l + 1)$ number. This proof tells us that there is an upper bound for how large a set of the natural numbers we must use. The second proof begins by finding multiply recurrent points and then combines this result with symbolic dynamics to show that in any infinite string of k symbols we will find recurrent points simultaneously for l transformations. A result that we could prove using the methods in the second proof is that van der Waerden's theorem can be applied to multiple dimensions. Thus we have two proofs that bring useful results. This is a common theme in proofs of van der Waerden's theorem: since it can be stated in many forms, each proof yeilds a slightly greater result.

2 Inductive Proof of van der Waerden's Theorem

Before we begin working on this proof I feel it necessary to give a definition and restate the theorem.

ARITHMETIC PROGRESSION: An *arithmetic progression* is a sequence of numbers whose successive elements are all the same distance apart. As an example the even numbers make an arithmetic sequence with distance 2 between consecutive elements.

VAN DER WAERDEN'S THEOREM: *Let k and l be two arbitrary natural numbers. Then there exists a natural number $n(k,l)$ such that, if an arbitrary segment, of length $n(k,l)$, of the sequence of natural numbers is divided in any manner into k classes (some of these classes may be empty), then an arithmetic progression of length l appears in at least one of these classes.* [6]

The proof I will be following is from a book by A.Y. Khinchin. Published in 1952, Khinchin wrote this proof originally as a letter in 1945 to a hospitalized soldier who had been a student of his [6].

Now since we are going to do an inductive proof of the theorem we need to establish our base case. Notice that if we let $l = 2$ the theorem is very trivial. We make our number $n(k,2) = k+1$. Now we see that if we have k classes of a segment of length $k+1$ of $\mathbb{N}$, we will have at least one segment of more than one number and this pair of numbers is an arithmetic progression of length 2. Now we shall induct on l to prove the theorem. We will assume that the theorem has been proved for some $l \geq 2$ and arbitrary k values and when we show that the theorem is still valid for $l+1$ values it will also hold for arbitrary k values. [6]

So what we 'know' (by our base case) is that for every $k \in \mathbb{N}$ there exists a $n(k,l) \in \mathbb{N}$ such that if we were to take an arbitrary segment of $\mathbb{N}$ of length $n(k,l)$ and divide it into k classes then we would find an arithmetic progression of length l in at least one of these k classes. Now we are trying to show that this assumption holds for $n(k,l+1)$. To do this we will construct a number $n(k,l+1)$. We will start by defining

$$q_0 = 1, n_0 = n(k,l)$$

And now if we already have q_{s-1} and n_{s-1} for some $s > 0$, we will say,

$$q_s = 2n_{s-1}q_{s-1}, n_s = n(k^{q_s}, l)(s = 1, 2, ...) \tag{1}$$

We see that our numbers q_s and n_s are defined for arbitrary $s \geq 0$. What we are doing here is defining a number q_k that is $n(k,l+1)$. So now we will be showing that if we take a segment of length q_k of $\mathbb{N}$ and divide it into k classes we can find an arithmetic progression of length $l+1$ in at least one of the classes.

I will define a term that will be useful, assume that Δ is our segment of length q_k of $\mathbb{N}$ that is divided into k classes.

Same Type: This definition can be applied to numbers and sets. Consider a segment X of the natural numbers that is split into k classes. We will say that the numbers $x, y \in X$ are of the *same type* if x and y belong to the same class. We will then write $x \approx y$. Now if we have two equally long subsegments of X, call them $\delta = (x, x+1, ..., x+r)$ and $\delta' = (y', y'+1, ..., y'+r)$, we say the segments δ and δ' are of the *same type* if $x \approx y', x+1 \approx y'+1, ..., x+r \approx y'+r$. We then write $\delta \approx \delta'$. [6] So we can see that if δ and δ' are the *same type* then we can find subsegments of δ and δ' that are also of the same type.

Example: Consider the set $\Delta = \{1, 2, 3, ..., 11, 12\}$, with $k = 3$. Now I will break Δ into 3 classes, $\delta_1 = \{1, 3, 5, 7\}, \delta_2 = \{2, 4, 6, 8\}$, and $\delta_3 = \{9, 10, 11, 12\}$. We can see that 1 and 3 are the *same type* because they both belong to δ_1, thus $1 \approx 3$. Similarly we can see that the subsegments $(5, 6, 7)$ and $(1, 2, 3)$ are the *same type* because $1 \approx 5$, $2 \approx 6$, and $3 \approx 7$.

Now let me point out something to keep in mind as we continue our proof, *classes* and *types* are competely different objects. We will be dealing with types far more than classes. It is good to try and remember the relationship between types and classes as we go on.

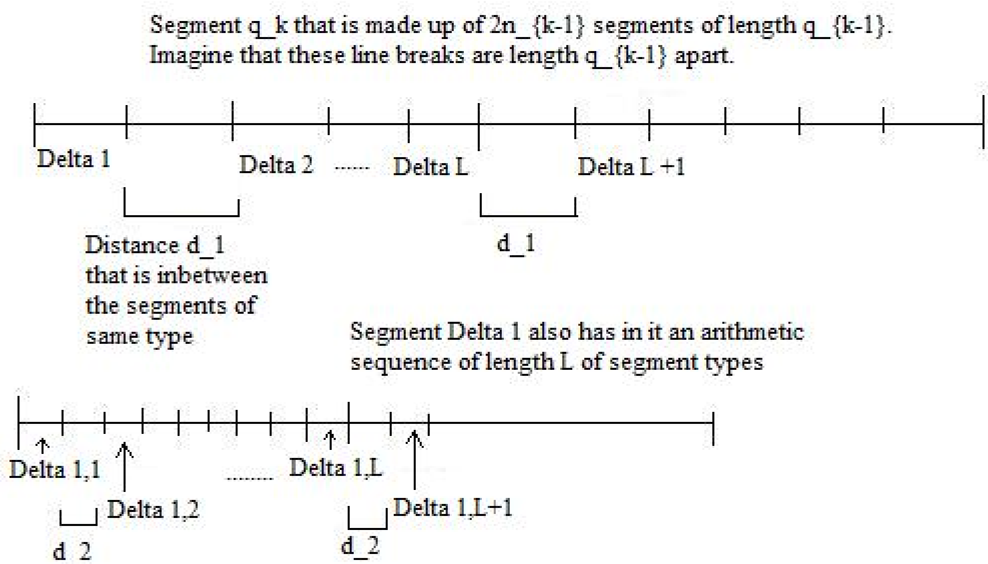
Notice this too about *same types*, that the total number of different *types* for numbers of a segment Δ is k. You can see this because if we have k classes of Δ then a and b are the *same type* only if they are in the same class, thus we have k different *types*. Now when we consider subsegments of length m of Δ, the different possible *types*, of segments of length m, is k^m. We can see this because a *type* must have each successive element from the same class, thus for each we have k classes from which we must select m elements, which gives us k^m different *types*.

Now recall from (1) that $q_k = 2n_{k-1}q_{k-1}$, with this we can call our segment Δ (that is of length q_k) a sequence of $2n_{k-1}$ subsegments of length q_{k-1}. Simply said we are taking $2n_{k-1}$ successive segments of length q_{k-1} of $\mathbb{N}$. To further clarify this important step, we can say that Δ is a segment of the natural number line of length q_k that is composed of (or can be broken into) $2n_{k-1}$ segments of length q_{k-1}. Now these q_{k-1} subsegments of Δ have $k^{q_{k-1}}$ different types (with respect to subsegments of q_{k-1} not numbers). And the left half of Δ contains n_{k-1} of these subsegments (simply because Δ is made up of $2n_{k-1}$ of the q_{k-1} subsegments, thus the left half, like the right half, is made up of n_{k-1} of these subsegments). And we know by (1) that $n_{k-1} = n(k^{q_{k-1}}, l)$, which we know (by our assumption that van der Waerden's theorem holds for l) means that in the left half of Δ there exists an arithmetic progression of l of these subsegments of the same type,

$$\Delta_1, \Delta_2, ..., \Delta_l \tag{2}$$

of length q_{k-1}. To recap; in the left hand side of Δ we know that there exists an arithmetic progression of length l in the **types**, of the segments of length q_{k-1}, by the the inductive hypothesis, which states that for a number $n_{k-1} = n(k^{q_{k-1}}, l)$ there exists an arithmetic sequence of length l in at least one of the $k^{q_{k-1}}$ classes (and these classes are actually the total types that a segment of length q_{k-1} can have). So now what we have is an arithmetic sequence of length l of partitions of Δ, and all of these partitions are of the same type. Thus the first elements of Δ_i and $\Delta_{i+1}(1 \leq i < l)$ are the same distance apart, call it d_1 (as are the second, third, etc elements of each set, this follows from the fact that the sets are of the same type). Now we have $\Delta_1, \Delta_2, ..., \Delta_l$, that are all of length q_{k-1} and their first elements all belong to the same class of Δ, and their second, third, ..., q_{k-1} elements all belong to the same class as well. So we have found q_{k-1} arithmetic sequences of length l where d_1 is the distance between the consecutive terms.

Now we shall add another term to the end of our Δ_i's. Call this new term Δ_{l+1}, with. Where the first element of Δ_{l+1} is d_1 away form the first element of Δ_l. Δ_{l+1} may project beyond the left side of Δ (or may be completely out of the the left side), but is obviously contained in the entire segment Δ. We have constructed a sequence of segments $\Delta_1, \Delta_2, ..., \Delta_l, \Delta_{l+1}$, that contain an arithmetic progression of length $l+1$ and difference d_1 between the segments of length q_{k-1}. Also we know that $\Delta_1, \Delta_2, ..., \Delta_l$ are of the same type, but we do not know anything about the type of Δ_{l+1}.



The next step is to select an arbitrary one of our Δ_i, in the sequence of segments that we just constructed, with $1 \leq i \leq l$. Call this selected segment Δ_{i_1}, we know that this segment is of length q_{k-1}, and by (1) we know $q_{k-1} = 2n_{k-2}q_{k-2}$. So the left half of Δ_{i_1} can be considered to be made up of n_{k-2} segments of length q_{k-2}, and we know that for subsegments of this length there is $k^{q_{k-2}}$ types possible (segment types). Also by (1) we see that $n_{k-2} = n(k^{q_{k-2}}, l)$ which means that Δ_{i_1} must contain an arithmetic progression of l of these subsegments of the same type, $\Delta_{i_1 i_2}$, of length q_{k-2}. Again these $\Delta_{i_1 i_2}$ elements are the same type because by the definition of $n(k^{q_{k-2}}, 1)$ we know that there exists an arithmetic sequence of length l of the $k^{q_{k-2}}$ types possible. We shall call d_2 the difference between the neighboring segments $\Delta_{i_1 i_2}$. And again we shall add the $l+1$ term, $\Delta_{i_1 l+1}$, that must not be totally contained in the left hand side of Δ_{i_1} but is contained in the whole of Δ_{i_1}. Again we know nothing of the type of $\Delta_{i_1 l+1}$. All we know is that the distance between the first element of $\Delta_{i_1 l}$ and $\Delta_{i_1 l'}$ is d_2.

Notice that when you see Δ_{i_1} it means that we have done our construction on Δ and now are left with a segment of length q_{k-1}. Thus when we have a segment described as $\Delta_{i_1 i_2 i_3}$ we can see by pattern that it will be of length q_{k-3}.

Also you can see that this construction that we have done on Δ_{i_1} can be done congruently to all the other segments $\Delta_{i_1}(1 \leq i_1 \leq l+1)$ of our original sequence of subsegments of length q_{k-1}. What is meant by this is that once we have split our Δ_{i_1} into it's arithmetic sequence of segments of same type we can partition all other Δ_{i_1} segments in the exact same way. It is obvious that all of our new segments are all of the same type because because all the Δ_{i_1} are of the same type, so if you take the first 5 elements of two Δ_{i_1} those two subsegments must be the same type since they come from the same place in two segments that are the same type. Thus we can state that when we take two arbitrary segments whose indices do not excede l, they are the same type:

$$\Delta_{i_1 i_2} \approx \Delta_{i'_1 i'_2}, \text{ with } (1 \leq i_1, i_2, i'_1, i'_2 \leq l)$$

And it follows that we get

$$\Delta_{i_1 i_2 \ldots i_k} \approx \Delta_{i'_1 i'_2 \ldots i'_k} \tag{3}$$
$$\text{with } (1 \leq i_1, i_2, ..., i_k, i'_1, i'_2, ...i'_k \leq l)$$

We get this result by expanding on our previous finding. We know that $\Delta_{i_1 i_2} \approx \Delta_{i'_1 i'_2}$ for $(1 \leq i_1, i_2, i'_1, i'_2 \leq l)$, thus when we go to $\Delta_{i_1 i_2 i_3}$ we know that we are getting the same construction of sets in every other $\Delta_{i'_1 i'_2}$. Which means that we are taking sets in the same position with respect to the sets they belong in, which are of the same type. This obviously implies that if any subsets of two different sets that are of the same type, are in the same position, then they are also of the same type.

When we do our process (going from Δ, to Δ_{i_1}, to $\Delta_{i_1 i_2}$, ...) k times the result is an element of the form $\Delta_{i_1 i_2 \ldots i_k}(1 \leq i_1, i_2, ..., i_k \leq l')$. Notice that the length of one of these elements is $q_{k-k} = q_0$, and by (3) we know that $q_0 = 1$. This means that $\Delta_{i_1 i_2 \ldots i_k}$ is just one number from our original segment Δ.

Here are two facts that will be useful in finishing our proof:

1). Taking our result from (3) and if we select some $1 \leq s < k$, and if $i_{s+1}, i_{s+2}, ..., i_k$ are arbitrary indices taken from the sequence of numbers $1, 2, ..., l, l+1$, then the number $\Delta_{i_1 i_2 \ldots i_s i_{s+1} \ldots i_k}$ appears in the same position in the segment $\Delta_{i_1 \ldots i_s}$ as the number $\Delta_{i'_1 i'_2 \ldots i'_s i_{s+1} \ldots i_k}$ does in the segment $\Delta_{i'_1 \ldots i'_s}$. This follows from the fact that we are doing the same steps that we have done in our original $\Delta_{i_1 i_2 \dots}$ to all other $\Delta_{i'_1 i'_2 \dots}$ sets. Thus when we take two sets who are the same type, like $\Delta_{i_1 \dots i_s}$ and $\Delta_{i'_1 \dots i'_s}$, then do the same steps $(s+1, s+2, ..., k)$ we get a number from the same place in each segment. And if the numbers belong in the same place in two segments that are the same type then we can write:

$$\Delta_{i_1, i_2, \dots, i_s, i_{s+1}, \dots, i_k} \approx \Delta_{i'_1, i'_2,, \dots, i'_s, i_{s+1},, \dots, i_k} \tag{4}$$

when $(1 \leq i_1, ..., i_s, i'_1, ..., i'_s \leq l)$ and $(1 \leq i_{s+1}, i_{s+2}, ..., i_k \leq l+1)$. See how the $i_{s+1}, ...i_k$ values can equal $l+1$, because the first steps we've taken before we do these $i_{s+1}, ...i_k$ steps ensure that we are dealing with a segment with the same type as all the others. Thus we can perform the next steps even on its $l+1$ set because it's $l+1$ set is the same type as all the other $\Delta_1, ..., \Delta_i$ with $i \leq l$.

2). For $s \leq k$, with $i'_s = i_s+1$, $\Delta_{i_1, \dots, i_{s-1},, i_s}$ and $\Delta_{i_1, \dots, i_{s-1}, i'_s}$ are neighboring segments in the s-th step of our construction. So if we continue our construction in both these neighboring segments with any arbitrary $i_{s+1}, ..., i_k$, the numbers $\Delta_{i_1, \dots, i_{s-1}, i_s, i_{s+1},, \dots, i_k}$ and $\Delta_{i_1, \dots, i_{s-1}, i'_s, i_{s+1}, \dots, i_k}$ appear in the same position with respect to their neighboring segments. Which means that:

$$\Delta_{i_1, \dots, i_{s-1}, i_s, i_{s+1}, \dots, i_k} - \Delta_{i_1, \dots, i_{s-1}, i'_s, i_{s+1}, \dots, i_k} = d_s \tag{5}$$

since the two numbers are in the same position in the neighboring segments in the s-th step of our construction, and everything that is in our sequence of length l during the s-th step of our construction is distance d_s apart. So (5) naturally follows.

Now after I was presented this following information the big "Aha" moment happened and I hope you get it too. So continuing on let us consider these $k+1$ numbers of the segment Δ:

$$a_0 = \Delta_{l+1, l+1, \dots, l+1}$$

$$a_1 = \Delta_{1, l+1, l+1, \dots, l+1}$$

$$a_2 = \Delta_{1, 1, l+1, \dots, l+1}$$

$$\dots.$$

$$a_k = \Delta_{1, 1, 1, \dots, 1}$$

If you recall Δ was split into k classes, which means that at least two of the $k+1$ numbers, mentioned above, must be in the same class. And if any two of these numbers are in the same class then it implies that at least one number that is in an $l+1$ segment in our final stage of constructing our $\Delta_{i_1 i_2 \dots i_k}$ belongs to an arithmetic progression that is now of length $l+1$.

Let's assume that our two numbers, $a_r = \Delta_{1, l+1, l+1, \dots, l+1}$ and $a_s = \Delta_{1, 1, l+1, \dots, l+1}$. This means that by (4) we know that these two numbers are in the same position of the segments $\Delta_{1,1}$ and $\Delta_{1,l+1}$. And we know that there is an arithmetic progression of length l of segments of the same type, $\Delta_{1,1}, \Delta_{1,2}, ..., \Delta_{1,l}$ where every number that is in the same position in each segment belongs to an arithmetic progression of length l with spacing d_2 between each number. But now we know that there is one number in our segment $\Delta_{1,l+1}$ that belongs to the same class as a number in the same position in segment $\Delta_{1,1}$. Which means that since $\Delta_{1,1} \approx \Delta_{1,2} \approx ... \approx \Delta_{1,l}$ that all the numbers in the same position belong to the same class, and we know that there is another number in the same position in $\Delta_{1,l+1}$ that belongs to the same class as all the numbers in the same position as it (with respect to their segments). So we see that we have found an arithmetic progression of length $l+1$ with distance d_2 in one class of our segment of $\mathbb{N}$ of length q_k.

It is not hard to see how this is true for any a_r and a_s that are in the same class. Given any two we can see by our construction that it will imply an arithmetic progression of length $l+1$ where all the numbers are the same distance apart and belong to the same class.

Thus, we can see that we have proved van der Waerden's theorem. We have also shown that there is an upper bound on the length the segment must be for the result to hold. I will show that the upper bounds we have calculated are much larger than necessary.

Example:

Now I am going to use the rules we set up from (1) to build the number $n(2,3)$. We know from the base case of our inductive proof that $n(k,2) = k+1$ we will take $n_0 = 2+1 = 3$, and recall that $q_0 = 1$. So now $q_1 = 2n_0q_0 = 2(3)(1) = 6$ and thus $n_1(2^6, 2) = 2^6 + 1$. Thus $q_2 = 2(2^6+1)(6) = 780$, and recall that we are taking q_k as our $n(k, l+1)$ number. Now 780 is a number that is much larger than necessary to get an arithmetic sequence of length 3 when considering only 2 classes, but it is large enough and recall that we constructed our q_k numbers with the foresight to use types to find our arithmetic sequences. So I will randomly partition a sequence of 780 numbers $(1, 2, 3, ..., 780)$ into two classes then find segments of length 6 of the same type in the left hand side of our 780 numbers. Using R I input the following line (1 : 780,rbinom(780, 1, .5)) to place each number with equal probability into class 0 or class 1. I then searched for segments of length 6 that were the same type. Here are three examples of segments of same type in arithmetic sequence of length 2 (simply stated: pairs of segments of the same type):

$$1 - \{10_1, 11_1, 12_1, 13_1, 14_1, 15_1\}, \{54_1, 55_1, 56_1, 57_1, 58_1, 59_1\}$$

$$2 - \{1_1, 2_0, 3_1, 4_0, 5_1, 6_0\}, \{82_1, 83_0, 84_1, 85_0, 86_1, 87_0\}$$

$$3 - \{264_0, 265_0, 266_0, 267_0, 268_0, 269_1\}, \{354_0, 355_0, 356_0, 357_0, 358_0, 359_1\}$$

The subscript of each number is the class that it was randomly selected to be in. The distance between the segments of 1 is 44, for 2 it's 81, and for 3 it's 90. So next I found the $l+1$ segment that was the same distance away from the segments we just found.

$$1 - \{98_0, 99_0, 100_0, 101_0, 102_1, 103_1\}$$

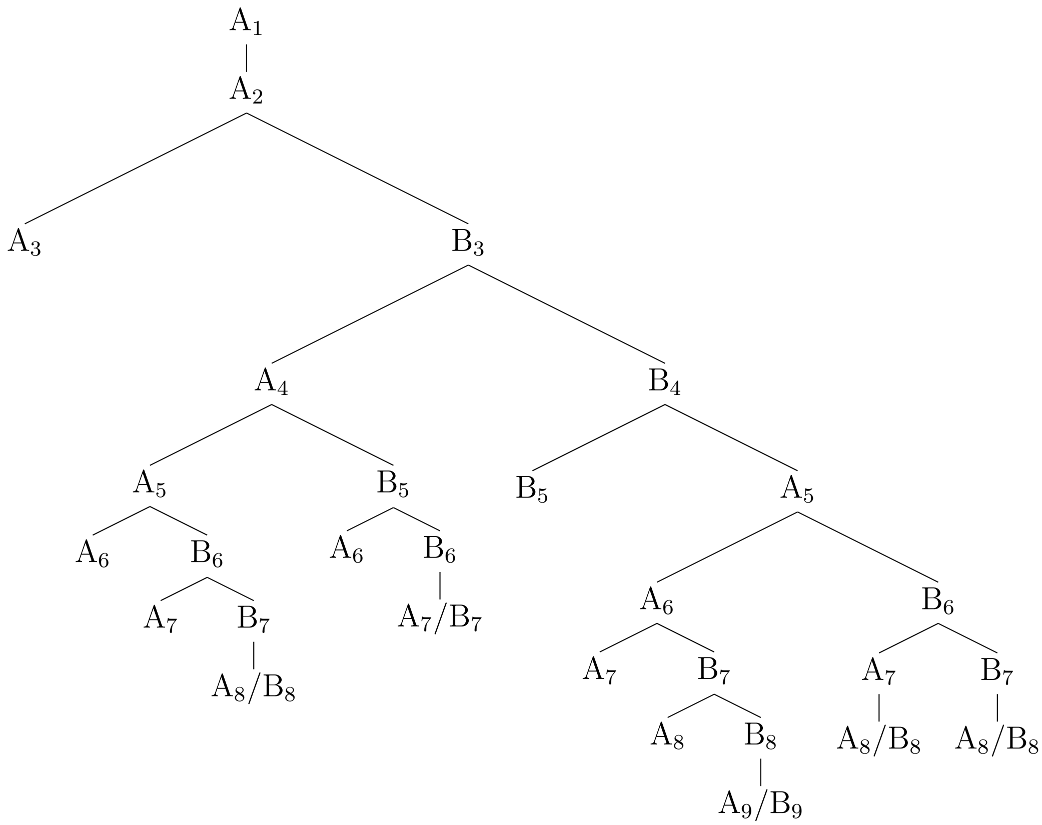
$$2 - \{163_1, 164_0, 165_0, 166_0, 167_0, 168_0\}$$

$$3 - \{444_0, 445_0, 456_1, 457_0, 448_1, 449_0\}$$

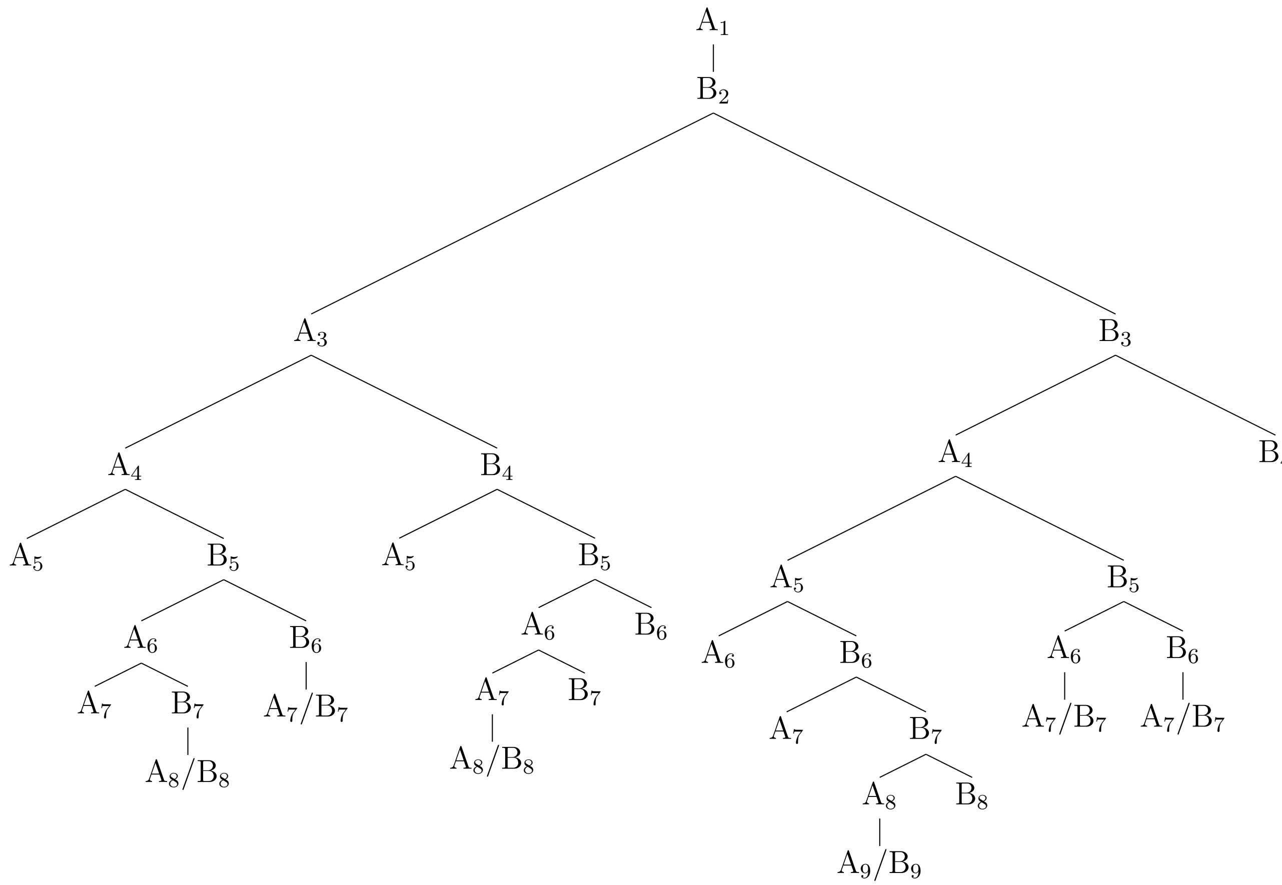
And we can see that we've found many arithmetic sequences of length 3 in the segments we've just found: $\{14_1, 58_1, 102_1\}$, $\{15_1, 59_1, 103_1\}$, $\{1_1, 82_1, 163_1\}$, $\{2_0, 83_0, 164_0\}$, $\{4_0, 85_0, 166_0\}$, $\{264_0, 354_0, 444_0\}$, $\{265_0, 355_0, 445_0\}$, and $\{267_0, 357_0, 447_0\}$. So it seems that our construction of $n(2,3)$ has given us a number far too large. Fortunately of the few known van der Waerden Numbers (the $n(k,l)$ values) the value of $n(2,3)$ is known to be 9 [7]. I will now provide my own proof of this fact.

One thing that we should notice before moving to prove (by construction or brute force) that $min(n(2,3)) = 9$, is the following, suppose we have an arithmetic sequence of classes, for example $S = \{A, \mathbf{A}, B, A, \mathbf{A}, B, B, \mathbf{A}\}$, where we have an arithmetic sequence in the class A of length 3 with distance 2 between the numbers. Then its reverse (taking the sequence of classes and reading it backwards) also contains an arithmetic sequence in class A of length 3 with distance 2 apart, as seen $S_r = \{\mathbf{A}, B, B, \mathbf{A}, A, B, \mathbf{A}, A\}$.

Now what we shall do is construct a tree of A's and B's where our goal is to make each branch as long as possible without containing an arithmetic progression of length 3. We will show that no branch can be longer than 9. We have two cases, depending on whether our tree begins with AA or AB. We are not considering BB and BA because the class labels can be switched from our AA and AB cases to get the same result. We will label each letter with a $\{1, 2, ..., 9\}$ to indicate the order of the sequence we are constructing, we will also make the last term an A, B, or A/B when there is an arithmetic sequence of length 3. So let us begin with the AA case:



So we only have one sequence of A's and B's that is of length 8 without an arithmetic sequence of length 3 in it. Specifically the sequence we found when starting with AA is $\{A, A, B, B, A, A, B, B, A/B\}$. Now let us construct a similar tree but starting with AB:



Here we also find just one sequence of length 9: $\{A, B, B, A, A, B, B, A, A/B\}$. Let us call the sequence of length 9 from the AA tree A_A, and the sequence of length 9 from the AB tree A_B. Also when considering the reverse of a sequence S we shall denote it S^r.

So now let us imagine that $n(2,3) \neq 9$, so let us assume that $n(2,3) > 9$. Then we should be able to find a sequence of 9 A's and B's that does not have an arithmetic sequence of length 3, such that when adding either an A or B to either end (because of symmetry it does not matter what end we add the tenth term) we get an arithmetic sequence of length 3. Now let us look at A_A and A_B more closely, we will take A_B^r in order to get the possibility of an AA start.

$$A_A = \{A, A, B, B, A, A, B, B, A/B\}$$
$$A_B^r = \{A/B, A, B, B, A, A, B, B, A\}$$

Notice that if we were to take the first and last element from both these sequences the sequences we have left are of the same type, which is $\{A, B, B, A, A, B, B\}$. Now since we are assuming that $n(2,3)$ is larger than 9 let us try to find a sequence of length 10. Now we know that if we add any two combinations of A and B to either end of our middle sequence we will find an arithmetic sequence of length 3 in a segment of length 9. So we can assume that if a sequence of A's and B's of length 10 existed that we could construct it by adding a combination of 3 A's and B's to the end of this middle. Though notice that if we were to place this combination of 3 A's and B's on the end of our middle we would get a sequence that starts with AB and we know that the the smallest length a sequence of A's and B's can be with an arithmetic sequence of length 3 in it with this beginning is 9. Similarly when we place any combination of 3 A's and B's on the beginning, by symmetry, we get a sequence that begins with AA and we know that this give us $n(2, 3) = 9$ again. This argument can be made for any 'middle' that does not have an arithmetic sequence of length 3 in it. You will find that in any sequence of A's and B's of length 10 there will be an arithmetic sequence of length 3 in the first or second string of 9 elements of the sequence, the element not belonging to this string can be deleted leaving us with a sequence of length 9.

3 Ergodic Proof of van der Waerden's Theorem

For this proof I will be following the proof given in H. Furstenberg's *Recurrence in Ergodic Theory and Combinatorial Number Theory* [5]. I will begin in the chapter titled van der Waerden's Theorem and move through it giving more examples and definitions. After completing section 2 I will illustrate the proof of why the Multiple Birkhoff Recurrence theorem implies van der Waerden's theorem in section 5 of the introduction.

I will begin by stating a few definitions that once understood will let us understand the proofs of some necessary lemmas.

Homeomorphism: A function $\mathbf{F}$ is a *homeomorphism* if it is (1) a bijection, which means it's both one to one and onto, (2) continuous, and (3) its inverse function, $\mathbf{F}^{-1}$ is also continuous.

Compact Metric Space: Metric Spaces are *compact* when every sequence in the space converges to a point. *Example:* The closed unit interval, $[0, 1]$ is a *compact metric space*. Can you see why? Because all sequences in the space can converge to the endpoints $0, 1$ or inside the interval. Thus when you read "compact metric space" think of the unit interval.

Group: A *group* is a set with an operation upon it. Call the set G and the operation be $\circ$ that can stand for multiplication, addition, etc. Then for $(G, \circ)$ to be a group they must satisfy these conditions:

(1) For all $a, b \in G$ we have $a \circ b \in G$. Known as closure.

(2) For $a, b, c \in G$ w have $(a \circ b) \circ c = a \circ (b \circ c)$. Known as associativity.

(3) There must exist an element $i \in G$ such that for any $a \in G$ we have $a \circ i = i \circ a = a$. i is called the identity element.

(4) And lastly for any $a \in G$ there must exist some $b \in G$ such that $a \circ b = b \circ a = i$. That is every element in G must have an inverse that takes it to the identity element [11].

Homomorphism: A *homomorphism* is a continuous map $\phi : X \to Y$ between two groups X and Y (with the same group operator G) that satisfies this condition: $\phi(gx) = g(\phi x)$.

Automorphism: An *automorphism* is a homomorphism from a group G to itself. So an automorphism preserves all the structure of the group.

Now we are ready to prove some lemmas that will build up to the Multiple Burkhoff Reccurence Theorem (theorem MBR). We will come across more important definitions later and will explain them when necessary.

Lemma 1. *Let T be a continuous map of a compact metric space X to itself ($T : X \to X$). Let $A \subset X$ with the property that for every $x \in A$, and $\epsilon > 0$, there exists* *a $y \in A$ and a $n \geq 1$ with $d(T^n y, x) < \epsilon$. Then for every $\epsilon > 0$, there exists a point $z \in A$ and an $n \geq 1$ with $d(T^n z, z) < \epsilon$. (To clarify when we say $d(a,b) < c$ it can be read; the distance between a and b is less than c).*

Proof: Let $\epsilon > 0$ be given. Set $\epsilon_1 = \dfrac{\epsilon}{2}$. We begin our proof by arbitrarily choosing a $z_0 \in A$ and select some $z_1 \in A$ with

$$d(T^{n_1} z_1, z_0) < \epsilon_1$$

for some n_1. Now select an ϵ_2 with $0 < \epsilon_2 \leq \epsilon_1$ such that when $d(z, z_1) < \epsilon_2$ we have

$$d(T^{n_1} z, z_0) < \epsilon_1.$$

This argument follows from the continuity of T. In the definition of continuity given any $\delta > 0$ there exists a $\eta > 0$ such that when we have two points $x_1, x_2 \in X$ such that $d(x_1, x_2) < \eta$ we know that $d(Tx_1, Tx_2) < \delta$. So our argument here is that we are finding the ϵ_2 such that when we take a point within ϵ_2 of z_1 we know by continuity that it must be within ϵ_1 of z_0. Now that we have found an ϵ_2 that satisfies this property, we will find a $z_2 \in A$ and $n_2 \geq 1$ that gives us

$$d(T^{n_2} z_2, z_1) < \epsilon_2.$$

Now let select an ϵ_3 with $0 < \epsilon_3 \leq \epsilon_1$ such that when $d(z, z_2) < \epsilon_3$ we have

$$d(T^{n_2} z, z_1) < \epsilon_2.$$

We will continue in defining $z_0, z_1, ..., z_k \in A$, $n_1, n_2, ..., n_k \in \mathbb{N}$, and $\epsilon_1, \epsilon_2, ..., \epsilon_k \in \left(0, \dfrac{\epsilon}{2}\right)$ in this manner. So we can simplify and say

$$d(T^{n_i} z_i, z_{i-1}) < \epsilon_i, i = 1, 2, ..., k$$

and select an ϵ_{k+1} with $0 < \epsilon_{k+1} \leq \epsilon_1$ such that

$$d(z, z_k) < \epsilon_{k+1} \Rightarrow d(T^{n_k} z, z_{k-1}) < \epsilon_k$$

We can then determine z_{k+1} and n_{k+1} to give

$$d(T^{n_{k+1}} z_{k+1}, z_k) < \epsilon_{k+1}.$$

Now let us step back and try to visualize what we have created. We began by finding an ϵ_2 such that when $z \in [z_1 - \epsilon_2, z_1 + \epsilon_2]$, which is what it means to have $d(z, z_1) < \epsilon_2$, then $T^{n_1} z \in [z_0 - \epsilon_1, z_0 + \epsilon_1]$. Continuing on we can see that we are finding segments in A such that when we take $T^{n_k + n_{k-1} + ... + n_2 + n_1}$ of some $z \in A$ with $d(z, z_k) < \epsilon_{k+1}$ we get $T^{n_k + n_{k-1} + ... + n_2 + n_1} z \in [z_0 - \epsilon_1, z_0 + \epsilon_1]$. This result comes from the more general statment that when $i < j$

$$d(T^{n_j + n_{j-1} + ... + n_{i+1}} z_j, z_i) < \epsilon_{i+1} \leq \epsilon_1 = \frac{\epsilon}{2}$$

Now since A is compact and we can continue splitting A into infinitely many subsegments (simply continue our process for $k \to \infty$) we can find some pair of $i, j, i < j$, for which $d(z_i, z_j) < \frac{\epsilon}{2}$. Another way to think about this fact is imagine if we split A into open segments of length $\frac{\epsilon}{2}$, then there must be a segment that contains more than one z_i, because A is a compact metric space. You can see this fact because A has infinitely many points in it and there are only finitely many segments covering A, thus if all the segments did not contain more than one z_i there would be infinitely many segments, but we know that this is not possible since a compact metric space can be covered by a finite number of open sets. Now if we combine this with our previous general result we get

$$d(T^{n_j+n_{j-1}+...+n_{i+1}}z_j, z_j) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

So what we've shown here is that when we find two elements of A in a segment that we've constructed (whose length is less than $\frac{\epsilon}{2}$) we know that when we take T^N (where $N = n_j + ... + n_i$, is the difference of the iterates of two elements we've found) of any element in this segment we will again be within ϵ_1 of the z_i element. But we know that z_i belongs in this segment, thus when we take $T^N z_i$ it will be in the interval $(z_i - \epsilon_1, z_i + \epsilon_1)$. Notice we have open brackets because the segment that we're taking our elements from has length $< \frac{\epsilon}{2}$. So we have proved Lemma 1.

LEMMA 2: *If A is a homogeneous set in the compact metric space X with repsect to a transformation T, and for any $\epsilon > 0$, we can find an $x, y \in A$ and an $n \geq 1$ so that $d(T^n y, x) < \epsilon$ then A is recurrent.*

To prove this we will first have to define what we mean by *homogeneous set* and a *recurrent set.*

HOMOGENEOUS SET: Let T be a continuous map of a compact metric space X to itself and let A be a closed subset of X. The set A is said to be homogeneous with respect to T if there exists a group G of homeomorphisms of X, each of which commutes with T (as in $T(g(x)) = g(T(x))$) and leaves A invariant (which means $g(A) = A$), and such that the dynamical system(A, G) is minimal.

A dynamical system is minimal if no closed subset of X is invariant under the action, which in our case is G. What this means is that there is no point $x\ inX$ such that for any $g \in G$ the orbit of x under g, $\{x, gx, g^2x, g^3x, ...\}$ doesn't come arbitrarily close to every point in X. This previous sentence is what it means for a set, in this case the orbit of x, to be dense in a space, which in this case is X. We can also deduce that a system (X, G) is minimal if for every open set $V \subset X$, there exist finitely many elements $g_i \in G$ with

$$\cup_{i=1}^{n} g_i^{-1} V = X.$$

If our system (X, G) wan not minimal then there would be some $Y \subset X$ that is invariant under G, now if we set $V = X - Y$ the equation above would not give us a point $x \in Y$, so $\cup_{i=1}^{n} g_i^{-1} V \neq X$. But this cannot happen by the first definition of minimal.

If we have a homeomorphism of X that commutes with T this homeomorphism is an automorphism of the system (X, T). From the definition of a homogeneous set we can see that these automorphisms of (X, G) will take any point in the homogeneous set A close to any other point in A.

Recurrent set: A closed subset A of a compact metric space X is said to be *recurrent* for a transformation $T : X \to X$, if for any $\epsilon > 0$ and any point $x \in A$, there exists a $y \in A$ and an $n \geq 1$ with $d(T^n y, x) < \epsilon$. Thus when a set is recurrent we know that for any point $x \in A$ and the transformation T there exists some other point in A who has an element in its forward orbit that comes within ϵ of x.

Now we can prove Lemma 2. *Proof:* Let G be a group of homeomorphisms that commute with T, leaving A invariant and so that the system (A, G) is minimal (here we are just stating what is necessary for A to be a homogeneous set). So if $g \in G$ and $a \in A$ we know, given T, that $T(g(a)) = g(T(a))$. Also we know that all the elements of G belong to a group and thus have group structure. Let the operation on G be function composition, it is then easy to see that G satisfies the group criteria.

Now for an $\epsilon > 0$ imagine a finite covering of A, call it $\{V_i\}$ and make it such that each V_i has diameter ϵ. Now since we're thinking of the unit interval as our compact metric space think of each of these V_i to be intervals with length ϵ. Now for each of these V_i we can find a finite set of elements in G, call the set $\{g_i\}$ (that stands for the $g \in G$ where the following property is satisfied for the V_i) such that

$$\cup g_i^{-1} V_i = A \tag{6}$$

Since $g \in G$ is a homeomorphism it is known that its inverse is also continuous. And since (A, G) is minimal we know that for any $a \in A$ and $g \in G$ the forward orbit of a with respect to g is dense in A. Thus the reverse orbit of a is also dense in A with respect to g. So it follows that for a given V_i there is a finite subset $G_0 \subset G$ that satisfy (6). So we can say that for any pair $x, y \in A$:

$$\min_{g \in G_0} d(gx, y) < \epsilon$$

Now find a $\delta > 0$ such that whenever $d(x_1, x_2) < \delta$ we have $d(gx_1, gx_2) < \epsilon$ for all $g \in G_0$. Now recall in Lemma 1 the final result was that for $x, y \in A$ and $n \geq 1$ we have $d(T^n y, x) < \epsilon$, now replace ϵ with δ to get $d(T^n y, x) < \delta$. Now by (6) we see that gx gets arbitrarily to at least one other element of A for any $g \in G_0$. So now select a $z \in A$ that is within ϵ of gx for a $g \in G_0$. So we have $d(gx, z) < \epsilon$. Now if we modify the result of Lemma 2 (that we can find an $x, y \in A$ and an $n \geq 1$ so that $d(T^n y, x) < \epsilon$) with δ, we have:

$$d(T^n y, x) < \delta$$

which means

$$d(gT^n y, gx) < \epsilon = d(T^n gy, gx) < \epsilon$$

and combine this with the fact that $d(z, gx) < \epsilon$ we have

$$d(T^n gy, z) < 2\epsilon.$$

What we did here is used the triangle inequality to show that if $d(a, b) < 2$ and $d(b, c) < 4$ then we must have $d(a, c) < 6$. Letting gy equal to the y used in the Lemma we have proved that A is recurrent.

Now with this notion of a reccurent set we can add this condition to Lemma 2 to get a new Lemma.

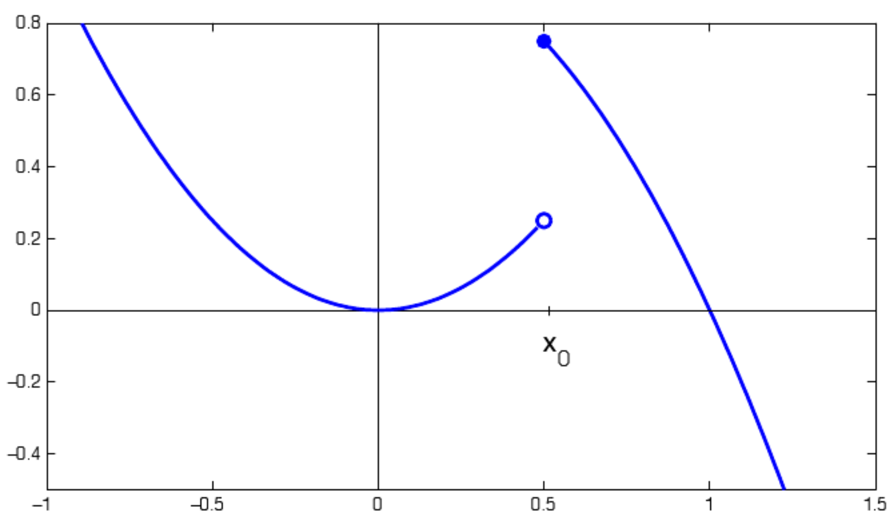
LEMMA 3: *Let A be a recurrent homogeneous set in X with respect to a transformation T. Then A contains a recurrent point for (X, T).*

In the proof of this Lemma we will have to consider the function:

$$F(x) = \inf_{n \geq 1} d(T^n x, x) \tag{7}$$

Let it be known that the function $F(x)$ is upper semicontinuous.

UPPER SEMICONTINUOUS: A function $F(x)$ is upper semicontinuous if given a x_0 there is a $\lambda > F(x_0)$, then there is a neighborhood U around x_0 such that $F(U) < \lambda$. This means that, if F is upper semicontinuous then, at all the points of discontinuity (where we have steps or jumps in the function) the closed part of the function belongs to the upper strand (at a point of discontinuity the filled dots will be above the empty dots on the function).



It is also known that a semicontinuous function's points of discontinuity lie in the union of countably many closed nowhere dense sets ([5], 39). An example of a nowhere dense set is the Cantor set, where every middle third of every interval is taken out. Since the Cantor set contains no intervals it must not have any intervals that have limit points.

Now we are ready to prove Lemma 3.

Proof: Since we have that A is a recurrent and homogeneous set in X with respect to T we know that Lemma 1 holds. This means that we can assume that $F(x)$ comes arbitrarily close to zero on A, since by Lemma 1 for every $\epsilon > 0$ we can find a $x \in A$ and an $n \geq 1$, such that $d(T^n x, x) < \epsilon$. Now since we know that (A, G) is minimal it holds that A is dense which implies that A must have points of continuity of $F(x)$ because if it didn't it would have to be a union of countably many nowhere dense sets. So select a point of continuity $x_0 \in A$.

Suppose that $F(x_0) > 0$, then select a δ such that $F(x) > \delta > 0$ in an open set $x_0 \in V \subset A$. Now since (A, G) is a minimal system we know by (6) that:

$$A \subset \cup_{g \in G_0} g^{-1} V,$$

where G_0 is a finite subset of G. Now let $\eta > 0$ be such that when $d(x_1, x_2) < \eta$ implies that $d(gx_1.gx_2) < \delta$ for all $g \in G_0$. We know that an $\eta > 0$ must exist for each $g \in G_0$ because g is a homoemorphism and they are continuous. The η we are using is the smallest for all g: the η selected works for all g. Then for every $x \in g^{-1}V$, we have $F(x) \geq \eta$. This must be the case because if $F(x) < \eta$, then we could find an n that satisfies $d(T^n x, x) < \eta$ which implies that $d(T^n gx, gx) < \delta$. This causes a problem because it means that $F(gx) < \delta$ for all $g \in G_0$, but this is impossible because for some $g \in G_0$ we have $gx \in V$ from by definition that $x \in g^{-1}V$. So it must be that $F(x) \geq \eta$ throughout A. If this was the case then we could set $\epsilon = \eta$ and then there would be no possible way to satisfy Lemma 1, which states that for any $\epsilon > 0$ there is a point $z \in A$ and an $n \geq q$ that satisfies $d(T^n z, z) < \epsilon = \eta$. Thus the assumption we made that $F(x_0) > 0$ must be wrong, so $F(x_0) = 0$, and x_0 is a recurrent point and we have proved Lemma 3. This does not state that all points in A are recurrent, only points in A that are points of continuity, and A is not required to be made up of only points of continuity. It should also be noted that a recurrent point returns close to itself and is not a periodic point that returns exactly to itself, thus the existence of recurrent points does not imply the existence of a periodic point.

Now we can put the three Lemmas that we have proved into one compact proposition:

PROPOSITION 1: *Let T be a continuous map of a compact metric space X to itself, and let $A \subset X$ be a homogeneous closed subset of X with respect to T. If for any $\epsilon > 0$, we can find an $x, y, \in A$ and a $n \geq 1$ such that $d(T^n y, x) < \epsilon$, then A contains a recurrent point for T.*

Our next step is to expand on the Lemmas that we have proved to get a stronger result, the Multiple Birkhoff Recurrence theorem (MBR). What will bring about MBR is taking the results we've proved and apply them to higher dimensions. To do this we will assume that $T_1, T_2, ..., T_l$ are continuous maps of our compact metric space X to itself. So our goal is to show that $(x, x, ..., x) \in X^l$ is recurrent for $T_1 \times T_2 \times ... \times T_l$. For simplicity call $T = T_1 \times T_2 \times ... \times T_l$ and call $x^l = (x, x, ..., x)$, and when we have Tx^l we get $(T_1x, T_2x, ..., T_lx)$.

PROPOSITION 2: *When we have a compact metric space X and $T_1, T_2, ..., T_l$ commuting homeomorphisms of X, then there exists a point $x \in X$ and a sequence $n_k \to \infty$ with $T_i^{n_k}x \to x$ simultaneously for $i = 1, 2, ..., l$.*

Before we begin the proof let me introduce the notion of uniform recurrence. A point x is said to be uniformly recurrent for the system (X, f) if for every open neighborhood U of x the set $\{n \geq 0 : f^n x \in U\}$ is syndetic. For a set to be syndetic there is a natural number N such that a block of N consecutive integers intersects the set. [1]

Proof: Let $T_1, T_2, ..., T_l$ be the generators for a group of homeomorphisms of X, call it G. For example a group generated by a, b, c is all the possible combinations of a, b, c under the group operator. In our case we have function composition as our group operator, thus the group G that we've created is infinite because we could compose T_1 with itself an infinite amount of times. Without loss of generality we will restrict our scope to any invariant subset of X under G but will write (X, G). So we have an invariant subset that is closed under the group of commuting homeomorphisms G, so this dynamical system that we are writing as (X, G) fulfills the requirements for it to be minimal. Now we will use induction to prove Proposition 2. In the $l = 1$ case we know the result by a theorem that states all points from a minimal closed T_1-invariant subset is uniformly recurrent ([5] 29). I will skip the proof of this theorem because our intuition after the first three Lemmas should lead us to believe this fact. Having satisfied the base case we will assume that the Proposition has been satistfied for $l-1$ transformations, so that we have x^{l-1} and $T' = T_1 \times T_2 \times ... \times T_{l-1}$ and $T^{n_k}x^{l-1} \to x^{l-1}$ as $n_k \to \infty$. Now consider l transformations, so $T = T_1 \times T_2 \times ... \times T_l$ and we will be dealing with x^l for any $x \in X$, notice that for $l = 2$ we are only dealing with values on the line $y = x$ since the y-axis is the space X like the x-axis. So our x^l we are interested in are on the diagonal for the l-dimensional space they belong to. Let us call $\Delta^{(l)}$ the subset of all diagonal l-tuples $(x, x, ..., x)$. Now let every element in $S \in G$ be written as $S \times S \times ... \times S$. Notice that since the T_i commute and generate G we know that T and $S \in G$ commute, also since T_i is a homeomorphism then any $S \in G$ is also a homeomorphism, so G is a group of homeomorphisms. And if you recall we are calling X our invariant set under G. It is easy to see that $\Delta^{(l)} \subset X^l$ and that $\Delta^{(l)}$ is a closed subset of X^l. Now we recognize that there must exist a set of $S \in G$ that leaves $\Delta^{(l)}$ invariant. So it follows that $\Delta^{(l)}$ is a homogeneous set. To finish the proof let $R_i = T_iT_l^{-1}, i = 1, 2, ..., l-1$, and let $y \in X$ be a point that satisfies $R_i^{n_k}y \to y$ for $i = 1, 2, ..., l-1$ and some sequence $n_k \to \infty$. Now set

$$x^* = (y, y, ..., y) \in \Delta^{(l)}$$

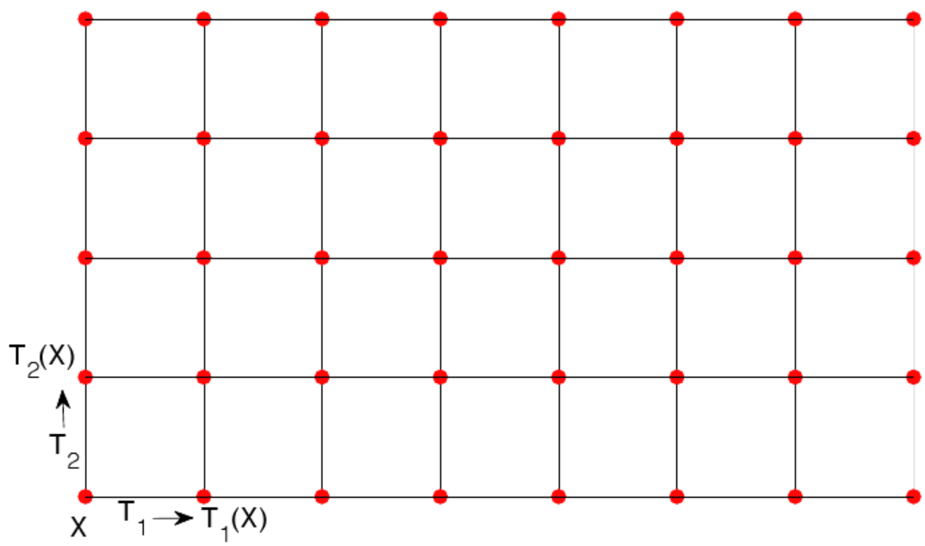
$$y_k^* = (T_l^{-nk}y, T_l^{-nk}y, ..., T_l^{-nk}y) \in \Delta^{(l)}$$

Then for large k, $T^{n_k}y_k^* = (R_1^{n_k}y, R_2^{n_k}y, ..., R_{l-1}^{n_k}y, y)$ and this must be close to x^*. We can draw this conclusion because by our inductive step we know that $(R_1^{n_k}y, ..., R_{l-1}^{n_k}y) \to (y, y, ..., y) \in \Delta^{(l-1)}$, thus $T^{n_k}y_k^* \to x^*$ as $n_k \to \infty$. So we have fulfilled Proposition 1 for l transformations, which means there is a recurrent point and we have proved Proposition 2.

With this result we can generalize the requirements and get a more broad result that will be MBR.

THEOREM MBR: *Let X be a compact metric space and $T_1, T_2, ..., T_l$ be commuting continuous maps of X to itself. Then there exists a point $x \in X$ and a sequence $n_k \to \infty$ with $T_i^{n_k}x \to x$ simultaneously for $i = 1, 2, ..., l$.*

Notice that we have T_i continuous maps rather than homeomorphisms, thus we cannot assume that the T_i have inverses. So the T_i do not generate a group, rather they generate a semigroup. Now to be able to use group properties consider $\mathbb{Z}^l$, where $\mathbb{Z}^2$ stands for only the integer points (including 0) of the space $\mathbb{R}^2$. In the diagram below we then apply a copy of the space X at each of the red points.



To keep things simple consider the $\mathbb{R}^2$ example where on the real line at every integer point we have a copy of our compact space (and recall that we are thinking of the unit interval as our compact space). In this example we have T_1 and T_2 as our continuous mappings and moving in the positive X direction (to the next integer value on the real line) means we have done T_1 to some $x \in X$ and moving in the positive Y direction is defined similarly. Call these moves S_i rather than T_i. Now notice that in the $\mathbb{R}^2$ case $(S_1 \times S_2)x = (S_2 \times S_1)x$ since we end up at the same point on the grid $(1, 1)$. Thus these S_i moves commute. Now all we must do to show that Proposition 2 holds is that there are inverses for the S_i.

Here I want to introduce some new notation, let $\omega(n_1, ..., n_l)$ be the representation of a point in X on grid with coordinates $(n_1, ..., n_l)$. So in our two dimensional grid above (where our $l = 2$) a $\omega(2, 1)$ would represent a $x \in X$ on the red circle $(2, 1)$ on the grid. Now let us compile this notation with our S_i moves. So when I write,

$$S_i\omega(n_1, n_2, ..., n_i, ..., n_l) = \omega(n_1, n_2, ..., n_{i+1}, ..., n_l)$$

for a certain $x_0 \in X$. Going back to our simple $l = 2$ case if we were given a point $x \in X$ represented by $\omega(1, 0)$, then $S_2\omega(1, 0) = \omega(1, 1)$.

Consider a space of all infinite sequences of numbers, call this space Y. Now call $\tilde{Y}$ the space of sequences that are forward orbits of points $x \in X$ satisfying,

$$S_i\omega(n_1, ..., n_l) = T_i(\omega(n_1, ...n_l)), i = 1, ..., l$$

Thus given a point $x_0 \in X$ we have a sequence $\tilde{y} \in \tilde{Y}$ that corresponds to the orbit of x_0. Let us go back to our $\mathbb{R}^2$ example, an orbit of a point can be written $\{(x_0, y_0), (Tx_0, y_0), (Tx_0, Ty_0), ...\}$. This orbit goes from the $(0, 0)$ lattice point to the $(1, 0)$ and then to the $(1, 1)$ lattice point. The second point in the orbit can be written as $S_1\omega(0, 0)$ with respect to the beginning point (x_0, y_0). Do not forget that whenever we have the notation $\omega(...)$ it is with respect to a single point in X. From here on whenever we say mention a point will be considering the full orbit of that point.

For any $x \in X$ and for $n \in N$ set

$$\omega_n(n_1, ..., n_l) = T_1^{n_1+n}T_2^{n_2+n}...T_l^{n_l+n}x$$

for $n_i \geq -n$. This is much simpler when explained in terms of our $\mathbb{R}^2$ example. So given a set $(4, 7)$ we can set our $n = \{-4, -3, -2, -1, -0, 1, ...\}$. Let $n = -3$, then we get $\omega_{-3}(4, 7) = (T_1^{4-3}x_0, T_2^{7-3}y_0) = (T^1x_0, T^4y_0)$. Again recall that we are doing this to the orbit of a point, thus essentially all we are doing is starting somewhere in the orbit and going back n points. Notice what this allows us to do, so long as our condition $n_i \geq -n$ is satisfied it lets us find the inverse of a point. Let me illustrate with an example. Consider the function

$$F(x) = 2x(mod 1).$$

$F(x)$ is continuous but not a homeomorphism. Consider the points $\frac{1}{4}$ and $\frac{3}{4}$. The orbits of these points go as follows

$$\left\{\frac{1}{4}, \frac{1}{2}, 0, 0, 0...\right\}$$

$$\left\{\frac{3}{4}, \frac{1}{2}, 0, 0, 0...\right\}.$$

Though notice that when considering the point $\dfrac{1}{4}$ that $\omega_{-1}\left(\dfrac{1}{2}\right) = \dfrac{1}{4}$. This is so because we are going back one iterate on the orbit of the point $\dfrac{1}{4}$ not the point $\dfrac{3}{4}$.

Now we see that we can take inverses with the operation S_i. So we have shown that the S_i commute and that they are invertible (again invertible here means going backwards in the orbit of the original point) thus these S_i are homeomorphisms. So now they satisfy the criteria for Proposition 2 and thus we have proved the theorem.

We are a few new terms away from showing you how this all applies to van der Waerden's theorem. I will present to you some symbolic dynamics that will prove very useful in completing the proof. Let $\Lambda = \{a, b, ..., q\}$ be a set of q symbols (try to think of them as symbols rather than letters in the alphabet because we can have more than 26 symbols). Now imagine all the infinite combinations of these q symbols, and call the set of these sequences $\bar{X}$. Now notice in one sequence $\bar{x_0} \in \bar{X}$ we have q partitions of the integers ($\mathbb{Z}$), you can see this by assigning one symbol to the number 0 then labeling the other symbols accordingly, then the numbers with the same symbols belong to the same class. Call these q sets B_i, then we have $\cup_{i=1}^{q} B_i = \mathbb{Z}$.

Next we must define a function to determine distance when considering the sequences in $\bar{X}$. Let $\bar{x}, \bar{x'} \in \bar{X}$, then

$$d(\bar{x}, \bar{x'}) = inf\left\{ \frac{1}{k+1} \middle| \bar{x}(i) = \bar{x'}(i) for |i| \leq k \right\}$$

What this is telling us is the size of the largest symmetrical interval around the central symbols of the sequences. Applying this, if we were given that $d(\bar{x}, \bar{x'}) = \frac{1}{5}$, we would know that if we went in either direction from the center symbol a distance of 4 symbols we would have identical 9 symbol intervals. Also when we have $d(\bar{x}, \bar{x'}) = 1$ we know that the two sequences have different central symbols, and if we have $d(\bar{x}, \bar{x'}) = \frac{1}{2}$ we know that only the central elements of each sequence correspond.

Our last item that we must introduce before showing how MBR implies van der Waerden's theorem is the shift map. Consider the sequence $\bar{x_0} = \{...bda.afgl...\}$ where the symbols come from $\Lambda = \{a, b, c, ..., q\}$, and the symbol to the right of the . represents the central symbol of the sequence. Now the shift map for a given $\bar{x} \in \bar{X} is$,

$$T(\bar{x}) = \{..., \lambda(i+1), ...\}$$

So the shift map on our previously mentioned, $\bar{x}$ gives us

$$T(\bar{x_0}) = \{...bd.aafgl...\}$$

means that we are moving the center of the whole sequence one point to the right, and then changing the integer values of all other points in the sequence accordingly. Notice that in the example that I have given we have $d(\bar{x}, T(\bar{x})) = \frac{1}{2}$ since the central terms of the sequence correspond. We can do the shift function multiple times to a sequence, where we get $\bar{x}, T(\bar{x}), T^2(\bar{x}), ..., T^n(\bar{x})$, where $T^n(\bar{x})$ means that we are shifting the integers that belong to the symbols and shifting them over n symbols.

Example: Let $\bar{x} = \{.....bdaa.cdbbadcc...\}$ where the symbol that comes after the . is the central symbol of the sequence. Now when we do $T^3(\bar{x})$ we get $\{...bdaacdb.badcc...\}$, and if we were to do $T^6(\bar{x})$ we get $\{...bdaacdbbad.cc...\}$.

What we are capable of doing now is writing van der Waerden's theorem using these new terms that we've defined. This will allow us to see how MBR can be used to imply van der Waerden's theorem. It is easy to see that this following assertion is the equivalent to the theorem: For any point $\bar{x_0} \in \bar{X}$ and any $l = 2, 3, ...$ there exists $m, n \in \mathbb{Z}$ such that $\bar{x_0}(m) = \bar{x_0}(m+n) = \bar{x_0}(m+2n) = ... = \bar{x_0}(m+ln)$. You can see how this is van der Waerden's theorem because we get an arithmetic sequence that has distance n between successive elements and that it is of length l. And we know that all the numbers belong in the same class because we have $\bar{x_0}(m) = \bar{x_0}(m+ln)$ for $l = 2, 3...$, which means that they are all the same symbol, which we took to mean class.

Now recall the distance function on our space $\bar{X}$. If we are given $d(\bar{x}, \bar{x'}) < 1$ then we know that we at least have $\bar{x}(0) = \bar{x'}(0)$, because the only way we can have $d(\bar{x}, \bar{x'}) < 1$ is if there is an identical interval of symbols about their $\bar{x}(0)$, and $\bar{x'}(0)$ points. Now notice that the shift map gives us $x(n) = T^n(x(0))$. Again we can rewrite van der Waerden's theorem:

If X is a compact metric space and $T : X \to X$ is a homeomorphism, then for an $x_0 \in X$ and any integer $l \geq l$ and a given $\epsilon > 0$, there is some point y on the orbit of x_0, which means we can write $y = T^m x_0$ for some m. And this y has the property that for some $n \geq 1$, the points $y, T^n y, T^{2n} y, ..., T^{ln} y$ are all within ϵ of each other.

This definition of van der Waerden's theorem follows from the MBR theorem we earlier proved. When proving this result we will use $\bar{X}$ as our space and the shift map T as our homeomorphism. To see how MBR implies this result let Y be the closure in $\bar{X}$ of the orbit of the point $\bar{x_0}$. This means that Y contains all the points on the orbit of $\bar{x_0}$ and it also contains all the points in $\bar{X}$ that the orbits of $\bar{x_0}$ get very close to but never become. Another way to say this is Y must also contain all the limit points of the orbit of $\bar{x_0}$. So we write:

$$Y = \overline{\{T^m \bar{x_0} | -\infty < m < \infty\}}.$$

Now let $T_1 = T, T_2 = T^2, ..., T_l = T^l$. We can see that the set Y is a T invariant compact set (where T represents all the possible combinations of the $T_i, i = 1, 2, ..., l$) because we know that no matter how many times you do T to $\bar{x_0}$ the resulting point must be in Y by our construction of Y. Also we must notice that the T_i are all commuting. This is not hard to see, imagine shifting a point m spaces, then shifting it n spaces, the result is a shift of $m+n$ spaces, which is obviously the same as a shift of $n+m$ spaces. So now we have commuting homeomorphisms on an invariant compact set Y, thus there exists a point $y' \in Y$ with some $T_1^{n_k} y' \to y', T_2^{n_k} y' \to y', ..., T_l^{n_k} y' \to y'$, by the MBR theorem. We can write our result in this manner, $y', T^n y', T^{2n} y', ..., T^{ln} y'$ are all within ϵ of each other. Now when we select an $0 < \epsilon < 1$ we know that the distance between points can at most be $\frac{1}{2}$ which means that all the points belong to the same class!

So what we have discovered is that given any sequence $\bar{x_0}$ (which is a partition of $\mathbb{Z}$ into q classes) then there is a point $q(i)$ on this sequence that for some $n \geq 1$ we have $q(i), T^n q(i), T^{2n} q(i), ..., T^{ln} q(i)$ all within ϵ of each other. And if this ϵ is less than 1 we know that all these values are the same symbol. Now since these symbols are shifted n spaces and the symbols acutally represent integers of $\mathbb{Z}$ we can see that we have found an arithmetic sequence of length l that is distance n apart and whose elements all belong to the same class. Furthermore since we found this arithmetic sequence for any $\bar{x_0} \in \bar{X}$ we know that we can find an arithmetic sequence of length l for *any* partitioning of $\mathbb{Z}$ into q classes.

You may notice that this result is slightly weaker because we cannot find an upper bound for how large of a set in $\mathbb{Z}$ we must use to ensure we find an arithmetic sequence of length l when partitioning into q classes.

4 Applications of van der Waerdens theorem

Unfortunately this theorem is not the most applicable result. One major goal ever since the theorem was proved was to find these van der Waerden numbers. The most minimal $n(k,l)$ that gives us an arithmetic progression of length l when splitting the set of numbers N with $1 \leq N \leq n(k,l)$. There have been a many discoveries of lower and upper bounds. Some examples of lower bounds for van der Waerden numbers are

$$n(l) > 2^{l-c(l\log l)^{\frac{1}{2}}}$$

where $n(l) = n(2,l)$ [4]. Another lower bound found is, given a prime p, [7]

$$n(2,p+1) > p2^p.$$

There have been many more lower bounds found for van der Waerden numbers.

One of the most recent discoveries with regards to van der Waerden numbers is the generalized result of the upper bound given by Tim Gowers in 1998, [7]

$$n(k,l) \leq e^{e^{k^{e^{e^{l+110}}}}}.$$

The trouble with these lower and upper bounds is that they are imprecise. It remains a very open problem in number theory as how to best approximate, or solve, $n(k,l)$.

One particularly interesting application of van der Waerdens theorem has been in finding prime numbers. Interestingly enough it was observed that $23143 + 30030l$ is a prime for $1 \leq l \leq 11$ ([4], pg. 224). The question that follows is can we find prime numbers in arithmetic progressions. It is known that there are an infinitely many prime triplets in arithmetic progression. It has been argued that if van der Waerden numbers we could be more precisely estimated then we could discover if there exists arbitrarily long strings of prime numbers that are consecutive members in some arithmetic progression [8].

Van der Waerden's theorem can also be relevant in explaining Bible Code. Bible Code is the idea that you can decifer hidden messages from the Bible, the main method in doing this is known as equidistant letter sequence method (ELS) [2]. To do ELS simply select a starting letter and then skip n letters, going forward or backwards, then at every n-th letter underline it (at some point you will have to stop the sequence). Then all the underlined letters can sometimes form a new word or sentence. As an example if I start at the **t** in my name Rothlisberger, and skip two letters going forward (stopping at the first **e**) we get Ro**t**hl**i**sb**e**rger. The bold letters spell out **tie**, so it can be said that my last name encodes the word tie with a forward skip of two. ELS is said to be more important when there are many encoded words found close together in the text, these words sometimes come together to form phrases which some believe can be predictors of the future.

Now it seems obvious that these ELS words are explained by van der Waerden's theorem. When doing ELS we are simply looking for arithmetic sequences that turn out to be words, rather than belonging to a class. So it follows that if we were to have a book with enough words (and all the letters of the alphabet) we would be garunteed to find many phrases, nevertheless words, by this process of ELS.

Van der Waerden's theorem is special because it can explain the Bible Code's findings. I believe that despite it's lack of mathematical applications (which is mainly because the generalized result of Szemeredi's theorem is more commonly used) the ability van der Waerden's theorem has to explain random patterns in the world makes it a very relevant theorem. As mathematics and science move to explain the world we live in I believe we will see more theorems, like van der Waerden's, that will describe the occurence of events we previously thought to be random.

5 Acknowledgments

I would like to thank Professor Ami Radunskaya. For her help in guiding me through difficult concepts and for her vast knowledge of LaTeXthat she gladly shared with me.

References

[1] Ethan Akin. *Recurrence in topological dynamics : Furstenberg families and Ellis actions*. Plenum Press, New York, 1997.

[2] Michael Drosnin. *The Bible Code*. Simon and Schuster, 1997.

[3] P. Erdos and P. Turan. On some sequences of integers. *Journal of the London Mathematical Society*, 1936.

[4] Paul Erdos. Some recent advances and current problems in number theory. *Lectures on Modern Mathematics*, 3:196–244, 1965.

[5] H. Furstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton University Press, Princeton, New Jersey, 1981.

[6] A. Y. Khinchin. *Three Pearls of Number Theory*. Graylock Press, Rochester, N.Y., 1952.

[7] Kevin O'Bryant and Eric W. Weisstein. *van der Waerden Number*. MathWorld–A Wolfram Web Resource., `http://mathworld.wolfram.com/vanderWaerdenNumber.html`, 2010.

[8] John R. Rabung and Van Der Waerden. On applications of van der waerden's theorem. *Mathematics Magazine*, pages 142–148, 1975.

[9] E. Szemeredi. On sets of integers containing no k elements in arithmetic progressions. *Acta Arithmetica*, 1975.

[10] B. L. van der Waerden. *How the Proof of Baudet's Conjecture Was Found*. Academic Press, London, 1971.

[11] S. Warner. *Topological Rings*. Elsevier Science Publishers, New York, 1993.