ABSTRACT

A Probabilistic Proof of the Vitali Covering Lemma

Ethan W. Gwaltney

Director: Paul Hagelstein, Ph.D.

The Vitali Covering Lemma states that, given a finite collection of balls in $\mathbb{R}^d$, there exists a disjoint subcollection that fills at least 3^{-d} of the measure of the union of the original collection. We present classical proofs of this lemma due to Banach and Garnett. Subsequently, we provide a new proof of this lemma that utilizes probabilistic "Erdös" type techniques and Padovan numbers.

APPROVED BY DIRECTOR OF HONORS THESIS:

____

Dr. Paul Hagelstein, Department of Mathematics

APPROVED BY THE HONORS PROGRAM:

____

Dr. Elizabeth Corey, Director

DATE: ______________________

A PROBABILISTIC PROOF OF THE VITALI COVERING LEMMA

A Thesis Submitted to the Faculty of

Baylor University

In Partial Fulfillment of the Requirements for the

Honors Program

By

Ethan Gwaltney

Waco, Texas

May 2017

TABLE OF CONTENTS

CHAPTER ONE

A Primer on Covering Lemmas

Two covering lemmas are helpful towards understanding the context of the problem we wish to address. The first, and simpler, is due to John B. Garnett, and gives a sharp bound for the proportion of a collection of intervals on the real line that can be covered by some subcollection of disjoint intervals. Here we state the theorem, give our version of the proof, and direct the reader to Garnett's book[1] for a more complete discussion.

Lemma 1 (Garnett Covering Lemma). *Let $\mathcal{I} = \{I_1, I_2, ..., I_n\}$ be a finite collection of bounded open intervals on the real line. Then there is a pairwise disjoint subcollection $\mathcal{I} \supseteq \mathcal{J} = \{J_1, J_2, ..., J_m\}$ such that*

$$\sum_{k=1}^{m} |J_k| \geq \frac{1}{2} \left| \bigcup_{k=1}^{n} I_k \right| .$$

Proof. First eliminate those intervals in $\mathcal{I}$ that are contained by other intervals in $\mathcal{I}$. Begin one subcollection by taking J_1 to be the leftmost interval in $\mathcal{I}$. If the right endpoint of J_1 does not intersect any interval in $\mathcal{I}$, then take K_1 to be the leftmost interval to the right of J_1. Otherwise, of all the inter-

[1]*Graduate Texts in Mathematics: Bounded Analytic Functions, Revised First Edition*, p. 24, John B. Garnett

vals in $\mathcal{I}$ intersecting the right endpoint of J_1, take K_1 to be the one with rightmost right endpoint. In either case, if the right endpoint of K_1 does not intersect any interval in $\mathcal{I}$, then take J_2 to be the leftmost interval to the right of K_1. Otherwise, of all the intervals in $\mathcal{I}$ intersecting K_1, take J_2 to be the one with rightmost right endpoint. Repeat this process, constructing two subcollections $\{J_i\}_{i=1}^{s}$ and $\{K_i\}_{i=1}^{t}$, until the entire collection is covered by the union of the new subcollections. Each of these subcollections is pairwise disjoint. Moreover, we have that

$$\sum_{i=1}^{s} |J_i| + \sum_{i=1}^{t} |K_i| \geq \left| \bigcup_{k=1}^{n} I_k \right| .$$

The subcollection of larger measure is sure to yield the desired inequality.

■

A more general result due to Banach is called the Vitali Covering Lemma. Here we will state a simplified version of the theorem and offer our version of the proof, as before. For more complete consideration, see Frank Jones's book's section[2] on the topic.

[2] *Lebesgue Integration on Euclidean Space, Revised Edition*, page 448, Frank Jones

Lemma 2 (Vitali Covering Lemma). *For* $\{C_i\}_{i=1}^n$ *a finite collection of balls in* $\mathbb{R}^d$*, there exists a pairwise disjoint subcollection* $\{D_i\}_{i=1}^m$ *such that*

$$\sum_{i=1}^m |D_i| \geq 3^{-d} \left| \bigcup_{i=1}^n C_i \right|.$$

Proof. Let $\{C_i\}$ be a collection of n balls in $\mathbb{R}^d$. Reindex $\{C_i\}$ such that $|C_1| \geq |C_2| \geq ... \geq |C_n|$. Then let $D_1 = C_1$. Now, suppose $\{D_1, D_2, ..., D_{k-1}\}$ have been chosen, for some $k \geq 2$. Let $D_k = C_l$, where $l = \min\{1 \leq j \leq n : C_j \cap \left(\bigcup_{i=1}^{k-1} D_i\right) = \emptyset\}$. Now consider a ball $C_j \notin \{D_i\}$. Consider why it is the case that C_j was not chosen in the disjoint collection. Clearly, C_j intersects some ball $D_l \in \{D_i\}$ that is larger than or equal to C_j in measure, by the construction of $\{D_i\}$. If rad B denotes the length of the radius of a ball B, then rad $C_j \leq$ rad D_l. By the triangle inequality, the distance from the center of D_l to any point in C_j is less than or equal to 3 rad D_l, and we have that $\bigcup_{i=1}^n C_i \subset \bigcup_{i=1}^m (3D_i)$. Hence,

$$\begin{aligned} \left| \bigcup_{i=1}^m 3D_i \right| \geq \left| \bigcup_{i=1}^n C_i \right| &\implies 3^d \left| \bigcup_{i=1}^m D_i \right| \geq \left| \bigcup_{i=1}^n C_i \right| \\ &\implies \sum_{i=1}^m |D_i| = \left| \bigcup_{i=1}^m D_i \right| \geq 3^{-d} \left| \bigcup_{i=1}^n C_i \right|. \end{aligned}$$

$\blacksquare$

CHAPTER TWO

A Simple Probabilistic Covering Theorem

Given a bounded subset $C \subset \mathbb{R}^d$ composed of a collection of balls, we wish to show that there exists a subcollection $S \subseteq C$ of pairwise disjoint balls such that the measure of this subcollection is greater than or equal to some constant proportion of the measure of C. The goal is the same as that in the above lemmas; our approach will differ significantly.

The proof of the Garnett Covering Lemma yields a proportion of $1/2$, but this result is unique to one-dimensional space, since it relies on the ordering of $\mathbb{R}$. The proof of the Vitali Covering Lemma uses the geometry of the set $C \subset \mathbb{R}^d$ and the triangle inequality to yield a proportion of 3^{-d}. These results are useful, but inadequate in certain cases of interest. We are interested in a probabilistic approach to the same line of inquiry. As a toy problem, we simplify the set C to the collection $\{[0, \frac{1}{n}], [\frac{1}{n}, \frac{2}{n}], \ldots, [\frac{n-1}{n}, 1]\}$. Note that the goal in introducing the following theorem is to demonstrate a result comparable to the above two lemmas by probabilistic methods (Garnett's method, for example, easily improves upon this result, but is not probabilistic in nature).

Theorem 1. *Let* $C = \{[0, \frac{1}{n}], [\frac{1}{n}, \frac{2}{n}], \ldots, [\frac{n-1}{n}, 1]\}$ *for some* $n < \infty$*. Then for sufficiently large* n*, there exists a pairwise disjoint subcollection* $S \subseteq C$ *such that* $|\bigcup_{I \in S} I| \geq \frac{1}{6}|\bigcup_{I \in C} I|$.

Proof. We first define K_n to be the number of disjoint subcollections of C and J_n to be the number of subcollections with measure greater than or equal to $\frac{1}{6}$. By the Pigeonhole Principle, what we wish to show reduces to the following inequality:

$$\frac{K_n}{2^n} + \frac{J_n}{2^n} \geq 1, \text{ or equivalently, } K_n + J_n \geq 2^n, \tag{1}$$

since 2^n is the number of subcollections of C. Notice K_n is the $(n+2)^{nd}$ member of the Fibonacci sequence, which is given by the recursive formula

$$F_0 = 0,\ F_1 = 1,\ F_n = F_{n-1} + F_{n-2},$$

and can be derived via generating functions in closed form. As such, we can also derive the closed form for K_n, as follows:

$$K_n = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^{n+2} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+2}}{\sqrt{5}}.$$

We can write J_n in terms of combinations as follows:

$$J_n = \sum_{k \geq \frac{n}{6}}^{n} \binom{n}{k} = \sum_{k=\lceil \frac{n}{6} \rceil}^{n} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{k} - \sum_{k=0}^{\lceil \frac{n}{6} \rceil - 1} \binom{n}{k} = 2^n - \sum_{k=0}^{\lceil \frac{n}{6} \rceil - 1} \binom{n}{k}.$$

We can rewrite (1) in terms of these forms for K_n and J_n to arrive at:

$$\frac{\left(\frac{1+\sqrt{5}}{2}\right)^{n+2} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+2}}{\sqrt{5}} - \sum_{k=0}^{\lceil \frac{n}{6} \rceil - 1} \binom{n}{k} \geq 0.$$

Thus we need to show

$$\frac{(1+\sqrt{5})^{n+2} - (1-\sqrt{5})^{n+2}}{4\sqrt{5} \cdot 2^n} \geq \sum_{k=0}^{\lceil \frac{n}{6} \rceil - 1} \binom{n}{k}.$$

Working with the ceiling function is unwieldy in computation, so instead we consider $n = 6m$ for some positive integer m, and rewrite (1) as

$$\frac{(1+\sqrt{5})^{6m+2} - (1-\sqrt{5})^{6m+2}}{4\sqrt{5} \cdot 2^{6m}} \geq \sum_{k=0}^{m-1} \binom{6m}{k}. \tag{2}$$

Now, for large n, this left hand side is dominated by the term

$$O\left(\frac{(1+\sqrt{5})^{6m+2} - (1-\sqrt{5})^{6m+2}}{4\sqrt{5} \cdot 2^{6m}}\right) = \left(\frac{1+\sqrt{5}}{2}\right)^{6m}.$$

The right hand side is dominated by the term

$$O\left(\sum_{k=0}^{m-1}\binom{6m}{k}\right) = m\binom{6m}{m} = \frac{m(6m)!}{m!(5m)!}.$$

We may approximate this term by Sterling's formula in the following way:

$$\begin{aligned} m\binom{6m}{m} = \frac{m(6m)!}{m!(5m)!} &\approx \frac{m\sqrt{12\pi m}\left(\frac{6m}{e}\right)^{6m}}{\sqrt{2\pi m}\left(\frac{m}{e}\right)^{m}\sqrt{10\pi m}\left(\frac{5m}{e}\right)^{5m}} \\ &= m\sqrt{\frac{3}{5\pi m}}\left(\frac{6^6}{5^5}\right)^m \leq \sqrt{m}\left(\frac{6^6}{5^5}\right)^m. \end{aligned}$$

Then since

$$\left(\left(\frac{1+\sqrt{5}}{2}\right)^6\right)^m > \sqrt{m}\left(\frac{6^6}{5^5}\right)^m,$$

as $m \to \infty$, (2) holds. ■

Now that we have illustrated how probabilistic methods can be helpful in proving covering lemmas, we are ready to introduce our main theorem.

CHAPTER THREE

A More General Result

We now extend the result from the simple situation of intervals of equal length covering $[0,1]$ to the more general situation of intervals of arbitrary length covering $[0,1]$. For such a given collection C of intervals the desired subcollection S must satisfy the same two criteria as before. First, the intervals in S must be pairwise disjoint. Second, S must fill some constant proportion of $[0,1]$. The following theorem summarizes the desired result:

Theorem 2. *Let C be a finite collection of intervals. Then there exists a pairwise disjoint subcollection $S \subseteq C$ such that $|\bigcup_{I\in S} I| \geq \frac{1}{4}|\bigcup_{I\in C} I|$.*

We first notice that the probability that the elements of a randomly selected subcollection of C will be pairwise disjoint does not differ from the simplified situation of Theorem 1. If C is composed of n intervals, then the probability a randomly selected subcollection is disjoint is the number of possible disjoint subcollections, which is given by the $(n+2)^{nd}$ Fibonacci number, divided by the total number of possible subcollections, which is just 2^n. However, as we are really interested only in subcollections yielding the largest possible proportion of the measure of the union of intervals in C, let us instead calculate the probability that a randomly selected subcollection of

C is *maximally* disjoint, that is, the subcollection is disjoint and should any additional interval be added to this subcollection, the subcollection would no longer be disjoint. For example, consider the following figures:

Figure A

Figure B

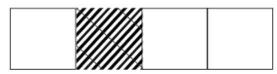
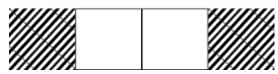


Figure C



Clearly, the subcollection in Figure A is not disjoint. The subcollection in Figure B is disjoint, but not maximally disjoint, as the fourth interval may be selected without violating disjointness. The subcollection in Figure C is maximally disjoint, since it is disjoint, and no other intervals may be added to the subcollection without violating disjointness. Now we claim that the number of maximally disjoint subcollections of C is the $(n+1)^{st}$ Padovan number, where the Padovan numbers are defined recursively as

$$P_0 = 1,\ P_1 = 1,\ P_2 = 1, \text{ and } P_k = P_{k-2} + P_{k-3}.$$

To see this, first let Q_n be the number of maximally disjoint subcollections of C containing the n^{th} interval. Recall that C has n intervals. Let R_n

be the number of maximally disjoint subcollections of C. Now, consider Q_n. If a subcollection of C contains the n^{th} interval, then in order to be disjoint it cannot contain the $(n-1)^{st}$ interval. In order to be *maximally* disjoint, either the $(n-2)^{nd}$ interval or the $(n-3)^{rd}$ interval must also be selected. Thus the total number of maximally disjoint subcollections of C containing the n^{th} interval reduces to the sum of the number of maximally disjoint subcollections containing the $(n-2)^{nd}$ interval and the number containing the $(n-3)^{rd}$ interval. Thus we arrive at the recursion formula

$$Q_n = Q_{n-2} + Q_{n-3}.$$

Moreover, for the cases $n = 1, 2, 3$, we have that $Q_n = 1, 1, 1$ respectively. Comparing this recursion to that given for the Padovan numbers shows that $Q_n = P_{n-1}$. Now let us turn our attention to R_n. Any maximally disjoint subcollection of C will contain either the n^{th} interval or the $(n-1)^{st}$ interval, but not both. Thus the number of maximally disjoint subcollections of C is the sum of Q_n and Q_{n-1}. Then, since $Q_n = P_{n-1}$, we have that

$$R_n = P_{n-1} + P_{n-2} = P_{n+1}.$$

The Padovan numbers are less well known than the Fibonacci numbers, but behave in similar fashion. For instance, it is easy to arrive at an exponential growth rate for the Padovan numbers in much the same way as the growth rate of the Fibonacci numbers is derived, namely, if $P_n = \lambda^n$ is a solution of the linear recursion $P_n = P_{n-2} + P_{n-3}$ for some λ, then

$$P_n = \lambda^n = \lambda^{n-2} + \lambda^{n-3} \implies \lambda^3 - \lambda - 1 = 0,$$

which we solve numerically. Thus we can show in a fashion similar to the treatment of the Fibonacci numbers that there exist constants $0 < c < C$ such that $c\lambda^n \leq P_n \leq C\lambda^n$, where λ is the positive solution to the above cubic. The other two solutions of the cubic equation $\lambda^3 - \lambda - 1 = 0$ are complex of magnitude less than one, and will not contribute to the asymptotic behavior of P_n.

Notice that the conditional probability that the k^{th} interval in the collection of n intervals, $I_{k,n}$, is chosen as a member of a maximally disjoint subcollection is the ratio of the number of maximally disjoint subcollections of intervals including $I_{k,n}$ and the total number of maximally disjoint subcollections. To find the numerator of the ratio, we simply multiply the number of maximally disjoint subcollections of $\{I_{1,n}, I_{2,n}, \ldots, I_{k-2,n}\}$ with the number of maximally disjoint subcollections of $\{I_{k+2,n}, I_{k+3,n}, \ldots, I_{n,n}\}$, and we have that

$$Pr(I_{k,n}) = \frac{P_{n-k} \cdot P_{k-1}}{P_{n+1}} \geq \frac{c^2 \lambda^{n-k} \cdot \lambda^{k-1}}{C\lambda^{n+1}} = \frac{c^2}{C\lambda^2}. \tag{3}$$

Thus for any interval in a collection of sufficiently many intervals, the probability of randomly selecting that interval as a member of a maximally disjoint subcollection is approximately $\frac{c^2}{C\lambda^2} > 0$. By an argument from the linearity of expectation that we explain in the conclusion to this manuscript, this is enough to show that, for sufficiently large n, there is a disjoint subcollection filling a $\frac{c^2}{C\lambda^2}$ proportion of the original collection. This is not too impressive, as c and C are both unknown. However, it demonstrates that a probabilistic proof of the sort we are interested in is possible.

CHAPTER FOUR

Non-asymptotic Bounds

The above argument yields an asymptotic result. We now want to show that for any collection of intervals, there exists a subcollection satisfying both the proportionality and disjointness conditions. To do this, we derive the generating function for the Padovan numbers in order to obtain a precise lower bound on the probability of, given an arbitrary interval in the collection, randomly selecting that interval as a member of a maximally disjoint set. We first recall the recursive relation defining the Padovan number P_n:

$$P_0 = P_1 = P_2 = 1, \;\; P_n = P_{n-2} + P_{n-3}.$$

Now we define the generating function, $P(x)$, to be the formal power series the coefficients of which are the Padovan numbers, as follows:

$$P(x) = \sum_{n=0}^{\infty} P_n x^n,$$

which yields

$$P(x) = 1 + x + x^2 + \sum_{n=3}^{\infty} P_n x^n.$$

Revisiting the recursive relation above, we replace the n^{th} Padovan number with $P_{n-2} + P_{n-3}$, yielding

$$\begin{aligned}
P(x) &= 1 + x + x^2 + \sum_{n=3}^{\infty} (P_{n-2} + P_{n-3})\, x^n \\
&= 1 + x + x^2 + \sum_{n=3}^{\infty} P_{n-2}x^n + \sum_{n=3}^{\infty} P_{n-3}x^n \\
&= 1 + x + x^2 + x^2 \sum_{n=3}^{\infty} P_{n-2}x^{n-2} + x^3 \sum_{n=3}^{\infty} P_{n-3}x^{n-3} \\
&= 1 + x + x^2 + x^2 \sum_{n=1}^{\infty} P_n x^n + x^3 \sum_{n=0}^{\infty} P_n x^n \\
&= 1 + x + x^2 + x^2 \left[\left(\sum_{n=0}^{\infty} P_n x^n \right) - 1 \right] + x^3 \sum_{n=0}^{\infty} P_n x^n \\
&= 1 + x + x^2 P(x) + x^3 P(x).
\end{aligned}$$

Solving for $P(x)$, we have

$$P(x) = \frac{1 + x}{1 - x^2 - x^3},$$

the generating function for the Padovan numbers. We now proceed to find the partial fraction decomposition of this generating function. The first step to this end is factoring the denominator. We factor by solving the equation

$$1 - x^2 - x^3 = 0,$$

by the method[1] due to Girolamo Cardano in the sixteenth century. As a first step, we depress this cubic polynomial using the method due to another sixteenth century mathematical icon, Niccolo Fontana, more famously known as Tartaglia. Accordingly, we use the substitution $x = y - \frac{1}{3}$. Then we have

$$1 - \left(y - \frac{1}{3}\right)^2 - \left(y - \frac{1}{3}\right)^3 = 0.$$

This leads to convenient cancellation, and yields the depressed cubic

$$y^3 - \frac{1}{3}y - \frac{25}{27} = 0.$$

The next step is due to del Ferro. We consider the system of equations

$$3st = -\frac{1}{3},\ s^3 - t^3 = -\frac{25}{27},$$

and note that $y = s - t$ gives the solution of the depressed cubic. Solving,

[1]The history of the general solution to the cubic equation is long and fascinating. For an effective summary of this history, as well as an interesting perspective on the solution, see William Dunham's *Journey through Genius: The Great Theorems of Mathematics*.

we have that

$$s = \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}},$$
$$t = -\sqrt[3]{\frac{25}{54} - \sqrt{\frac{23}{108}}},$$

and thus the real solution to the cubic equation $1 - x^2 - x^3 = 0$ is

$$x = \alpha := \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}} + \sqrt[3]{\frac{25}{54} - \sqrt{\frac{23}{108}}} - \frac{1}{3} \approx 0.75488.$$

By long division of the original denominator of the generating function by the factor $(x - \alpha)$, we have the following form for the generating function:

$$P(x) = \frac{1 + x}{(x - \alpha)(-x^2 - (1 + \alpha)x - (1 + \alpha)\alpha)}.$$

We may further factor this denominator using the quadratic formula, such that our fully factored quotient is as follows:

$$P(x) = -\frac{1 + x}{(x - \alpha)(x - \beta)(x - \gamma)},$$

where $\beta = -\frac{1+\alpha+i\sqrt{4\alpha(1+\alpha)-(1+\alpha)^2}}{2}$, and γ is the complex conjugate of β.

Now, by partial fraction decomposition, we set

$$P(x) = -\frac{1+x}{(x-\alpha)(x-\beta)(x-\gamma)} = -\left(\frac{A}{x-\alpha} + \frac{B}{x-\beta} + \frac{C}{x-\gamma}\right) \quad (4)$$

and solve for the constants A, B, and C by simply inputting $x = \alpha$, $x = \beta$, and $x = \gamma$ successively. This easily yields

$$A = \frac{1+\alpha}{(\alpha-\beta)(\alpha-\gamma)}$$
$$B = \frac{1+\beta}{(\beta-\alpha)(\beta-\gamma)}$$
$$C = \frac{1+\gamma}{(\gamma-\alpha)(\gamma-\beta)}.$$

We may manipulate each quotient from (4) into a constant multiple of a geometric series. For instance,

$$\begin{aligned}\frac{A}{x-\alpha} &= \frac{\alpha\frac{A}{\alpha}}{x-\alpha} = \frac{A}{\alpha}\left(\frac{\alpha}{x-\alpha}\right) = \frac{A}{\alpha}\left(\frac{1}{\frac{x}{\alpha}-1}\right) \\ &= -\frac{A}{\alpha}\left(\frac{1}{1-\frac{x}{\alpha}}\right) = -\frac{A}{\alpha}\sum_{n=0}^{\infty}\alpha^{-n}x^n.\end{aligned}$$

Similarly, we have that

$$\frac{B}{x-\beta} = -\frac{B}{\beta}\sum_{n=0}^{\infty}\beta^{-n}x^n, \text{ and } \frac{C}{x-\gamma} = -\frac{C}{\gamma}\sum_{n=0}^{\infty}\gamma^{-n}x^n.$$

Combining these terms, we have the following expression for $P(x)$:

$$\begin{aligned} P(x) &= -\left(-\frac{A}{\alpha}\sum_{n=0}^{\infty}\alpha^{-n}x^n - \frac{B}{\beta}\sum_{n=0}^{\infty}\beta^{-n}x^n - \frac{C}{\gamma}\sum_{n=0}^{\infty}\gamma^{-n}x^n\right) \\ &= \sum_{n=0}^{\infty}\left(\frac{A}{\alpha^{n+1}} + \frac{B}{\beta^{n+1}} + \frac{C}{\gamma^{n+1}}\right)x^n. \end{aligned}$$

Then, since earlier we defined $P(x) = \sum_{n=0}^{\infty} P_n x^n$, we have a closed form for the n^{th} Padovan number, namely,

$$P_n = \frac{A}{\alpha^{n+1}} + \frac{B}{\beta^{n+1}} + \frac{C}{\gamma^{n+1}}, \tag{5}$$

where the constants A, B, C, α, β, and γ are defined above. With this closed form for the Padovan numbers in hand, we want to lower bound the quantity from (3). Straightforward optimization is unwieldy. Instead, we may bound the Padovan numbers themselves above and below for $k \geq 4$ as follows:

$$\frac{L}{\alpha^{k+1}} \leq P_k = \frac{A}{\alpha^{k+1}} + \frac{B}{\beta^{k+1}} + \frac{C}{\gamma^{k+1}} \leq \frac{U}{\alpha^{k+1}}, \tag{6}$$

where L and U are constants. We now show by induction that (6) is satisfied by the constants $L = A + B\left(\frac{\alpha}{\beta}\right)^5 + C\left(\frac{\alpha}{\gamma}\right)^5$ and $U = A + B\left(\frac{\alpha}{\beta}\right)^7 + C\left(\frac{\alpha}{\gamma}\right)^7$

(to give some context, note that $L \approx 0.49$ and $U \approx 0.56$). First, we have

$$\begin{aligned}
&\frac{L}{\alpha^5} = \frac{A}{\alpha^5} + \frac{B}{\beta^5} + \frac{C}{\gamma^5} = 2 = P_4, \\
&\frac{L}{\alpha^6} \approx 2.649 \leq 3 = P_5, \text{ and} \\
&\frac{L}{\alpha^7} \approx 3.510 \leq 4 = P_6.
\end{aligned}$$

Now, suppose

$$\frac{L}{\alpha^{n-2}} \leq P_{n-3}, \; \frac{L}{\alpha^{n-1}} \leq P_{n-2}, \text{ and } \frac{L}{\alpha^n} \leq P_{n-1}.$$

Then, it is enough to show that $\frac{L}{\alpha^{n+1}} \leq P_n$, which follows quickly from the recursion formula for the Padovan numbers and the fact that α is the real solution to the equation $1 - x^2 - x^3 = 0$:

$$\begin{aligned}
1 - \alpha^2 - \alpha^3 = 0 &\Longleftrightarrow \frac{1}{\alpha^{n+1}} - \frac{1}{\alpha^{n-1}} - \frac{1}{\alpha^{n-2}} = 0 \\
&\Longleftrightarrow \frac{L}{\alpha^{n+1}} = \frac{L}{\alpha^{n-1}} + \frac{L}{\alpha^{n-2}} \leq P_{n-2} + P_{n-3} = P_n.
\end{aligned}$$

A similar argument demonstrates that U satisfies (6). Now, for $5 \leq k \leq n-4$, we may revisit the function from (3) and bound in the following way:

$$Pr(I_{k,n}) = \frac{P_{n-k} \cdot P_{k-1}}{P_{n+1}} \geq \frac{\frac{L}{\alpha^{n-k+1}} \cdot \frac{L}{\alpha^k}}{\frac{U}{\alpha^{n+2}}} = \frac{L^2\alpha}{U} \approx 0.3247.$$

Thus it only remains to find the minimum probability for $1 \leq k \leq 4$, by the symmetry of (3). This simplifies the optimization problem significantly. What was once a multivariate optimization over n and k is now optimization over a single variable n. We may iterate the same argument as above to find an interval on which the probability with fixed k is bounded below. Recalling the inequality from (6), we have that

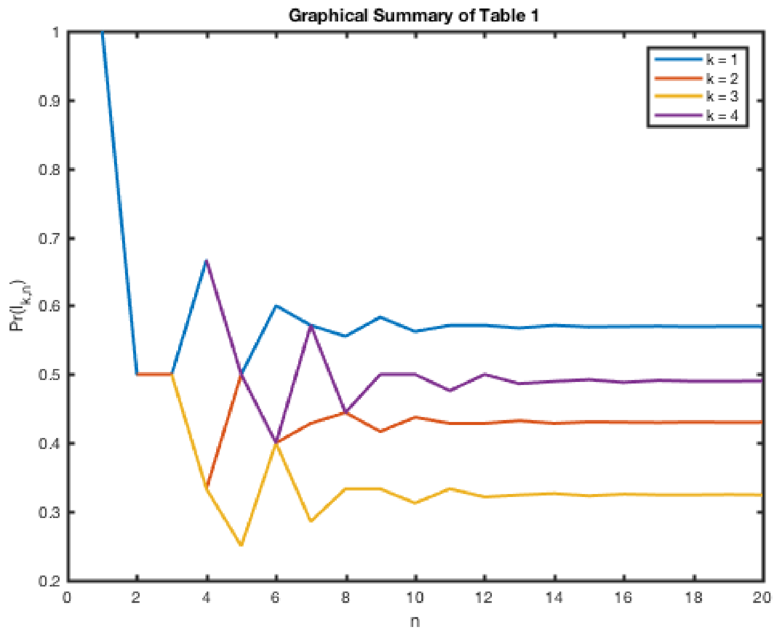
$$\begin{aligned}
Pr(I_{1,n}) &= \frac{P_{n-1} \cdot P_0}{P_{n+1}} = \frac{P_{n-1}}{P_{n+1}} \geq \frac{\frac{L}{\alpha^n}}{\frac{U}{\alpha^{n+2}}} = \frac{L\alpha^2}{U} \approx 0.5000, \text{ for } n \geq 5, \\
Pr(I_{2,n}) &= \frac{P_{n-2} \cdot P_1}{P_{n+1}} = \frac{P_{n-2}}{P_{n+1}} \geq \frac{\frac{L}{\alpha^{n-1}}}{\frac{U}{\alpha^{n+2}}} = \frac{L\alpha^3}{U} \approx 0.3774, \text{ for } n \geq 6, \\
Pr(I_{3,n}) &= \frac{P_{n-3} \cdot P_2}{P_{n+1}} = \frac{P_{n-3}}{P_{n+1}} \geq \frac{\frac{L}{\alpha^{n-2}}}{\frac{U}{\alpha^{n+2}}} = \frac{L\alpha^4}{U} \approx 0.2849, \text{ for } n \geq 7, \text{ and} \\
Pr(I_{4,n}) &= \frac{P_{n-4} \cdot P_3}{P_{n+1}} = \frac{2P_{n-4}}{P_{n+1}} \geq \frac{\frac{2L}{\alpha^{n-3}}}{\frac{U}{\alpha^{n+2}}} = \frac{2L\alpha^5}{U} \approx 0.4302, \text{ for } n \geq 8.
\end{aligned}$$

This leaves at most 4 + 4 + 4 + 4 = 16 values for n left to check (since $n \geq k$), which are computed in the following table.

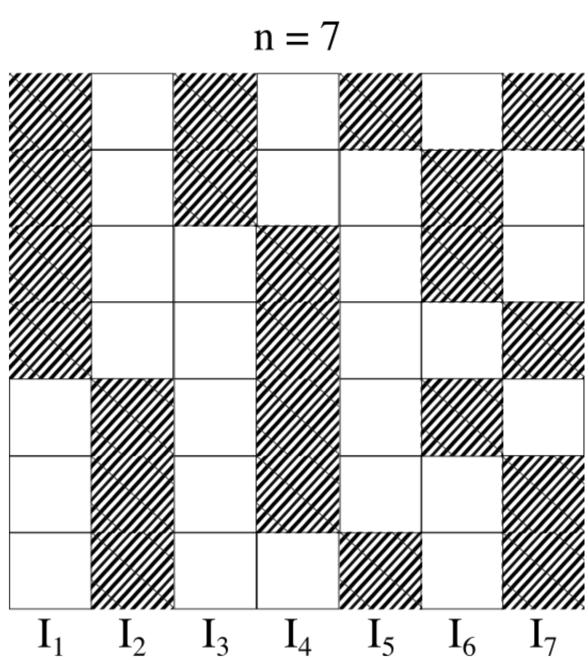
Table 1: $Pr(I_{k,n})$ for small n, k

$Pr(I_{k,n})$	k = 1	k = 2	k = 3	k = 4
n = 1	1			
n = 2	1/2	1/2		
n = 3	1/2	1/2	1/2	
n = 4	2/3	1/3	1/3	2/3
n = 5	1/2	1/2	1/4	1/2
n = 6	3/5	2/5	2/5	2/5
n = 7	4/7	3/7	2/7	4/7
n = 8	5/9	4/9	1/3	4/9

The gray entry is the minimum probability among all the entries in the table as well as the lower bounds for large n. This probability is less than the lower bounds for large n listed above. Thus, for an arbitrary collection of intervals, given an arbitrary interval in the collection, the probability that the given interval will appear in a randomly selected maximally disjoint subcollection of intervals is at least $\frac{1}{4}$. The following graph more concisely describes the situation presented by Table 1.



To better understand the final move required to complete the argument, it is useful to consider a graphical representation of our current situation. Consider the following diagram, which represents the case where the collection given consists of seven intervals.



Here, the horizontal axis denotes an interval within the collection. The vertical axis denotes a possible maximally disjoint subcollection. Shaded intervals are included in the subcollection, empty intervals are excluded. The exhaustive list of possible subcollections given by the diagram confirms the above theory that the number of maximally disjoint subcollections of n intervals is the $(n+1)^{st}$ Padovan number. The diagram lists seven possible subcollections; the eighth Padovan number is seven.

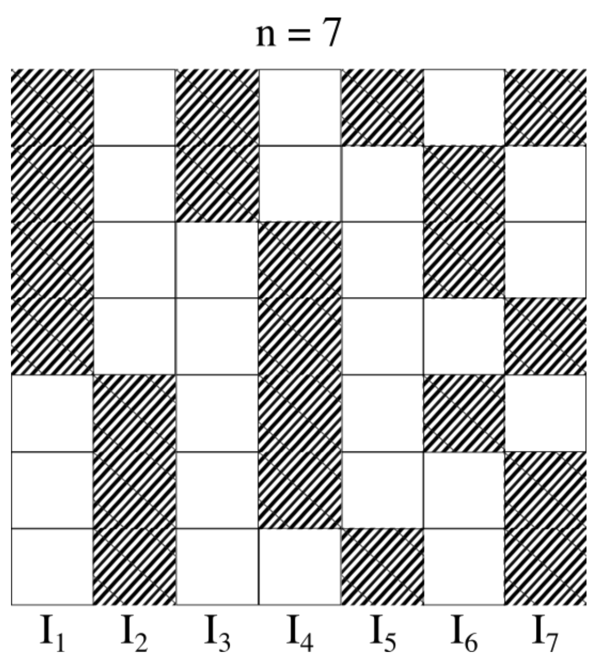
Now, consider an arbitrary point x contained by the original collection of seven intervals. Then x is contained by at least one of the intervals in the collection. Say x is contained by I_3 (it is easy to see this is tied for the worst case scenario with the situation in which x is contained by only I_5). Then the diagram makes it easy to see that x is a member of two of the seven possible maximally disjoint subcollections. Without loss of generality, we may assume the intervals cover $[0, 1]$, since we may easily rescale the diagram otherwise. Then, since the diagram exhaustively lists the possible maximally disjoint subcollections, we may regard our diagram as the probability space $[0, 1] \times [0, 1]$. Call the subset of the probability space given by intervals chosen in maximally disjoint subcollections (the shaded region in the diagram) Δ_7. Then we have for all $x \in [0,1]$ that

$$\frac{2}{7} \leq \int_{y=0}^{1} \chi_{\Delta_7}(x,y)dy.$$

Moreover, since the absolute minimum of $Pr(I_{k,n})$ is $\frac{1}{4}$, by Fubini's theorem we have that

$$\begin{aligned}\frac{1}{4} \leq \inf_{x\in[0,1]} \int_{y=0}^{1} \chi_{\Delta_n}(x,y)dy &\leq \int_{x=0}^{1}\int_{y=0}^{1} \chi_{\Delta_n}(x,y)dy\ dx \\ &= \int_{y=0}^{1}\int_{x=0}^{1} \chi_{\Delta_n}(x,y)dx\ dy.\end{aligned}$$

In order to digest this final result, recall the following figure depicting the case where $n = 7$:



Let each of the I_i represent a student enrolling in one of seven courses. The courses are represented by the seven different spaces along the vertical axis. A shaded square in the column corresponding to I_j represents a course taken by student I_j. As we know that no one student takes fewer than $\frac{2}{7}$ the total available courses, it must also be the case that there is at least one course taken by no fewer than $\frac{2}{7}$ the total number of students.

To see this, suppose that the contrary holds, namely that every course is taken by fewer than $\frac{2}{7}$ the total number of students. So each course is taken by at most one student. Once each student is assigned one course, if any one student receives an additional assigment to another course, that course will have at least two students enrolled, contradicting our supposition. Thus there is at least one course in which at least two of the students will enroll. In much the same way, we have shown that there is at least one maximally disjoint subcollection in which at least $\frac{1}{4}$ of the total length of the collection is filled.

REFERENCES

Alon, Noga, and Joel H. Spencer. *The Probabilistic Method*. 4th ed. N.p.: John Wiley & Sons, 2016. Print. Wiley.

Bateman, Michael. *Kakeya sets and directional maximal operators in the plane*. Duke Math. J. 2009. Print.

Bateman, Michael, and Katz, N. H. *Kakeya sets in Cantor directions*. Math. Res. Lett. 2008. Print.

Cordoba, Alex, and Fefferman, Robert. *A geometric proof of the strong maximal theorem*. Ann. of Math. 1975. Print.

Dunham, William. *Journey through Genius: The Great Theorems of Mathematics*. New York: Penguin, 1991. Print.

Evans, Lawrence, and Gariepy, Ronald. *Measure theory and fine properties of functions*. Boca Raton: CRC Press, 2015. Print.

Falconer, Kenneth J. *The Geometry of Fractal Sets*. Vol. 85. Cambridge: Cambridge UP, 1986. Print. Cambridge Tracts in Mathematics. Print.

Fefferman, Robert. *Some topics from harmonic analysis and partial differential equations*. Essays on Fourier analysis in honor of Elias M. Stein, Princeton, NJ, 1991, 175–210, Princeton Math. Ser. 42. Princeton University Press, 1995. Print.

Garnett, John B. *Bounded Analytic Functions*. New York: Springer, 2011. Print.

de Guzmán, Miguel. *Differentiation of Integrals in* $\mathbb{R}^n$. Lecture Notes in Math. Springer, 1975. Print.

Jones, Frank. *Lebesgue Integration on Euclidean Space*. Revised ed. Boston, Mass.: Jones and Bartlett, 2006. Print.

Padovan, Richard. *Dom Hans Van Der Laan and the plastic number*. Nexus IV: Architecture and Mathematics, 181-193. Florence: Kim Williams Books, 2002. Print.

Spivak, Michael. *Calculus*. Spivak. 3rd ed. New York: Benjamin, 1967. Print.

Stein, Elias. *Singular Integrals and Differentiability Properties of Functions*. Princeton University Press, 1970. Print.

Stewart, Ian. *Tales of a neglected number*. Sci. Amer. 92-93, 1996. Print.