A Short Course on Approximation Theory

N. L. Carothers
Department of Mathematics and Statistics
Bowling Green State University

Preface

These are notes for a topics course offered at Bowling Green State University on a variety of occasions. The course is typically offered during a somewhat abbreviated six week summer session and, consequently, there is a bit less material here than might be associated with a full semester course offered during the academic year. On the other hand, I have tried to make the notes self-contained by adding a number of short appendices and these might well be used to augment the course.

The course title, approximation theory, covers a great deal of mathematical territory. In the present context, the focus is primarily on the approximation of real-valued continuous functions by some simpler class of functions, such as algebraic or trigonometric polynomials. Such issues have attracted the attention of thousands of mathematicians for at least two centuries now. We will have occasion to discuss both venerable and contemporary results, whose origins range anywhere from the dawn of time to the day before yesterday. This easily explains my interest in the subject. For me, reading these notes is like leafing through the family photo album: There are old friends, fondly remembered, fresh new faces, not yet familiar, and enough easily recognizable faces to make me feel right at home.

The problems we will encounter are easy to state and easy to understand, and yet their solutions should prove intriguing to virtually anyone interested in mathematics. The techniques involved in these solutions entail nearly every topic covered in the standard undergraduate curriculum. From that point of view alone, the course should have something to offer both the beginner and the veteran. Think of it as an opportunity to take a grand tour of undergraduate mathematics (with the occasional side trip into graduate mathematics) with the likes of Weierstrass, Gauss, and Lebesgue as our guides.

Approximation theory, as you might guess from its name, has both a pragmatic side, which is concerned largely with computational practicalities, precise estimations of error, and so on, and also a theoretical side, which is more often concerned with existence and uniqueness questions, and "applications" to other theoretical issues. The working professional in the field moves easily between these two seemingly disparate camps; indeed, most modern books on approximation theory will devote a fair number of pages to both aspects of the subject. Being a well-informed amateur rather than a trained expert on the subject, however, my personal preferences have been the driving force behind my selection of topics. Thus, although we will have a few things to say about computational considerations, the primary focus here is on the theory of approximation.

By way of prerequisites, I will freely assume that the reader is familiar with basic notions from linear algebra and advanced calculus. For example, I will assume that the reader is familiar with the notions of a basis for a vector space, linear transformations (maps) defined on a vector space, determinants, and so on; I will also assume that the reader is familiar with the notions of pointwise and uniform convergence for sequence of real-valued functions,

"ε-δ" and "ε-N" proofs (for continuity of a function, say, and convergence of a sequence), closed and compact subsets of the real line, normed vector spaces, and so on. If one or two of these phrases is unfamiliar, don't worry: Many of these topics are reviewed in the text; but if several topics are unfamiliar, please speak with me as soon as possible.

For my part, I have tried to carefully point out thorny passages and to offer at least a few hints or reminders whenever details beyond the ordinary are needed. Nevertheless, in order to fully appreciate the material, it will be necessary for the reader to actually work through certain details. For this reason, I have peppered the notes with a variety of exercises, both big and small, at least a few of which really must be completed in order to follow the discussion.

In the final chapter, where a rudimentary knowledge of topological spaces is required, I am forced to make a few assumptions that may be unfamiliar to some readers. Still, I feel certain that the main results can be appreciated without necessarily following every detail of the proofs.

Finally, I would like to stress that these notes borrow from a number of sources. Indeed, the presentation draws heavily from several classic textbooks, most notably the wonderful books by Natanson [41], de La Vallée Poussin [37], and Cheney [12] (numbers refer to the References at the end of these notes), and from several courses on related topics that I took while a graduate student at The Ohio State University offered by Professor Bogdan Baishanski, whose prowess at the blackboard continues to serve as an inspiration to me. I should also mention that these notes began, some 20 years ago as I write this, as a supplement to Rivlin's classic introduction to the subject [45], which I used as the primary text at the time. This will explain my frequent references to certain formulas or theorems in Rivlin's book. While the notes are no longer dependent on Rivlin, *per se*, it would still be fair to say that they only supplement his more thorough presentation. In fact, wherever possible, I would encourage the interested reader to consult the original sources cited throughout the text.

Contents

Chapter 1

Preliminaries

Introduction

In 1853, the great Russian mathematician, P. L. Chebyshev (Čebyšev), while working on a problem of *linkages*, devices which translate the linear motion of a steam engine into the circular motion of a wheel, considered the following problem:

> Given a continuous function f defined on a closed interval $[\,a,b\,]$ and a positive integer n, can we "represent" f by a polynomial $p(x) = \sum_{k=0}^{n} a_k x^k$, of degree at most n, in such a way that the maximum error at any point x in $[\,a,b\,]$ is controlled? In particular, is it possible to construct p so that the error $\max_{a \le x \le b} |f(x) - p(x)|$ is minimized?

This problem raises several questions, the first of which Chebyshev himself ignored:

– Why should such a polynomial even *exist*?

– If it does, can we hope to *construct* it?

– If it exists, is it also *unique*?

– What happens if we change the measure of error to, say, $\int_a^b |f(x) - p(x)|^2\,dx$?

Exercise 1.1. How do we know that $C[\,a,b\,]$ contains non-polynomial functions? Name one (and explain why it isn't a polynomial)!

Best Approximations in Normed Spaces

Chebyshev's problem is perhaps best understood by rephrasing it in modern terms. What we have here is a problem of *best approximation* in a *normed linear space*. Recall that a *norm* on a (real) *vector space* X is a nonnegative function on X satisfying

$$\|x\| \ge 0, \text{ and } \|x\| = 0 \text{ if and only if } x = 0,$$

$$\|\alpha x\| = |\alpha|\|x\| \text{ for any } x \in X \text{ and } \alpha \in \mathbb{R},$$

$$\|x + y\| \le \|x\| + \|y\| \text{ for any } x,\, y \in X.$$

Any norm on X induces a *metric* or distance function by setting $\operatorname{dist}(x, y) = \|x - y\|$. The abstract version of our problem(s) can now be restated:

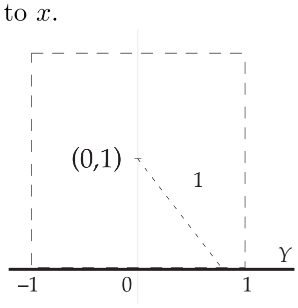
> Given a subset (or even a *subspace*) Y of X and a point $x \in X$, is there an element $y \in Y$ that is *nearest* to x? That is, can we find a vector $y \in Y$ such that $\|x - y\| = \min_{z \in Y} \|x - z\|$? If there is such a *best approximation* to x from elements of Y, is it unique?

It's not hard to see that a satisfactory answer to this question will require that we take Y to be a *closed* set in X, for otherwise points in $\overline{Y} \setminus Y$ (sometimes called the *boundary* of the set Y) will not have nearest points. Indeed, which point in the interval $[0, 1)$ is nearest to 1? Less obvious is that we typically need to impose additional requirements on Y in order to insure the existence (and certainly the uniqueness) of nearest points. For the time being, we will consider the case where Y is a closed *subspace* of a normed linear space X.

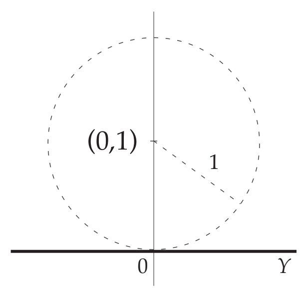
Examples 1.2.

1. As we'll soon see, in $X = \mathbb{R}^n$ with its usual norm $\|(x_k)_{k=1}^n\|_2 = \left(\sum_{k=1}^n |x_k|^2\right)^{1/2}$, the problem has a complete solution for any subspace (or, indeed, any *closed convex set*) Y. This problem is often considered in calculus or linear algebra where it is called "least-squares" approximation. A large part of the current course will be taken up with least-squares approximations, too. For now let's simply note that the problem changes character dramatically if we consider a different norm on $\mathbb{R}^n$, as evidenced by the following example.

2. Consider $X = \mathbb{R}^2$ under the norm $\|(x, y)\| = \max\{|x|, |y|\}$, and consider the subspace $Y = \{(0, y) : y \in \mathbb{R}\}$ (i.e., the y-axis). It's not hard to see that the point $x = (1, 0) \in \mathbb{R}^2$ has infinitely many nearest points in Y; indeed, every point $(0, y)$, $-1 \le y \le 1$, is nearest to x.



sphere of radius 1, max norm



sphere of radius 1, usual norm

3. There are many norms we might consider on $\mathbb{R}^n$. Of particular interest are the ℓ_p-norms; that is, the scale of norms:

$$\|(x_i)_{i=1}^n\|_p = \left(\sum_{k=1}^n |x_k|^p\right)^{1/p}, \quad 1 \le p < \infty,$$

and

$$\|(x_i)_{i=1}^n\|_\infty = \max_{1 \le i \le n} |x_i|.$$

It's easy to see that $\|\cdot\|_1$ and $\|\cdot\|_\infty$ define norms. The other cases take a bit more work; for full details see Appendix A.

4. The ℓ_2-norm is an example of a norm induced by an *inner product* (or "dot" product). You will recall that the expression

$$\langle\, x, y \,\rangle = \sum_{i=1}^{n} x_i y_i,$$

where $x = (x_i)_{i=1}^n$ and $y = (y_i)_{i=1}^n$, defines an inner product on $\mathbb{R}^n$ and that the norm in $\mathbb{R}^n$ satisfies

$$\|x\|_2 = \sqrt{\langle\, x, x \,\rangle}\,.$$

In this sense, the usual norm on $\mathbb{R}^n$ is actually induced by the inner product. More generally, any inner product will give rise to a norm in this same way. (But not vice versa. As we'll see, inner product norms satisfy a number of special properties that aren't enjoyed by all norms.)

The presence of an inner product in an abstract space opens the door to geometric arguments that are remarkably similar to those used in $\mathbb{R}^n$. (See Appendix D for more details.) Luckily, inner products are easy to come by in practice. By way of one example, consider this: Given a positive Riemann integrable *weight function* $w(x)$ defined on some interval $[\,a, b\,]$, it's not hard to check that the expression

$$\langle\, f, g \,\rangle = \int_a^b f(t)\, g(t)\, w(t)\, dt$$

defines an inner product on $C[\,a, b\,]$, the space of all continuous, real-valued functions $f : [\,a, b\,] \to \mathbb{R}$, with associated norm

$$\|f\|_2 = \left(\int_a^b |f(t)|^2 w(t)\, dt \right)^{1/2}.$$

We will take full advantage of this fact in later chapters (in particular, Chapters 8–9).

5. Our original problem concerns the space $X = C[\,a, b\,]$ under the *uniform norm* $\|f\| = \max_{a \le x \le b} |f(x)|$. The adjective "uniform" is used here because convergence in this norm is the same as uniform convergence on $[\,a, b\,]$:

$$\|f_n - f\| \to 0 \iff f_n \to f \text{ uniformly on } [\,a, b\,]$$

(which we will frequently abbreviate by writing $f_n \rightrightarrows f$ on $[\,a, b\,]$). This, by the way, is the norm of choice on $C[\,a, b\,]$ (largely because continuity is preserved by uniform limits). In this particular case we're interested in approximations by elements of $Y = \mathcal{P}_n$, the *subspace* of all polynomials of degree at most n in $C[\,a, b\,]$. It's not hard to see that $\mathcal{P}_n$ is a finite-dimensional subspace of $C[\,a, b\,]$ of dimension exactly $n + 1$. (Why?)

6. If we consider the subspace $Y = \mathcal{P}$ consisting of *all* polynomials in $X = C[\,a, b\,]$, we readily see that the existence of best approximations can be problematic. It follows from the Weierstrass theorem, for example, that each $f \in C[\,a, b\,]$ has distance 0 from $\mathcal{P}$ but, because not every $f \in C[\,a, b\,]$ is a polynomial (why?), we can't hope for a best approximating polynomial to exist in every case. For example, the function

$f(x) = x \sin(1/x)$ is continuous on $[0,1]$ but can't possibly agree with any polynomial on $[0,1]$. (Why?) As you may have already surmised, the problem here is that every element of $C[a,b]$ is the (uniform) limit of a sequence from $\mathcal{P}$; in other words, the closure of Y equals X; in symbols, $\overline{Y} = X$.

Finite-Dimensional Vector Spaces

The key to the problem of polynomial approximation is the fact that each of the spaces $\mathcal{P}_n$, described in Examples 1.2 (5), is *finite-dimensional*. To see how finite-dimensionality comes into play, it will be most efficient to consider the abstract setting of finite-dimensional subspaces of arbitrary normed spaces.

Lemma 1.3. *Let V be a finite-dimensional vector space. Then, all norms on V are equivalent. That is, if $\|\cdot\|$ and $|||\cdot|||$ are norms on V, then there exist constants $0 < A,\ B < \infty$ such that*

$$A\,\|x\| \le |||x||| \le B\,\|x\|$$

for all vectors $x \in V$.

Proof. Suppose that V is n-dimensional and that $\|\cdot\|$ is a norm on V. Fix a basis $e_1, \ldots, e_n$ for V and consider the norm

$$\left\| \sum_{i=1}^n a_i e_i \right\|_1 = \sum_{i=1}^n |a_i| = \|(a_i)_{i=1}^n\|_1$$

for $x = \sum_{i=1}^n a_i e_i \in V$. Because $e_1, \ldots, e_n$ is a basis for V, it's not hard to see that $\|\cdot\|_1$ is, indeed, a norm on V. [Notice that we've actually set-up a correspondence between $\mathbb{R}^n$ and V; specifically, the map $(a_i)_{i=1}^n \mapsto \sum_{i=1}^n a_i e_i$ is obviously both one-to-one and onto. In fact, this correspondence is an *isometry* between $(\mathbb{R}^n, \|\cdot\|_1)$ and $(V, \|\cdot\|_1)$.]

It now suffices to show that $\|\cdot\|$ and $\|\cdot\|_1$ are equivalent. (Why?)

One inequality is easy to show; indeed, notice that

$$\left\| \sum_{i=1}^n a_i e_i \right\| \le \sum_{i=1}^n |a_i|\,\|e_i\| \le \left(\max_{1 \le i \le n} \|e_i\| \right) \sum_{i=1}^n |a_i| = B \left\| \sum_{i=1}^n a_i e_i \right\|_1 .$$

The real work comes in establishing the other inequality.

Now the inequality we've just established shows that the function $x \mapsto \|x\|$ is *continuous* on the space $(V, \|\cdot\|_1)$; indeed,

$$\big|\, \|x\| - \|y\| \,\big| \le \|x - y\| \le B\,\|x - y\|_1$$

for any $x,\ y \in V$. Thus, $\|\cdot\|$ assumes a *minimum* value on the *compact* set

$$S = \{x \in V : \|x\|_1 = 1\}.$$

(Why is S compact?) In particular, there is some $A > 0$ such that $\|x\| \ge A$ whenever $\|x\|_1 = 1$. (Why can we assume that $A > 0$?) The inequality we need now follows from the homogeneity of the norm:

$$\left\| \frac{x}{\|x\|_1} \right\| \ge A \Longrightarrow \|x\| \ge A\,\|x\|_1. \qquad \square$$

Corollary 1.4. *Every finite-dimensional normed space is complete (that is, every Cauchy seqeuence converges). In particular, if Y is a finite-dimensional subspace of a normed linear space X, then Y is a closed subset of X.*

Corollary 1.5. *Let Y be a finite-dimensional normed space, let $x \in Y$, and let $M > 0$. Then, any closed ball $\{\, y \in Y : \|x - y\| \le M \,\}$ is compact.*

Proof. Because translation is an isometry, it clearly suffices to show that the set $\{\, y \in Y : \|y\| \le M \,\}$ (i.e., the ball about 0) is compact.

Suppose now that Y is n-dimensional and that $e_1, \ldots, e_n$ is a basis for Y. From Lemma 1.3 we know that there is some constant $A > 0$ such that

$$A \sum_{i=1}^{n} |a_i| \le \left\| \sum_{i=1}^{n} a_i e_i \right\|$$

for all $x = \sum_{i=1}^{n} a_i e_i \in Y$. In particular,

$$A\,|a_i| \le \left\| \sum_{i=1}^{n} a_i e_i \right\| \le M \implies |a_i| \le M/A \quad \text{for} \quad i = 1, \ldots, n.$$

Thus, $\{\, y \in Y : \|y\| \le M \,\}$ is a *closed* subset (why?) of the *compact* set

$$\left\{ x = \sum_{i=1}^{n} a_i e_i : |a_i| \le M/A,\ i = 1, \ldots, n \right\} = [\,-M/A,\, M/A\,]^n. \qquad \square$$

Theorem 1.6. *Let Y be a finite-dimensional subspace of a normed linear space X, and let $x \in X$. Then, there exists a (not necessarily unique) vector $y^* \in Y$ such that*

$$\|x - y^*\| = \min_{y \in Y} \|x - y\|$$

for all $y \in Y$. That is, there is a best approximation to x by elements from Y.

Proof. First notice that because $0 \in Y$, we know that any nearest point y^* will satisfy $\|x - y^*\| \le \|x\| = \|x - 0\|$. Thus, it suffices to look for y^* in the *compact* set

$$K = \{\, y \in Y : \|x - y\| \le \|x\,\| \}.$$

To finish the proof, we need only note that the function $f(y) = \|x - y\|$ is *continuous*:

$$|f(y) - f(z)| = \Big|\, \|x - y\| - \|x - z\| \,\Big| \le \|y - z\|,$$

and hence attains a minimum value at some point $y^* \in K$. $\square$

Corollary 1.7. *For each $f \in C[\,a, b\,]$ and each positive integer n, there is a (not necessarily unique) polynomial $p_n^* \in \mathcal{P}_n$ such that*

$$\|f - p_n^*\| = \min_{p \in \mathcal{P}_n} \|f - p\|.$$

Example 1.8. Nothing in Corollary 1.7 says that p_n^* will be a polynomial of degree *exactly* n—rather, it's a polynomial of degree *at most* n. For example, the best approximation to $f(x) = x$ by a polynomial of degree at most 3 is, of course, $p(x) = x$. Even examples of nonpolynomial functions are easy to come by; for instance, the best linear approximation to $f(x) = |x|$ on $[-1, 1]$ is actually the constant function $p(x) = 1/2$, and this makes for an entertaining **exercise**.

Before we leave these "soft" arguments behind, let's discuss the problem of *uniqueness* of best approximations. First, let's see why we might like to have unique best approximations:

Lemma 1.9. *Let Y be a finite-dimensional subspace of a normed linear space X, and suppose that each $x \in X$ has a unique nearest point $y_x \in Y$. Then the nearest point map $x \mapsto y_x$ is continuous.*

Proof. Let's write $P(x) = y_x$ for the nearest point map, and let's suppose that $x_n \to x$ in X. We want to show that $P(x_n) \to P(x)$, and for this it's enough to show that there is a subsequence of $(P(x_n))$ that converges to $P(x)$. (Why?)

Because the sequence (x_n) is bounded in X, say $\|x_n\| \le M$ for all n, we have

$$\|P(x_n)\| \le \|P(x_n) - x_n\| + \|x_n\| \le 2\|x_n\| \le 2M.$$

Thus, $(P(x_n))$ is a bounded sequence in Y, a finite-dimensional space. As such, by passing to a subsequence, we may suppose that $(P(x_n))$ converges to some element $P_0 \in Y$. (How?) Now we need to show that $P_0 = P(x)$. But

$$\|P(x_n) - x_n\| \le \|P(x) - x_n\|$$

for any n. (Why?) Hence, letting $n \to \infty$, we get

$$\|P_0 - x\| \le \|P(x) - x\|.$$

Because nearest points in Y are unique, we must have $P_0 = P(x)$. □

Exercise 1.10. Let X be a metric (or normed) space and let $f : X \to X$. Show that f is continuous at $x \in X$ if and only if, whenever $x_n \to x$ in X, some subsequence of $(f(x_n))$ converges to $f(x)$. [Hint: The forward direction is easy; for the backward implication, suppose that $(f(x_n))$ fails to converge to $f(x)$ and work toward a contradiction.]

It should be pointed out that the nearest point map is, in general, *nonlinear* and, as such, can be very difficult to work with. Later we'll see at least one case in which nearest point maps always turn out to be linear.

In spite of any potential difficulties with the nearest point map, we next observe that the set of best approximations has a well-behaved, almost-linear structure.

Theorem 1.11. *Let Y be a subspace of a normed linear space X, and let $x \in X$. The set Y_x, consisting of all best approximations to x out of Y, is a bounded* convex *set.*

Proof. As we've seen, the set Y_x is a subset of the ball $\{ y \in X : \|x - y\| \le \|x\| \}$ and, as such, is bounded. (More generally, the set Y_x is a subset of the sphere $\{ y \in X : \|x - y\| = d \}$, where $d = \operatorname{dist}(x, Y) = \inf_{y \in Y} \|x - y\|$.)

Next recall that a subset K of a vector space V is said to be *convex* if K contains the line segment joining any pair of its points. Specifically, K is convex if

$$x, y \in K,\ 0 \le \lambda \le 1 \implies \lambda x + (1 - \lambda) y \in K.$$

Thus, given $y_1, y_2 \in Y_x$ and $0 \le \lambda \le 1$, we want to show that the vector $y^* = \lambda y_1 + (1-\lambda) y_2 \in Y_x$. But $y_1, y_2 \in Y_x$ means that

$$\|x - y_1\| = \|x - y_2\| = \min_{y \in Y} \|x - y\|.$$

Hence,

$$\begin{aligned} \|x - y^*\| &= \|x - (\lambda y_1 + (1-\lambda) y_2)\| \\ &= \|\lambda(x - y_1) + (1-\lambda)(x - y_2)\| \\ &\le \lambda \|x - y_1\| + (1-\lambda)\|x - y_2\| \\ &= \min_{y \in Y} \|x - y\|. \end{aligned}$$

Consequently, $\|x - y^*\| = \min_{y \in Y} \|x - y\|$; that is, $y^* \in Y_x$. □

Exercise 1.12. If, in Theorem 1.11, we also assume that Y is finite-dimensional, show that Y_x is *closed* (hence a *compact* convex set).

If Y_x contains more than one point, then, in fact, it contains an entire line segment. Thus, Y_x is either empty, contains exactly one point, or contains infinitely many points. This observation gives us a *sufficient condition* for uniqueness of nearest points: If the normed space X contains no line segments on any sphere $\{x \in X : \|x\| = r\}$, then best approximations (out of any convex subset Y) will necessarily be unique.

A norm $\|\cdot\|$ on a vector space X is said to be *strictly convex* if, for any pair of points $x \ne y \in X$ with $\|x\| = r = \|y\|$, we always have $\|\lambda x + (1-\lambda) y\| < r$ for all $0 < \lambda < 1$. That is, the open line segment between any pair of points on the sphere of radius r lies entirely within the open ball of radius r; in other words, only the endpoints of the line segment can hit the sphere. For simplicity, we often say that the space X is strictly convex, with the understanding that we're actually referring to a property of the norm in X. In any such space, we get an immediate corollary to our last result:

Corollary 1.13. *If X has a strictly convex norm, then, for any subspace Y of X and any point $x \in X$, there can be at most one best approximation to x out of Y. That is, Y_x is either empty or consists of a single point.*

In order to arrive at a condition that's somewhat easier to check, let's translate our original definition into a statement about the triangle inequality in X.

Lemma 1.14. *A normed space X has a strictly convex norm if and only if the triangle inequality is strict on nonparallel vectors; that is, if and only if*

$$x \ne \alpha y, \;\; y \ne \alpha x, \;\; \text{all } \alpha \in \mathbb{R} \implies \|x + y\| < \|x\| + \|y\|.$$

Proof. First suppose that X is strictly convex, and let x and y be nonparallel vectors in X. Then, in particular, the vectors $x/\|x\|$ and $y/\|y\|$ must be different. (Why?) Hence,

$$\left\| \left(\frac{\|x\|}{\|x\| + \|y\|} \right) \frac{x}{\|x\|} \; + \; \left(\frac{\|y\|}{\|x\| + \|y\|} \right) \frac{y}{\|y\|} \right\| < 1.$$

That is, $\|x + y\| < \|x\| + \|y\|$.

Next suppose that the triangle inequality is strict on nonparallel vectors, and let $x \neq y \in X$ with $\|x\| = r = \|y\|$. If x and y are parallel, then we must have $y = -x$. (Why?) In this case,

$$\|\lambda x + (1 - \lambda)\, y\| = |2\lambda - 1|\, \|x\| < r$$

because $-1 < 2\lambda - 1 < 1$ whenever $0 < \lambda < 1$. Otherwise, x and y are nonparallel. Thus, for any $0 < \lambda < 1$, the vectors λx and $(1 - \lambda)\, y$ are likewise nonparallel and we have

$$\|\lambda x + (1 - \lambda)\, y\| < \lambda\|x\| + (1 - \lambda)\|y\| = r. \qquad \square$$

Examples 1.15.

1. The usual norm on $C[\,a, b\,]$ is *not* strictly convex (and so the problem of uniqueness of best approximations is all the more interesting to tackle). For example, if $f(x) = x$ and $g(x) = x^2$ in $C[\,0, 1\,]$, then $f \neq g$ and $\|f\| = 1 = \|g\|$, while $\|f + g\| = 2$. (Why?)

2. The usual norm on $\mathbb{R}^n$ is strictly convex, as is any one of the norms $\|\cdot\|_p$ for $1 < p < \infty$. (See Problem 10.) The norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$, on the other hand, are *not* strictly convex. (Why?)

Problems

[Problems marked ($\triangleright$) are essential to a full understanding of the course. Problems marked ($*$) are of general interest and are offered as a contribution to your personal growth. Unmarked problems are just for fun.]

The most important collection of functions for our purposes is the space $C[\,a,b\,]$, consisting of all continuous functions $f : [\,a,b\,] \to \mathbb{R}$. It's easy to see that $C[\,a,b\,]$ is a vector space under the usual pointwise operations on functions: $(f+g)(x) = f(x) + g(x)$ and $(\alpha f)(x) = \alpha f(x)$ for $\alpha \in \mathbb{R}$. Actually, we will be most interested in the finite-dimensional subspaces $\mathcal{P}_n$ of $C[\,a,b\,]$, consisting of all *algebraic polynomials* of degree at most n.

$\triangleright$ 1. The subspace $\mathcal{P}_n$ has dimension exactly $n+1$. Why?

Another useful subset of $C[\,a,b\,]$ is the collection $\operatorname{lip}_K\alpha$, consisting of those f which satisfy a *Lipschitz condition* of order $\alpha > 0$ with constant $0 < K < \infty$; i.e., those f for which $|f(x) - f(y)| \le K\,|x-y|^\alpha$ for all x, y in $[\,a,b\,]$. [Some authors would say that f is *Hölder continuous* with exponent α.]

$*$ 2. (a) Show that $\operatorname{lip}_K\alpha$ is, indeed, a subset of $C[\,a,b\,]$.

(b) If $\alpha > 1$, show that $\operatorname{lip}_K\alpha$ contains only the constant functions.

(c) Show that $\sqrt{x}$ is in $\operatorname{lip}_1(1/2)$ and that $\sin x$ is in $\operatorname{lip}_1 1$ on $[\,0,1\,]$.

(d) Show that the collection $\operatorname{lip}\alpha$, consisting of all those f which are in $\operatorname{lip}_K\alpha$ for some K, is a subspace of $C[\,a,b\,]$.

(e) Show that $\operatorname{lip} 1$ contains all the polynomials.

(f) If $f \in \operatorname{lip}\alpha$ for some $\alpha > 0$, show that $f \in \operatorname{lip}\beta$ for all $0 < \beta < \alpha$.

(g) Given $0 < \alpha < 1$, show that x^α is in $\operatorname{lip}_1\alpha$ on $[\,0,1\,]$ but *not* in $\operatorname{lip}\beta$ for any $\beta > \alpha$.

The vector space $C[\,a,b\,]$ is most commonly endowed with the *uniform* or *sup norm*, defined by $\|f\| = \max_{a\le x\le b}|f(x)|$. Some authors use $\|f\|_u$ or $\|f\|_\infty$ here, and some authors refer to this as the *Chebyshev norm*. Whatever the notation used, it is the norm of choice on $C[\,a,b\,]$.

$*$ 3. Show that $\mathcal{P}_n$ and $\operatorname{lip}_K\alpha$ are closed subsets of $C[\,a,b\,]$ (under the sup norm). Is $\operatorname{lip}\alpha$ closed? A bit harder: Show that $\operatorname{lip} 1$ is both first category and dense in $C[\,a,b\,]$.

4. Fix n and consider the norm $\|p\|_1 = \sum_{k=0}^n |a_k|$ for $p(x) = a_0 + a_1x + \cdots + a_nx^n \in \mathcal{P}_n$, considered as a subset of $C[\,a,b\,]$. Show that there are constants $0 < A_n,\ B_n < \infty$ such that $A_n\|p\|_1 \le \|p\| \le B_n\|p\|_1$, where $\|p\| = \max\limits_{a\le x\le b}|p(x)|$. Do A_n and B_n really depend on n? Do they depend on the underlying interval $[\,a,b\,]$?

5. Fill-in any missing details from Example 1.8.

We will occasionally consider spaces of real-valued functions defined on finite sets; that is, we will consider $\mathbb{R}^n$ under various norms. (Why is this the same?) We define a scale of norms on $\mathbb{R}^n$ by setting $\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$, where $x = (x_1,\ldots,x_n)$ and $1 \le p < \infty$. (We need $p \ge 1$ in order for this expression to be a legitimate norm, but the expression makes perfect sense for any $p > 0$, and even for $p < 0$ provided no x_i is 0.) Notice, please, that the usual norm on $\mathbb{R}^n$ is given by $\|x\|_2$.

6. Show that $\lim_{p\to\infty} \|x\|_p = \max_{1\le i\le n} |x_i|$. For this reason we define
$$\|x\|_\infty = \max_{1\le i\le n} |x_i|.$$
Thus $\mathbb{R}^n$ under the norm $\|\cdot\|_\infty$ is the same as $C(\{1,2,\ldots,n\})$ under its usual norm.
7. Assuming $x_i \neq 0$ for $i = 1,\ldots,n$, compute $\lim_{p\to 0^+} \|x\|_p$ and $\lim_{p\to -\infty} \|x\|_p$.
8. Consider $\mathbb{R}^2$ under the norm $\|x\|_p$. Draw the graph of the unit sphere $\{x : \|x\|_p = 1\}$ for various values of p (especially $p = 1$, 2, ∞).
9. In a normed space normed space $(X, \|\cdot\|)$, prove that the following are equivalent:
   (a) $\|x+y\| = \|x\| + \|y\|$ always implies that x and y lie in the same direction; that is, either $x = \alpha y$ or $y = \alpha x$ for some nonnegative scalar α.
   (b) If x, $y \in X$ are nonparallel, then $\left\|\dfrac{x+y}{2}\right\| < \dfrac{\|x\|+\|y\|}{2}$.
   (c) If $x \neq y \in X$ with $\|x\| = 1 = \|y\|$, then $\left\|\dfrac{x+y}{2}\right\| < 1$.
   (d) X is strictly convex (as defined on page 7).

We write ℓ_p^n to denote the vector space of sequences of length n endowed with the p-norm; that is, $\mathbb{R}^n$ supplied with the norm $\|\cdot\|_p$. And we write ℓ_p to denote the vector space of infinite length sequences $x = (x_n)_{n=1}^\infty$ for which $\|x\|_p < \infty$. In each space, the usual algebraic operations are defined pointwise (or coordinatewise) and the norm is understood to be $\|\cdot\|_p$.

10. Show that ℓ_p (and hence ℓ_p^n) is strictly convex for $1 < p < \infty$. Show also that this fails in cases $p = 1$ and $p = \infty$. [Hint: Show that the function $f(t) = |t|^p$ satisfies $f((s+t)/2) < (f(s)+f(t))/2$ whenever $s \neq t$ and $1 < p < \infty$.]

$*$ 11. Let X be a normed space and let $B = \{x \in X : \|x\| \le 1\}$. Show that B is a closed convex set.

12. Consider $\mathbb{R}^2$ under the norm $\|\cdot\|_\infty$. Let $B = \{y \in \mathbb{R}^2 : \|y\|_\infty \le 1\}$ and let $x = (2, 0)$. Show that there are infinitely many points in B nearest to x.
13. Let K be a *compact* convex set in a strictly convex space X and let $x \in X$. Show that x has a *unique* nearest point $y_0 \in K$.
14. Let K be a closed subset of a complete normed space X. Prove that K is convex if and only if K is *midpoint convex*; that is, if and only if $(x+y)/2 \in K$ whenever x, $y \in K$. Is this result true in more general settings? For example, can you prove it without assuming completeness? Or, for that matter, is it true for arbitrary sets in any vector space (i.e., without even assuming the presence of a norm)?

Chapter 2

Approximation by Algebraic Polynomials

The Weierstrass Theorem

Let's begin with some notation. Throughout this chapter, we'll be concerned with the problem of best (uniform) approximation of a given function $f \in C[\,a,b\,]$ by elements from $\mathcal{P}_n$, the subspace of algebraic polynomials of degree at most n in $C[\,a,b\,]$. We know that the problem has a solution (possibly more than one), which we've chosen to write as p_n^*. We set

$$E_n(f) = \min_{p \in \mathcal{P}_n} \|f - p\| = \|f - p_n^*\|.$$

Because $\mathcal{P}_n \subset \mathcal{P}_{n+1}$ for each n, it's clear that $E_n(f) \geq E_{n+1}(f)$ for each n. Our goal in this chapter is to prove that $E_n(f) \to 0$. We'll accomplish this by proving:

Theorem 2.1. (The Weierstrass Approximation Theorem, 1885) *Let $f \in C[\,a,b\,]$. Then, for every $\varepsilon > 0$, there is a polynomial p such that $\|f - p\| < \varepsilon$.*

It follows from the Weierstrass theorem that, for some sequence of polynomials (q_k), we have $\|f - q_k\| \to 0$. We may suppose that $q_k \in \mathcal{P}_{n_k}$ where (n_k) is increasing. (Why?) Whence it follows that $E_n(f) \to 0$; that is, $p_n^* \rightrightarrows f$. (Why?) This is an important first step in determining the exact nature of $E_n(f)$ as a function of f and n. We'll look for much more precise information in later sections.

Now there are many proofs of the Weierstrass theorem (a mere three are outlined in the exercises, but there are hundreds!), and all of them start with one simplification: The underlying interval $[\,a,b\,]$ is of no consequence.

Lemma 2.2. *If the Weierstrass theorem holds for $C[\,0,1\,]$, then it also holds for $C[\,a,b\,]$, and conversely. In fact, $C[\,0,1\,]$ and $C[\,a,b\,]$ are, for all practical purposes, identical: They are linearly isometric as normed spaces, order isomorphic as lattices, and isomorphic as algebras* (rings).

Proof. We'll settle for proving only the first assertion; the second is outlined in the exercises (and uses a similar argument). See Problem 1.

First, notice that the function

$$\sigma(x) = a + (b - a)x, \quad 0 \leq x \leq 1,$$

defines a continuous, one-to-one map from $[0,1]$ onto $[a,b]$. Given $f \in C[a,b]$, it follows that $g(x) = f(\sigma(x))$ defines an element of $C[0,1]$. Moreover,

$$\max_{0 \le x \le 1} |g(x)| = \max_{a \le t \le b} |f(t)|.$$

Now, given $\varepsilon > 0$, suppose that we can find a polynomial p such that $\|g - p\| < \varepsilon$; in other words, suppose that

$$\max_{0 \le x \le 1} \big| f\big(a + (b-a)x\big) - p(x) \big| < \varepsilon.$$

Then,

$$\max_{a \le t \le b} \left| f(t) - p\left(\frac{t-a}{b-a}\right) \right| < \varepsilon.$$

(Why?) But if $p(x)$ is a polynomial in x, then $q(t) = p\left(\frac{t-a}{b-a}\right)$ is a polynomial in t satisfying $\|f - q\| < \varepsilon$.

The proof of the converse is entirely similar: If $g(x)$ is an element of $C[0,1]$, then $f(t) = g\left(\frac{t-a}{b-a}\right)$, $a \le t \le b$, defines an element of $C[a,b]$. Moreover, if $q(t)$ is a polynomial in t approximating $f(t)$, then $p(x) = q(a + (b-a)x)$ is a polynomial in x approximating $g(x)$. The remaining details are left as an **exercise**. □

The point to our first result is that it suffices to prove the Weierstrass theorem for any interval we like; $[0,1]$ and $[-1,1]$ are popular choices, but it hardly matters which interval we use.

Bernstein's Proof

The proof of the Weierstrass theorem we present here is due to the great Russian mathematician S. N. Bernstein in 1912. Bernstein's proof is of interest to us for a variety of reasons; perhaps most important is that Bernstein actually *displays* a sequence of polynomials that approximate a given $f \in C[0,1]$. Moreover, as we'll see later, Bernstein's proof generalizes to yield a powerful, unifying theorem, called the Bohman-Korovkin theorem (see Theorem 2.9).

If f is any *bounded* function on $[0,1]$, we define the sequence of *Bernstein polynomials* for f by

$$\big(B_n(f)\big)(x) = \sum_{k=0}^{n} f(k/n) \cdot \binom{n}{k} x^k (1-x)^{n-k}, \qquad 0 \le x \le 1.$$

Please note that $B_n(f)$ is a polynomial of degree at most n. Also, it's easy to see that $\big(B_n(f)\big)(0) = f(0)$, and $\big(B_n(f)\big)(1) = f(1)$. In general, $\big(B_n(f)\big)(x)$ is an *average* of the numbers $f(k/n)$, $k = 0, \ldots, n$. Bernstein's theorem states that $B_n(f) \rightrightarrows f$ for each $f \in C[0,1]$. Surprisingly, the proof actually only requires that we check three easy cases:

$$f_0(x) = 1, \quad f_1(x) = x, \quad \text{and} \quad f_2(x) = x^2.$$

This, and more, is the content of the following lemma.

Lemma 2.3. (i) $B_n(f_0) = f_0$ *and* $B_n(f_1) = f_1$.

(ii) $B_n(f_2) = \left(1 - \frac{1}{n}\right) f_2 + \frac{1}{n} f_1$, *and hence* $B_n(f_2) \rightrightarrows f_2$.

(iii) $\displaystyle\sum_{k=0}^{n}\Big(\frac{k}{n}-x\Big)^2\binom{n}{k}x^k(1-x)^{n-k}=\frac{x(1-x)}{n}\leq\frac{1}{4n},$ *if* $0\leq x\leq 1.$

(iv) *Given* $\delta>0$ *and* $0\leq x\leq 1$, *let* F *denote the set of* k *in* $\{0,\dots,n\}$ *for which* $\left|\dfrac{k}{n}-x\right|\geq\delta$. *Then* $\displaystyle\sum_{k\in F}\binom{n}{k}x^k(1-x)^{n-k}\leq\frac{1}{4n\delta^2}.$

Proof. That $B_n(f_0)=f_0$ follows from the binomial formula:

$$\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}=[x+(1-x)]^n=1.$$

To see that $B_n(f_1)=f_1$, first notice that for $k\geq 1$ we have

$$\frac{k}{n}\binom{n}{k}=\frac{(n-1)\,!}{(k-1)\,!\,(n-k)\,!}=\binom{n-1}{k-1}.$$

Consequently,

$$\begin{aligned}\sum_{k=0}^{n}\frac{k}{n}\binom{n}{k}x^k(1-x)^{n-k} &= x\sum_{k=1}^{n}\binom{n-1}{k-1}x^{k-1}(1-x)^{n-k}\\ &= x\sum_{j=0}^{n-1}\binom{n-1}{j}x^j(1-x)^{(n-1)-j}\;=\;x.\end{aligned}$$

Next, to compute $B_n(f_2)$, we rewrite twice:

$$\begin{aligned}\left(\frac{k}{n}\right)^2\binom{n}{k}=\frac{k}{n}\binom{n-1}{k-1} &= \frac{n-1}{n}\cdot\frac{k-1}{n-1}\binom{n-1}{k-1}+\frac{1}{n}\binom{n-1}{k-1},\ \text{if } k\geq 1\\ &= \left(1-\frac{1}{n}\right)\binom{n-2}{k-2}+\frac{1}{n}\binom{n-1}{k-1},\ \text{if } k\geq 2.\end{aligned}$$

Thus,

$$\begin{aligned}&\sum_{k=0}^{n}\left(\frac{k}{n}\right)^2\binom{n}{k}x^k(1-x)^{n-k}\\ &\quad=\left(1-\frac{1}{n}\right)\sum_{k=2}^{n}\binom{n-2}{k-2}x^k(1-x)^{n-k}\;+\;\frac{1}{n}\sum_{k=1}^{n}\binom{n-1}{k-1}x^k(1-x)^{n-k}\\ &\quad=\left(1-\frac{1}{n}\right)x^2+\frac{1}{n}\,x,\end{aligned}$$

which establishes (ii) because $\|B_n(f_2)-f_2\|=\frac{1}{n}\|f_1-f_2\|\to 0$ as $n\to\infty$.

To prove (iii) we combine the results in (i) and (ii) and simplify. Because $((k/n)-x)^2=(k/n)^2-2x(k/n)+x^2$, we get

$$\begin{aligned}\sum_{k=0}^{n}\left(\frac{k}{n}-x\right)^2\binom{n}{k}x^k(1-x)^{n-k} &= \left(1-\frac{1}{n}\right)x^2+\frac{1}{n}x-2x^2+x^2\\ &= \frac{1}{n}\,x(1-x)\;\leq\;\frac{1}{4n},\end{aligned}$$

for $0 \le x \le 1$.

Finally, to prove (iv), note that $1 \le ((k/n) - x)^2/\delta^2$ for $k \in F$, and hence

$$\begin{aligned}\sum_{k\in F}\binom{n}{k}x^k(1-x)^{n-k} &\le \frac{1}{\delta^2}\sum_{k\in F}\Big(\frac{k}{n}-x\Big)^2\binom{n}{k}x^k(1-x)^{n-k}\\ &\le \frac{1}{\delta^2}\sum_{k=0}^{n}\Big(\frac{k}{n}-x\Big)^2\binom{n}{k}x^k(1-x)^{n-k}\\ &\le \frac{1}{4n\delta^2},\quad \text{from (iii).} \qquad \square\end{aligned}$$

Now we're ready for *the proof of Bernstein's theorem*:

Proof. Let $f \in C[\,0,1\,]$ and let $\varepsilon > 0$. Then, because f is uniformly continuous, there is a $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon/2$ whenever $|x - y| < \delta$. Now we use the previous lemma to estimate $\|f - B_n(f)\|$. First notice that because the numbers $\binom{n}{k}x^k(1-x)^{n-k}$ are nonnegative and sum to 1, we have

$$\begin{aligned}|f(x) - B_n(f)(x)| &= \left| f(x) - \sum_{k=0}^{n}\binom{n}{k} f\left(\frac{k}{n}\right) x^k(1-x)^{n-k}\right|\\ &= \left|\sum_{k=0}^{n}\left(f(x) - f\left(\frac{k}{n}\right)\right)\binom{n}{k}x^k(1-x)^{n-k}\right|\\ &\le \sum_{k=0}^{n}\left|f(x) - f\left(\frac{k}{n}\right)\right|\binom{n}{k}x^k(1-x)^{n-k},\end{aligned}$$

Now fix n (to be specified in a moment) and let F denote the set of k in $\{0, \dots, n\}$ for which $|(k/n) - x| \ge \delta$. Then $|f(x) - f(k/n)| < \varepsilon/2$ for $k \notin F$, while $|f(x) - f(k/n)| \le 2\|f\|$ for $k \in F$. Thus,

$$\begin{aligned}&\big|f(x) - \big(B_n(f)\big)(x)\big|\\ &\quad\le \frac{\varepsilon}{2}\sum_{k\notin F}\binom{n}{k}x^k(1-x)^{n-k} + 2\|f\|\sum_{k\in F}\binom{n}{k}x^k(1-x)^{n-k}\\ &\quad< \frac{\varepsilon}{2}\cdot 1 \;+\; 2\|f\|\cdot\frac{1}{4n\delta^2},\quad \text{from (iv) of the Lemma,}\\ &\quad< \varepsilon,\quad \text{provided that } n > \|f\|/\varepsilon\delta^2. \qquad \square\end{aligned}$$

Landau's Proof

Just because it's good for us, let's give *a second proof of Weierstrass's theorem*. This one is due to Landau in 1908. First, given $f \in C[\,0,1\,]$, notice that it suffices to approximate $f - p$, where p is any polynomial. (Why?) In particular, by subtracting the *linear* function $f(0) + x(f(1) - f(0))$, we may suppose that $f(0) = f(1) = 0$ and, hence, that $f \equiv 0$ outside $[\,0,1\,]$. That is, we may suppose that f is defined and uniformly continuous on all of $\mathbb{R}$.

Again we will display a sequence of polynomials that converge uniformly to f; this time we define

$$L_n(x) = c_n\int_{-1}^{1} f(x+t)\,(1-t^2)^n\,dt,$$

where c_n is chosen so that

$$c_n \int_{-1}^{1} (1-t^2)^n \, dt = 1.$$

Note that by our assumptions on f, we may rewrite $L_n(x)$ as

$$L_n(x) = c_n \int_{-x}^{1-x} f(x+t)\,(1-t^2)^n \, dt = c_n \int_0^1 f(t)\,(1-(t-x)^2)^n \, dt.$$

Written this way, it's clear that L_n is a polynomial in x of degree at most n.

We first need to estimate c_n. An easy induction argument will convince you that $(1-t^2)^n \geq 1 - nt^2$, and so we get

$$\int_{-1}^{1} (1-t^2)^n \, dt \geq 2 \int_0^{1/\sqrt{n}} (1-nt^2) \, dt = \frac{4}{3\sqrt{n}} > \frac{1}{\sqrt{n}},$$

from which it follows that $c_n < \sqrt{n}$. In particular, for any $0 < \delta < 1$,

$$c_n \int_{\delta}^{1} (1-t^2)^n \, dt < \sqrt{n}\,(1-\delta^2)^n \to 0 \qquad (n \to \infty),$$

which is the inequality we'll need.

Next, let $\varepsilon > 0$ be given, and choose $0 < \delta < 1$ such that

$$|f(x) - f(y)| \leq \varepsilon/2 \text{ whenever } |x-y| \leq \delta.$$

Then, because $c_n(1-t^2)^n \geq 0$ and integrates to 1, we get

$$\begin{aligned}
|L_n(x) - f(x)| &= \left| c_n \int_{-1}^{1} \big[f(x+t) - f(x) \big] (1-t^2)^n \, dt \right| \\
&\leq c_n \int_{-1}^{1} |f(x+t) - f(x)| (1-t^2)^n \, dt \\
&\leq \frac{\varepsilon}{2}\, c_n \int_{-\delta}^{\delta} (1-t^2)^n \, dt + 4\|f\|\, c_n \int_{\delta}^{1} (1-t^2)^n \, dt \\
&\leq \frac{\varepsilon}{2} + 4\|f\|\,\sqrt{n}\,(1-\delta^2)^n < \varepsilon,
\end{aligned}$$

provided that n is sufficiently large. $\square$

A third proof of the Weierstrass theorem, due to Lebesgue in 1898, is outlined in the problems at the end of the chapter (see Problem 7). Lebesgue's proof is of historical interest because it inspired Stone's version of the Weierstrass theorem, which we'll discuss in Chapter 11.

Before we go on, let's stop and make an observation or two: While the Bernstein polynomials $B_n(f)$ offer a convenient and explicit polynomial approximation to f, they are by no means the best approximations. Indeed, recall that if $f_1(x) = x$ and $f_2(x) = x^2$, then $B_n(f_2) = (1 - \frac{1}{n}) f_2 + \frac{1}{n} f_1 \neq f_2$. Clearly, the best approximation to f_2 out of $\mathcal{P}_n$ should be f_2 itself whenever $n \geq 2$. On the other hand, because we always have

$$E_n(f) \leq \|f - B_n(f)\| \qquad \text{(why?)},$$

a detailed understanding of Bernstein's proof will lend insight into the general problem of polynomial approximation. Our next project, then, is to improve upon our estimate of the error $\|f - B_n(f)\|$.

Improved Estimates

To begin, we will need a bit more notation. The *modulus of continuity* of a bounded function f on the interval $[\,a,b\,]$ is defined by

$$\omega_f(\delta) = \omega_f([\,a,b\,];\delta) = \sup\big\{\, |f(x)-f(y)| : x,y \in [\,a,b\,],\ |x-y| \le \delta \,\big\}$$

for any $\delta > 0$. Note that $\omega_f(\delta)$ is a measure of the "ε" that goes along with δ (in the definition of uniform continuity); literally, we have written $\varepsilon = \omega_f(\delta)$ as a function of δ.

Here are a few easy facts about the modulus of continuity:

Exercise 2.4.

1. We always have $|f(x)-f(y)| \le \omega_f(\,|x-y|\,)$ for any $x \ne y \in [\,a,b\,]$.
2. If $0 < \delta' \le \delta$, then $\omega_f(\delta') \le \omega_f(\delta)$.
3. f is *uniformly continuous* if and only if $\omega_f(\delta) \to 0$ as $\delta \to 0^+$. [Hint: The statement that $|f(x)-f(y)| \le \varepsilon$ whenever $|x-y| \le \delta$ is equivalent to the statement that $\omega_f(\delta) \le \varepsilon$.]
4. If $f\,'$ exists and is bounded on $[\,a,b\,]$, then $\omega_f(\delta) \le K\delta$ for some constant K.
5. We say that f satisfies a *Lipschitz condition* of order α with constant K, where $0 < \alpha \le 1$ and $0 \le K < \infty$, if $|f(x)-f(y)| \le K|x-y|^\alpha$ for all x, y. We abbreviate this statement by writing: $f \in \operatorname{lip}_K\alpha$. Check that if $f \in \operatorname{lip}_K\alpha$, then $\omega_f(\delta) \le K\delta^\alpha$ for all $\delta > 0$.

For the time being, we actually need only one simple fact about $\omega_f(\delta)$:

Lemma 2.5. *Let f be a bounded function on $[\,a,b\,]$ and let $\delta > 0$. Then, $\omega_f(n\delta) \le n\,\omega_f(\delta)$ for $n = 1,2,\ldots$. Consequently, $\omega_f(\lambda\delta) \le (1+\lambda)\,\omega_f(\delta)$ for any $\lambda > 0$.*

Proof. Given $x < y$ with $|x-y| \le n\,\delta$, split the interval $[\,x,y\,]$ into n pieces, each of length at most δ. Specifically, if we set $z_k = x + k(y-x)/n$, for $k = 0,1,\ldots,n$, then $|z_k - z_{k-1}| \le \delta$ for any $k \ge 1$, and so

$$\begin{aligned}
|f(x)-f(y)| &= \left|\sum_{k=1}^{n} f(z_k) - f(z_{k-1})\right| \\
&\le \sum_{k=1}^{n} |f(z_k) - f(z_{k-1})| \\
&\le n\,\omega_f(\delta).
\end{aligned}$$

Thus, $\omega_f(n\delta) \le n\,\omega_f(\delta)$.

The second assertion follows from the first (and one of our exercises). Given $\lambda > 0$, choose an integer n so that $n-1 < \lambda \le n$. Then,

$$\omega_f(\lambda\delta) \le \omega_f(n\,\delta) \le n\,\omega_f(\delta) \le (1+\lambda)\,\omega_f(\delta). \qquad \square$$

We next repeat the proof of Bernstein's theorem, making a few minor adjustments here and there.

Theorem 2.6. *For any bounded function f on $[0,1]$ we have*

$$\|f - B_n(f)\| \leq \frac{3}{2}\,\omega_f\left(\frac{1}{\sqrt{n}}\right).$$

In particular, if $f \in C[0,1]$, then $E_n(f) \leq \frac{3}{2}\,\omega_f(\frac{1}{\sqrt{n}}) \to 0$ as $n \to \infty$.

Proof. We first do some term juggling:

$$\begin{aligned}
|f(x) - B_n(f)(x)| &= \left|\sum_{k=0}^{n}\left(f(x) - f\left(\frac{k}{n}\right)\right)\binom{n}{k}x^k(1-x)^{n-k}\right| \\
&\leq \sum_{k=0}^{n}\left|f(x) - f\left(\frac{k}{n}\right)\right|\binom{n}{k}x^k(1-x)^{n-k} \\
&\leq \sum_{k=0}^{n}\omega_f\left(\left|x - \frac{k}{n}\right|\right)\binom{n}{k}x^k(1-x)^{n-k} \\
&\leq \omega_f\left(\frac{1}{\sqrt{n}}\right)\sum_{k=0}^{n}\left[1 + \sqrt{n}\left|x - \frac{k}{n}\right|\right]\binom{n}{k}x^k(1-x)^{n-k} \\
&= \omega_f\left(\frac{1}{\sqrt{n}}\right)\left[1 \;+\; \sqrt{n}\sum_{k=0}^{n}\left|x - \frac{k}{n}\right|\binom{n}{k}x^k(1-x)^{n-k}\right],
\end{aligned}$$

where the third inequality follows from Lemma 2.5 (by taking $\lambda = \sqrt{n}\,\left|x - \frac{k}{n}\right|$ and $\delta = \frac{1}{\sqrt{n}}$). All that remains is to estimate the sum, and for this we'll use Cauchy-Schwarz (and our earlier observations about Bernstein polynomials). Because each of the terms $\binom{n}{k}x^k(1-x)^{n-k}$ is nonnegative, we have

$$\begin{aligned}
&\sum_{k=0}^{n}\left|x - \frac{k}{n}\right|\binom{n}{k}x^k(1-x)^{n-k} \\
&\qquad = \sum_{k=0}^{n}\left|x - \frac{k}{n}\right|\left[\binom{n}{k}x^k(1-x)^{n-k}\right]^{1/2}\cdot\left[\binom{n}{k}x^k(1-x)^{n-k}\right]^{1/2} \\
&\qquad \leq \left[\sum_{k=0}^{n}\left|x - \frac{k}{n}\right|^2\binom{n}{k}x^k(1-x)^{n-k}\right]^{1/2}\cdot\left[\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}\right]^{1/2} \\
&\qquad \leq \left[\frac{1}{4n}\right]^{1/2} = \frac{1}{2\sqrt{n}}.
\end{aligned}$$

Finally,

$$|f(x) - B_n(f)(x)| \leq \omega_f\left(\frac{1}{\sqrt{n}}\right)\left[1 \;+\; \sqrt{n}\cdot\frac{1}{2\sqrt{n}}\right] = \frac{3}{2}\,\omega_f\left(\frac{1}{\sqrt{n}}\right). \qquad \square$$

Examples 2.7.

1. If $f \in \mathrm{lip}_K\alpha$, it follows that $\|f - B_n(f)\| \leq \frac{3}{2}Kn^{-\alpha/2}$ and hence that $E_n(f) \leq \frac{3}{2}Kn^{-\alpha/2}$.

2. As a particular case of the first example, consider $f(x) = \left|x - \frac{1}{2}\right|$ on $[\,0,1\,]$. Then $f \in \text{lip}_1 1$, and so $\|f - B_n(f)\| \le \frac{3}{2}\, n^{-1/2}$. But, as Rivlin points out (see Remark 3 on p. 16 of [45]), $\|f - B_n(f)\| > \frac{1}{2}\, n^{-1/2}$. Thus, we can't hope to improve on the power of n in this estimate. Nevertheless, we will see an improvement in our estimate of $E_n(f)$.

The Bohman-Korovkin Theorem

The real value to us in Bernstein's approach is that the map $f \mapsto B_n(f)$, while providing a simple formula for an approximating polynomial, is also *linear* and *positive*. In other words,

$$\begin{aligned}
&B_n(f+g) = B_n(f) + B_n(g),\\
&B_n(\alpha f) = \alpha B_n(f),\ \ \alpha \in \mathbb{R},\\
&\text{and}\\
&B_n(f) \ge 0 \text{ whenever } f \ge 0.
\end{aligned}$$

(See Problem 15 for more on this.) As it happens, any positive, linear map $T : C[\,0,1\,] \to C[\,0,1\,]$ is automatically continuous!

Lemma 2.8. *If $T : C[\,a,b\,] \to C[\,a,b\,]$ is both positive and linear, then T is continuous.*

Proof. First note that a positive, linear map is also *monotone*. That is, T satisfies $T(f) \le T(g)$ whenever $f \le g$. (Why?) Thus, for any $f \in C[\,a,b\,]$, we have

$$-f,\, f \le |f| \implies -T(f),\, T(f) \le T(|f|);$$

that is, $|T(f)| \le T(|f|)$. But now $|f| \le \|f\| \cdot \mathbf{1}$, where $\mathbf{1}$ denotes the constant 1 function, and so we get

$$|T(f)| \le T(|f|) \le \|f\|\, T(\mathbf{1}).$$

Thus,

$$\|T(f)\| \le \|f\|\, \|T(\mathbf{1})\|$$

for any $f \in C[\,a,b\,]$. Finally, because T is linear, it follows that T is Lipschitz with constant $\|T(\mathbf{1})\|$:

$$\|T(f) - T(g)\| = \|T(f-g)\| \le \|T(\mathbf{1})\|\, \|f-g\|.$$

Consequently, T is continuous. $\square$

Now positive, linear maps abound in analysis, so this is a fortunate turn of events. What's more, Bernstein's theorem generalizes very nicely when placed in this new setting. The following elegant theorem was proved (independently) by Bohman and Korovkin in, roughly, 1952.

Theorem 2.9. *Let $T_n : C[\,0,1\,] \to C[\,0,1\,]$ be a sequence of positive, linear maps, and suppose that $T_n(f) \to f$ uniformly in each of the three cases*

$$f_0(x) = 1, \quad f_1(x) = x, \quad \text{and} \quad f_2(x) = x^2.$$

Then, $T_n(f) \to f$ uniformly for every $f \in C[\,0,1\,]$.

The proof of the Bohman-Korovkin theorem is essentially identical to the proof of Bernstein's theorem except, of course, we write $T_n(f)$ in place of $B_n(f)$. For full details, see [12]. Rather than proving the theorem, let's settle for a quick application.

Example 2.10. Let $f \in C[\,0,1\,]$ and, for each n, let $L_n(f)$ be the polygonal approximation to f with nodes at k/n, $k = 0, 1, \ldots, n$. That is, $L_n(f)$ is linear on each subinterval $[\,(k-1)/n,\, k/n\,]$ and agrees with f at each of the endpoints: $L_n(f)(k/n) = f(k/n)$. Then $L_n(f) \to f$ uniformly for each $f \in C[\,0,1\,]$. This is actually an easy calculation all by itself, but let's see why the Bohman-Korovkin theorem makes short work of it.

That $L_n(f)$ is positive and linear is (nearly) obvious; that $L_n(f_0) = f_0$ and $L_n(f_1) = f_1$ are really easy because, in fact, $L_n(f) = f$ for any linear function f. We just need to show that $L_n(f_2) \rightrightarrows f_2$. But a picture will convince you that the maximum distance between $L_n(f_2)$ and f_2 on the interval $[\,(k-1)/n,\, k/n\,]$ is at most

$$\left(\frac{k}{n}\right)^2 - \left(\frac{k-1}{n}\right)^2 = \frac{2k-1}{n^2} \le \frac{2}{n}.$$

That is, $\|f_2 - L_n(f_2)\| \le 2/n \to 0$ as $n \to \infty$.

[Note that L_n is a linear projection from $C[\,0,1\,]$ onto the subspace of polygonal functions based on the nodes k/n, $k = 0, \ldots, n$. An easy calculation, similar in spirit to the example above, will show that $\|f - L_n(f)\| \le 2\,\omega_f(1/n) \to 0$ as $n \to \infty$ for any $f \in C[\,0,1\,]$. See Problem 8.]

Problems

▷∗ 1. Define $\sigma : [\,0,1\,] \to [\,a,b\,]$ by $\sigma(t) = a + t(b-a)$ for $0 \le t \le 1$, and define a transformation $T_\sigma : C[\,a,b\,] \to C[\,0,1\,]$ by $(T_\sigma(f))(t) = f(\sigma(t))$. Prove that T_σ satisfies:

   (a) $T_\sigma(f+g) = T_\sigma(f) + T_\sigma(g)$ and $T_\sigma(cf) = c\,T_\sigma(f)$ for $c \in \mathbb{R}$.

   (b) $T_\sigma(fg) = T_\sigma(f)\,T_\sigma(g)$. In particular, T_σ maps polynomials to polynomials.

   (c) $T_\sigma(f) \le T_\sigma(g)$ if and only if $f \le g$.

   (d) $\|T_\sigma(f)\| = \|f\|$.

   (e) T_σ is both one-to-one and onto. Moreover, $(T_\sigma)^{-1} = T_{\sigma^{-1}}$.

▷∗ 2. Bernstein's Theorem shows that the polynomials are *dense* in $C[\,0,1\,]$. Using the results in Problem 1, conclude that the polynomials are also dense in $C[\,a,b\,]$.

▷∗ 3. How do we know that there are non-polynomial elements in $C[\,0,1\,]$? In other words, is it possible that every element of $C[\,0,1\,]$ agrees with some polynomial on $[\,0,1\,]$?

4. Let (Q_n) be a sequence of polynomials of degree m_n, and suppose that (Q_n) converges uniformly to f on $[\,a,b\,]$, where f is *not* a polynomial. Show that $m_n \to \infty$.

5. Use induction to show that $(1+x)^n \ge 1 + nx$, for all $n = 1, 2, \ldots$, whenever $x \ge -1$. Conclude that $(1-t^2)^n \ge 1 - nt^2$ whenever $-1 \le t \le 1$.

A *polygonal function* is a piecewise linear, continuous function; that is, a continuous function $f : [\,a,b\,] \to \mathbb{R}$ is a polygonal function if there are finitely many distinct points $a = x_0 < x_1 < \cdots < x_n = b$, called *nodes*, such that f is linear on each of the intervals $[\,x_{k-1}, x_k\,]$, $k = 1, \ldots, n$.

Fix distinct points $a = x_0 < x_1 < \cdots < x_n = b$ in $[\,a,b\,]$, and let S_n denote the set of all polygonal functions having nodes at the x_k. It's not hard to see that S_n is a vector space. In fact, it's relatively clear that S_n must have dimension exactly $n+1$ as there are $n+1$ "degrees of freedom" (each element of S_n is completely determined by its values at the x_k). More convincing, perhaps, is the fact that we can easily display a *basis* for S_n. (see Natanson [41]).

∗ 6. (a) Show that S_n is an $(n+1)$-dimensional subspace of $C[\,a,b\,]$ spanned by the constant function $\varphi_0(x) = 1$ and the "angles" $\varphi_{k+1}(x) = |x - x_k| + (x - x_k)$ for $k = 0, \ldots, n-1$. Specifically, show that each $h \in S_n$ can be uniquely written as $h(x) = c_0 + \sum_{i=1}^n c_i \varphi_i(x)$. [Hint: Because each side of the equation is an element of S_n, it's enough to show that the system of equations $h(x_0) = c_0$ and $h(x_k) = c_0 + 2\sum_{i=1}^k c_i(x_k - x_{i-1})$ for $k = 1, \ldots, n$ can be solved (uniquely) for the c_i.]

   (b) Each element of S_n can be written as $\sum_{i=1}^{n-1} a_i |x - x_i| + bx + d$ for some choice of scalars $a_1, \ldots, a_{n-1}$, b, d.

∗ 7. Given $f \in C[\,0,1\,]$, show that f can be uniformly approximated by a polygonal function. Specifically, given a positive integer n, let $L_n(x)$ denote the unique polygonal function with nodes at $(k/n)_{k=0}^n$ that agrees with f at each of these nodes. Show that $\|f - L_n\|$ is small provided that n is sufficiently large.

8. (a) Let f be in $\mathrm{lip}_C 1$; that is, suppose that f satisfies $|f(x) - f(y)| \le C|x - y|$ for some constant C and all x, y in $[0,1]$. In the notation of Problem 7, show that $\|f - L_n\| \le 2\,C/n$. [Hint: Given x in $[k/n, (k+1)/n)$, check that $|f(x) - L_n(x)| = |f(x) - f(k/n) + L_n(k/n) - L_n(x)| \le |f(x) - f(k/n)| + |f((k+1)/n) - f(k/n)|$.]

   (b) More generally, prove that $\|f - L_n(f)\| \le 2\,\omega_f(1/n) \to 0$ as $n \to \infty$ for any $f \in C[0,1]$.

In light of the results in Problems 6 and 7, Lebesgue noted that he could fashion a proof of Weierstrass's Theorem provided he could prove that $|x - c|$ can be uniformly approximated by polynomials on any interval $[a,b]$. (Why is this enough?) But thanks to the result in Problem 1, for this we need only show that $|x|$ can be uniformly approximated by polynomials on the interval $[-1,1]$.

$*$ 9. Here's an elementary proof that there is a sequence of polynomials (P_n) converging uniformly to $|x|$ on $[-1,1]$.

   (a) Define (P_n) recursively by $P_{n+1}(x) = P_n(x) + [x - P_n(x)^2]/2$, where $P_0(x) = 0$. Clearly, each P_n is a polynomial.

   (b) Check that $0 \le P_n(x) \le P_{n+1}(x) \le \sqrt{x}$ for $0 \le x \le 1$. Use Dini's theorem to conclude that $P_n(x) \rightrightarrows \sqrt{x}$ on $[0,1]$.

   (c) $P_n(x^2)$ is also a polynomial, and $P_n(x^2) \rightrightarrows |x|$ on $[-1,1]$.

10. If $f \in C[-1,1]$ is an *even* function, show that f may be uniformly approximated by even polynomials (that is, polynomials of the form $\sum_{k=0}^n a_k x^{2k}$).

11. If $f \in C[0,1]$ and if $f(0) = f(1) = 0$, show that the sequence of polynomials $\sum_{k=0}^n \left[\binom{n}{k} f(k/n)\right] x^k (1-x)^{n-k}$ having *integer* coefficients converges uniformly to f (where $[x]$ denotes the greatest integer in x). The same trick works for any $f \in C[a,b]$ provided that $0 < a < b < 1$.

12. If p is a polynomial and $\varepsilon > 0$, prove that there is a polynomial q with *rational* coefficients such that $\|p - q\| < \varepsilon$ on $[0,1]$. Conclude that $C[0,1]$ is *separable* (that is, $C[0,1]$ has a countable dense subset).

13. Let (x_i) be a sequence of numbers in $(0,1)$ such that $\lim\limits_{n\to\infty} \frac{1}{n}\sum_{i=1}^n x_i^k$ exists for every $k = 0, 1, 2, \ldots$. Show that $\lim\limits_{n\to\infty} \frac{1}{n}\sum_{i=1}^n f(x_i)$ exists for every $f \in C[0,1]$.

14. If $f \in C[0,1]$ and if $\int_0^1 x^n f(x)\,dx = 0$ for each $n = 0, 1, 2, \ldots$, show that $f \equiv 0$. [Hint: Using the Weierstrass theorem, show that $\int_0^1 f^2 = 0$.]

$\triangleright$ 15. Show that $|B_n(f)| \le B_n(|f|)$, and that $B_n(f) \ge 0$ whenever $f \ge 0$. Conclude that $\|B_n(f)\| \le \|f\|$.

16. If f is a bounded function on $[0,1]$, show that $B_n(f)(x) \to f(x)$ at each point of continuity of f.

17. Find $B_n(f)$ for $f(x) = x^3$. [Hint: $k^2 = (k-1)(k-2) + 3(k-1) + 1$.] The same method of calculation can be used to show that $B_n(f) \in \mathcal{P}_m$ whenever $f \in \mathcal{P}_m$ and $n > m$.

$*$ 18. Let f be continuously differentiable on $[a,b]$, and let $\varepsilon > 0$. Show that there is a polynomial p such that $\|f - p\| < \varepsilon$ *and* $\|f' - p'\| < \varepsilon$.

19. Suppose that $f \in C[\,a, b\,]$ is twice continuously differentiable and has $f'' > 0$. Prove that the best linear approximation to f on $[\,a, b\,]$ is $a_0 + a_1 x$ where $a_0 = f'(c)$, $a_1 = [f(a) + f(c) + f'(c)(a + c)]/2$, and where c is the unique solution to $f'(c) = (f(b) - f(a))/(b - a)$.

20. Let $f : [\,a, b\,] \to \mathbb{R}$ be a bounded function. Prove that

$$\omega_f([\,a, b\,]; \delta) = \sup\big\{\, \text{diam}\big(f(I)\big) : I \subset [\,a, b\,],\ \text{diam}(I) \le \delta \,\big\}$$

where I denotes a closed subinterval of $[\,a, b\,]$ and where $\text{diam}(A)$ denotes the diameter of the set A.

21. If the graph of $f : [\,a, b\,] \to \mathbb{R}$ has a jump of magnitude $\alpha > 0$ at some point x_0 in $[\,a, b\,]$, then $\omega_f(\delta) \ge \alpha$ for all $\delta > 0$.

22. Calculate ω_g for $g(x) = \sqrt{x}$.

23. If $f \in C[\,a, b\,]$, show that $\omega_f(\delta_1 + \delta_2) \le \omega_f(\delta_1) + \omega_f(\delta_2)$ and that $\omega_f(\delta) \downarrow 0$ as $\delta \downarrow 0$. Use this to show that ω_f is continuous for $\delta \ge 0$. Finally, show that the modulus of continuity of ω_f is again ω_f.

▷ 24. (a) Let $f : [-1, 1\,] \to \mathbb{R}$. If $x = \cos\theta$, where $-1 \le x \le 1$, and if $g(\theta) = f(\cos\theta)$, show that $\omega_g([-\pi, \pi\,], \delta) = \omega_g([\,0, \pi\,], \delta) \le \omega_f([-1, 1\,]; \delta)$.

    (b) If $h(x) = f(ax+b)$ for $c \le x \le d$, show that $\omega_h([\,c, d\,]; \delta) = \omega_f([\,ac+b, ad+b\,]; a\delta)$.

25. (a) Let f be continuously differentiable on $[\,0, 1\,]$. Show that $(B_n(f)')$ converges uniformly to f' by showing that $\|B_n(f') - (B_{n+1}(f))'\,\| \le \omega_{f'}(1/(n+1))$.

    (b) In order to see why this is of interest, find a uniformly convergent sequence of polynomials whose derivatives fail to converge uniformly.

    [Compare this result with Problem 18.]

Chapter 3

Trigonometric Polynomials

Introduction

A (real) *trigonometric polynomial*, or trig polynomial for short, is a function of the form

$$a_0 + \sum_{k=1}^{n} \left(a_k \cos kx + b_k \sin kx \right), \tag{3.1}$$

where $a_0, \ldots, a_n$ and $b_1, \ldots, b_n$ are real numbers. The *degree* of a trig polynomial is the highest frequency occurring in any representation of the form (3.1); thus, (3.1) has degree n provided that one of a_n or b_n is nonzero. We will use $\mathcal{T}_n$ to denote the collection of trig polynomials of degree at most n, and $\mathcal{T}$ to denote the collection of all trig polynomials (i.e., the union of the $\mathcal{T}_n$ over all n).

It is convenient to take the space of all continuous 2π-periodic functions on $\mathbb{R}$ as the containing space for $\mathcal{T}_n$; a space we denote by $C^{2\pi}$. The space $C^{2\pi}$ has several equivalent descriptions. For one, it's obvious that $C^{2\pi}$ is a subspace of $C(\mathbb{R})$, the space of all continuous functions on $\mathbb{R}$. But we might also consider $C^{2\pi}$ as a subspace of $C[\,0, 2\pi\,]$ in the following way: The 2π-periodic continuous functions on $\mathbb{R}$ may be identified with the set of functions $f \in C[\,0, 2\pi\,]$ satisfying $f(0) = f(2\pi)$. Each such f extends to a 2π-periodic element of $C(\mathbb{R})$ in an obvious way, and it's not hard to see that the condition $f(0) = f(2\pi)$ defines a (closed) subspace of $C[\,0, 2\pi\,]$. As a third description, it is often convenient to identify $C^{2\pi}$ with the collection $C(\mathbb{T})$, consisting of all continuous real-valued functions on $\mathbb{T}$, where $\mathbb{T}$ denotes the unit circle in the complex plane $\mathbb{C}$. That is, we simply make the identifications

$$\theta \longleftrightarrow e^{i\theta} \qquad \text{and} \qquad f(\theta) \longleftrightarrow f(e^{i\theta}).$$

In any case, each $f \in C^{2\pi}$ is uniformly continuous and uniformly bounded on all of $\mathbb{R}$, and is completely determined by its values on any interval of length 2π. In particular, we may (and will) endow $C^{2\pi}$ with the sup norm:

$$\|f\| = \max_{0 \le x \le 2\pi} |f(x)| = \max_{x \in \mathbb{R}} |f(x)|.$$

Our goal in this chapter is to prove what is sometimes called *Weierstrass's second theorem* (also from 1885).

Theorem 3.1. (Weierstrass's Second Theorem, 1885) *Let $f \in C^{2\pi}$. Then, for every $\varepsilon > 0$, there exists a trig polynomial T such that $\|f - T\| < \varepsilon$.*

Ultimately, we will give several different proofs of this theorem. Weierstrass gave a separate proof of this result in the same paper containing his theorem on approximation by algebraic polynomials, but it was later pointed out by Lebesgue [38] that the two theorems are, in fact, equivalent. Lebesgue's proof is based on several elementary observations. We will outline these elementary facts, supplying a few proofs here and there, but leaving full details to the reader.

We first justify the use of the word "polynomial" in the phrase "trig polynomial."

Lemma 3.2. *$\cos nx$ and $\sin(n+1)x/\sin x$ can be written as polynomials of degree exactly n in $\cos x$ for any integer $n \ge 0$.*

Proof. Using the recurrence formula

$$\cos kx + \cos(k-2)x = 2\cos(k-1)x\cos x$$

it's not hard to see that $\cos 2x = 2\cos^2 x - 1$, $\cos 3x = 4\cos^3 x - 3\cos x$, and $\cos 4x = 8\cos^4 x - 8\cos^2 x + 1$. More generally, by induction, $\cos nx$ is a polynomial of degree n in $\cos x$ with leading coefficient 2^{n-1}. Using this fact and the identity

$$\sin(k+1)x - \sin(k-1)x = 2\cos kx\,\sin x$$

(along with another easy induction argument), it follows that $\sin(n+1)x$ can be written as $\sin x$ times a polynomial of degree n in $\cos x$ with leading coefficient 2^n.

Alternatively, notice that by writing $(i\sin x)^{2k} = (\cos^2 x - 1)^k$ we have

$$\begin{aligned}
\cos nx &= \operatorname{Re}\left[(\cos x + i\sin x)^n\right] = \operatorname{Re}\left[\sum_{k=0}^{n}\binom{n}{k}(i\sin x)^k\cos^{n-k} x\right] \\
&= \sum_{k=0}^{[n/2]}\binom{n}{2k}(\cos^2 x - 1)^k\cos^{n-2k} x.
\end{aligned}$$

The coefficient of $\cos^n x$ in this expansion is then

$$\sum_{k=0}^{[n/2]}\binom{n}{2k} = \frac{1}{2}\sum_{k=0}^{n}\binom{n}{k} = 2^{n-1}.$$

(The sum of all the binomial coefficients is $(1+1)^n = 2^n$, but the even or odd terms, taken separately, sum to exactly half this amount because $(1+(-1))^n = 0$.)

Similarly,

$$\begin{aligned}
\sin(n+1)x &= \operatorname{Im}\left[(\cos x + i\sin x)^{n+1}\right] \\
&= \operatorname{Im}\left[\sum_{k=0}^{n+1}\binom{n+1}{k}(i\sin x)^k\cos^{n+1-k} x\right] \\
&= \sum_{k=0}^{[n/2]}\binom{n+1}{2k+1}(\cos^2 x - 1)^k\cos^{n-2k} x\sin x,
\end{aligned}$$

where we've written $(i \sin x)^{2k+1} = i(\cos^2 x - 1)^k \sin x$. The coefficient of $\cos^n x \sin x$ is

$$\sum_{k=0}^{[n/2]} \binom{n+1}{2k+1} = \frac{1}{2} \sum_{k=0}^{n+1} \binom{n+1}{k} = 2^n. \qquad \square$$

Corollary 3.3. *Any real trig polynomial (3.1) may be written as $P(\cos x) + Q(\cos x) \sin x$, where P and Q are algebraic polynomials of degree at most n and $n-1$, respectively. If the sum (3.1) represents an even function, then it can be written using only cosines.*

Corollary 3.4. *The collection $\mathcal{T}$, consisting of all trig polynomials, is both a subspace and a subring of $C^{2\pi}$* (that is, $\mathcal{T}$ is closed under both linear combinations and products). *In other words, $\mathcal{T}$ is a* subalgebra *of $C^{2\pi}$.*

It's not hard to see that the procedure we've described above can be reversed; that is, each algebraic polynomial in $\cos x$ and $\sin x$ can be written in the form (3.1). For example, $4\cos^3 x = 3\cos x + \cos 3x$. But, rather than duplicate our efforts, let's use a bit of linear algebra. First, the $2n+1$ functions

$$\mathcal{A} = \{\, 1,\ \cos x,\ \cos 2x,\ \ldots,\ \cos nx,\ \sin x,\ \sin 2x,\ \ldots,\ \sin nx \,\},$$

are *linearly independent*; the easiest way to see this is to notice that we may define an *inner product* on $C^{2\pi}$ under which these functions are *orthogonal*. Specifically,

$$\langle f, g \rangle = \int_0^{2\pi} f(x)\, g(x)\, dx = 0, \qquad \langle f, f \rangle = \int_0^{2\pi} f(x)^2\, dx \neq 0$$

for any pair of functions $f \neq g \in \mathcal{A}$. (We'll pursue this direction in greater detail later in the course.) Second, we've shown that each element of $\mathcal{A}$ lies in the space spanned by the $2n+1$ functions

$$\mathcal{B} = \{\, 1,\ \cos x,\ \cos^2 x,\ \ldots,\ \cos^n x,\ \sin x,\ \cos x \sin x,\ \ldots,\ \cos^{n-1} x \sin x \,\}.$$

That is,

$$\mathcal{T}_n \ \equiv\ \operatorname{span} \mathcal{A} \ \subset\ \operatorname{span} \mathcal{B}.$$

By comparing dimensions, we have

$$2n+1 = \dim \mathcal{T}_n = \dim(\operatorname{span} \mathcal{A}) \le \dim(\operatorname{span} \mathcal{B}) \le 2n+1,$$

and hence we must have $\operatorname{span} \mathcal{A} = \operatorname{span} \mathcal{B}$. The point here is that $\mathcal{T}_n$ is a finite-dimensional subspace of $C^{2\pi}$ of dimension $2n+1$, and we may use either one of these sets of functions as a basis for $\mathcal{T}_n$.

Before we leave these issues behind, let's summarize the situation for *complex* trig polynomials; i.e., the case where we allow complex coefficients in equation (3.1). To begin, it's clear that every sum of the form (3.1), whether real or complex coefficients are used, can be written as

$$\sum_{k=-n}^{n} c_k e^{ikx}, \tag{3.2}$$

where the c_k are *complex*; that is, a trig polynomial is actually a polynomial (over $\mathbb{C}$) in $z = e^{ix}$ and $\bar{z} = e^{-ix}$:

$$\sum_{k=-n}^{n} c_k e^{ikx} = \sum_{k=-n}^{n} c_k z^k = \sum_{k=0}^{n} c_k z^k + \sum_{k=1}^{n} c_{-k} \bar{z}^k.$$

Conversely, every sum of the form (3.2) can be written in the form (3.1), using *complex* a_k and b_k. Thus, the complex trig polynomials of degree at most n form a vector space of dimension $2n+1$ over $\mathbb{C}$ (hence of dimension $2(2n+1)$ when considered as a vector space over $\mathbb{R}$).

But, not every polynomial in z and $\bar{z}$ represents a *real* trig polynomial. Rather, the real trig polynomials are the *real parts* of the complex trig polynomials. To see this, notice that (3.2) represents a real-valued function if and only if

$$\sum_{k=-n}^{n} c_k e^{ikx} = \overline{\sum_{k=-n}^{n} c_k e^{ikx}} = \sum_{k=-n}^{n} \bar{c}_{-k} e^{ikx};$$

that is, we must have $c_k = \bar{c}_{-k}$ for each k. In particular, c_0 must be real, and hence

$$\begin{aligned}\sum_{k=-n}^{n} c_k e^{ikx} &= c_0 + \sum_{k=1}^{n}(c_k e^{ikx} + c_{-k} e^{-ikx}) \\ &= c_0 + \sum_{k=1}^{n}(c_k e^{ikx} + \bar{c}_k e^{-ikx}) \\ &= c_0 + \sum_{k=1}^{n}\big[\,(c_k + \bar{c}_k)\cos kx + i(c_k - \bar{c}_k)\sin kx\,\big] \\ &= c_0 + \sum_{k=1}^{n}\big[\,2\,\mathrm{Re}(c_k)\cos kx - 2\,\mathrm{Im}(c_k)\sin kx\,\big],\end{aligned}$$

which is of the form (3.1) with a_k and b_k real.

Conversely, given any real trig polynomial (3.1), we have

$$a_0 + \sum_{k=1}^{n}\big(\,a_k\cos kx + b_k \sin kx\,\big) = a_0 + \sum_{k=1}^{n}\left[\,\left(\frac{a_k - ib_k}{2}\right)e^{ikx} + \left(\frac{a_k + ib_k}{2}\right)e^{-ikx}\,\right],$$

which of of the form (3.2) with $c_k = \bar{c}_{-k}$ for each k.

It's time we returned to approximation theory! Because we've been able to identify $C^{2\pi}$ with a subspace of $C[\,0, 2\pi\,]$, and because $\mathcal{T}_n$ is a finite-dimensional subspace of $C^{2\pi}$, we have

Corollary 3.5. *Each $f \in C^{2\pi}$ has a best approximation* (on all of $\mathbb{R}$) *out of $\mathcal{T}_n$. If f is an even function, then it has a best approximation which is also even.*

Proof. We only need to prove the second claim, so suppose that $f \in C^{2\pi}$ is even and that $T^* \in \mathcal{T}_n$ satisfies

$$\|f - T^*\| = \min_{T\in\mathcal{T}_n} \|f - T\|.$$

Then, because f is even, $\widetilde{T}(x) = T^*(-x)$ is also a best approximation to f out of $\mathcal{T}_n$; indeed,

$$\begin{aligned}\|f - \widetilde{T}\,\| &= \max_{x\in\mathbb{R}} |f(x) - T^*(-x)| \\ &= \max_{x\in\mathbb{R}} |f(-x) - T^*(x)| \\ &= \max_{x\in\mathbb{R}} |f(x) - T^*(x)| = \|f - T^*\|.\end{aligned}$$

But now, the *even* trig polynomial

$$\widehat{T}(x) = \frac{\widetilde{T}(x) + T^*(x)}{2} = \frac{T^*(-x) + T^*(x)}{2}$$

is also a best approximation out of $\mathcal{T}_n$ because

$$\|f - \widehat{T}\,\| = \left\| \frac{(f - \widetilde{T}\,) + (f - T^*)}{2} \right\| \le \frac{\|f - \widetilde{T}\,\| + \|f - T^*\|}{2} = \min_{T \in \mathcal{T}_n} \|f - T\|. \qquad \square$$

Weierstrass's Second Theorem

We next give (de La Vallée Poussin's version of) Lebesgue's proof of Weierstrass's second theorem; specifically, we will deduce the second theorem from the first.

Theorem 3.6. *Let $f \in C^{2\pi}$ and let $\varepsilon > 0$. Then, there is a trig polynomial T such that $\|f - T\| = \max_{x \in \mathbb{R}} |f(x) - T(x)| < \varepsilon$.*

Proof. We will prove that Weierstrass's first theorem for $C[-1,1\,]$ implies his second theorem for $C^{2\pi}$.

Step 1. If $f \in C^{2\pi}$ is even, then f may be uniformly approximated by even trig polynomials.

If f is even, then it's enough to approximate f on the interval $[\,0, \pi\,]$. In this case, we may consider the function $g(y) = f(\arccos y)$, $-1 \le y \le 1$, in $C[-1,1\,]$. By Weierstrass's first theorem, there is an algebraic polynomial $p(y)$ such that

$$\max_{-1 \le y \le 1} |f(\arccos y) - p(y)| = \max_{0 \le x \le \pi} |f(x) - p(\cos x)| < \varepsilon.$$

But $T(x) = p(\cos x)$ is an even trig polynomial! Hence,

$$\|f - T\| = \max_{x \in \mathbb{R}} |f(x) - T(x)| < \varepsilon.$$

Let's agree to abbreviate $\|f - T\| < \varepsilon$ as $f \approx T + \varepsilon$.

Step 2. Given $f \in C^{2\pi}$, there is a trig polynomial T such that $2f(x)\sin^2 x \approx T(x) + 2\varepsilon$.

Each of the functions $f(x) + f(-x)$ and $[f(x) - f(-x)]\sin x$ is *even*. Thus, we may choose even trig polynomials T_1 and T_2 such that

$$f(x) + f(-x) \approx T_1(x) \qquad \text{and} \qquad [f(x) - f(-x)]\sin x \approx T_2(x).$$

Multiplying the first expression by $\sin^2 x$, the second by $\sin x$, and adding, we get

$$2f(x)\sin^2 x \approx T_1(x)\sin^2 x + T_2(x)\sin x \equiv T_3(x),$$

where $T_3(x)$ is still a trig polynomial, and where $f \approx T_3 + 2\varepsilon$ because $|\sin x\,| \le 1$.

Step 3. Given $f \in C^{2\pi}$, there is a trig polynomial T such that $2f(x)\cos^2 x \approx T(x) + 2\varepsilon$.

Repeat Step 2 for $f(x - \pi/2)$ and translate: We first choose a trig polynomial $T_4(x)$ such that

$$2f\left(x - \frac{\pi}{2}\right)\sin^2 x \approx T_4(x).$$

That is,

$$2f(x)\cos^2 x \approx T_5(x),$$

where $T_5(x)$ is a trig polynomial.

Finally, by combining the conclusions of Steps 2 and 3, we find that there is a trig polynomial $T_6(x)$ such that $f \approx T_6(x) + 2\varepsilon$. □

Just for fun, let's complete the circle and show that Weierstrass's second theorem for $C^{2\pi}$ implies his first theorem for $C[-1,1]$. Because it's possible to give an independent proof of the second theorem, as we'll see later, this is a meaningful endeavor.

Theorem 3.7. *Given $f \in C[-1,1]$ and $\varepsilon > 0$, there exists an algebraic polynomial p such that $\|f - p\| < \varepsilon$.*

Proof. Given $f \in C[-1,1]$, the function $f(\cos x)$ is an even function in $C^{2\pi}$. By our proof of Weierstrass's second theorem (Step 1 of the proof), we may approximate $f(\cos x)$ by an even trig polynomial:

$$f(\cos x) \approx a_0 + a_1 \cos x + a_2 \cos 2x + \cdots + a_n \cos nx.$$

But, as we've seen, $\cos kx$ can be written as an algebraic polynomial in $\cos x$. Hence, there is some algebraic polynomial p such that $f(\cos x) \approx p(\cos x)$. That is,

$$\max_{0 \le x \le \pi} |f(\cos x) - p(\cos x)| = \max_{-1 \le t \le 1} |f(t) - p(t)| < \varepsilon. \qquad \square$$

The algebraic polynomials $T_n(x)$ satisfying

$$T_n(\cos x) = \cos nx, \text{ for } n = 0, 1, 2, \ldots,$$

are called the *Chebyshev polynomials of the first kind.* Please note that this formula *uniquely* defines T_n as a polynomial of degree exactly n (with leading coefficient 2^{n-1}), and hence uniquely determines the values of $T_n(x)$ for $|x| > 1$, too. The algebraic polynomials $U_n(x)$ satisfying

$$U_n(\cos x) = \frac{\sin(n+1)x}{\sin x}, \text{ for } n = 0, 1, 2, \ldots,$$

are called the *Chebyshev polynomials of the second kind.* Likewise, note that this formula uniquely defines U_n as a polynomial of degree exactly n (with leading coefficient 2^n).

We will discover many intriguing properties of the Chebyshev polynomials in the next chapter. For now, let's settle for just one: The recurrence formula we gave earlier

$$\cos nx = 2\cos x\, \cos(n-1)x - \cos(n-2)x$$

now becomes

$$T_n(x) = 2x\, T_{n-1}(x) - T_{n-2}(x), \quad n \ge 2,$$

where $T_0(x) = 1$ and $T_1(x) = x$. This recurrence relation (along with the initial cases T_0 and T_1) may be taken as a definition for the Chebyshev polynomials of the first kind. At any rate, it's now easy to list any number of the Chebyshev polynomials T_n; for example, the next few are $T_2(x) = 2x^2 - 1$, $T_3(x) = 4x^3 - 3x$, $T_4(x) = 8x^4 - 8x^2 + 1$, and $T_5(x) = 16x^5 - 20x^3 + 5x$.

Problems

▷ 1. (a) Use induction to show that $|\sin nx/\sin x| \le n$ for all $0 \le x \le \pi$. [Hint: $\sin(n+1)x = \sin nx \cos x + \sin x \cos nx$.]

(b) By examining the proof of (a), show that $|\sin nx/\sin x| = n$ can only occur when $x = 0$ or $x = \pi$.

2. Let $f \in C^{2\pi}$ be continuously differentiable and let $\varepsilon > 0$. Show that there is a trig polynomial T such that $\|f - T\| < \varepsilon$ *and* $\|f' - T'\| < \varepsilon$.

▷ 3. Establish the following properties of $T_n(x)$.

(i) Show that the zeros of $T_n(x)$ are real, simple, and lie in the open interval $(-1, 1)$.

(ii) We know that $|T_n(x)| \le 1$ for $-1 \le x \le 1$, but when does equality occur; that is, when is $|T_n(x)| = 1$ on $[-1, 1]$?

(iii) Evaluate $\displaystyle\int_{-1}^{1} T_n(x)\, T_m(x)\, \frac{dx}{\sqrt{1-x^2}}$.

(iv) Show that $T_n'(x) = nU_{n-1}(x)$.

(v) Show that T_n is a solution to $(1-x^2)y'' - xy' + n^2 y = 0$.

4. Find analogues of properties (i)–(vi) in Problem 3 (if possible) for $U_n(x)$, the Chebyshev polynomials of the second kind.

5. Show that the Chebyshev polynomials (T_n) are linearly independent over any interval $[\,a, b\,]$ in $\mathbb{R}$. Is the same true of the sequence (U_n)?

Chapter 4

Characterization of Best Approximation

Introduction

We next discuss Chebyshev's solution to the problem of best polynomial approximation from 1854. Given that there was no reason to believe that the problem even had a solution, let alone a unique solution, Chebyshev's accomplishment should not be underestimated. Chebyshev might very well have been able to prove Weierstrass's result—30 years early—had the thought simply occurred to him! Chebyshev's original papers are apparently rather sketchy. It wasn't until 1903 that full details were given by Kirchberger [34]. Curiously, Kirchberger's proofs foreshadow very modern techniques such as convexity and separation arguments. The presentation we'll give owes much to Haar and to de La Vallée Poussin (both from around 1918).

We begin with an easy observation:

Lemma 4.1. *Let $f \in C[\,a,b\,]$ and let $p = p_n^*$ be a best approximation to f out of $\mathcal{P}_n$. Then there are at least two distinct points $x_1,\ x_2 \in [\,a,b\,]$ such that*

$$f(x_1) - p(x_1) = -(f(x_2) - p(x_2)) = \|f - p\|.$$

That is, $f - p$ attains each of the values $\pm\|f - p\|$.

Proof. Let's write $E = E_n(f) = \|f - p\| = \max_{a \le x \le b} |f(x) - p(x)|$. If the conclusion of the Lemma is false, then we might as well suppose that $f(x_1) - p(x_1) = E$, for some x_1, but that

$$e = \min_{a \le x \le b} (f(x) - p(x)) > -E.$$

In particular, $E + e \neq 0$ and so $q = p + (E + e)/2$ is an element of $\mathcal{P}_n$ with $q \neq p$. We claim that q is a better approximation to f than p. Here's why:

$$E - \left(\frac{E+e}{2}\right) \;\ge\; f(x) - p(x) - \left(\frac{E+e}{2}\right) \;\ge\; e - \left(\frac{E+e}{2}\right),$$

or

$$\left(\frac{E-e}{2}\right) \;\ge\; f(x) - q(x) \;\ge\; -\left(\frac{E-e}{2}\right).$$

That is,

$$\|f-q\| \;\le\; \left(\frac{E-e}{2}\right) \;<\; E \;=\; \|f-p\|,$$

a contradiction. □

Corollary 4.2. *The best approximating constant to $f \in C[\,a,b\,]$ is*

$$p_0^* \;=\; \frac{1}{2}\left[\max_{a\le x\le b} f(x) + \min_{a\le x\le b} f(x)\right],$$

and

$$E_0(f) \;=\; \frac{1}{2}\left[\max_{a\le x\le b} f(x) - \min_{a\le x\le b} f(x)\right].$$

Proof. **Exercise**.

Now all of this is meant as motivation for the general case, which essentially repeats the observation of our first Lemma inductively. A little experimentation will convince you that a best linear approximation, for example, would imply the existence of *three* (or more) points at which $f-p_1^*$ alternates between $\pm\|f-p_1^*\|$.

A bit of notation will help us set up the argument for the general case: Given g in $C[\,a,b\,]$, we'll say that $x \in [\,a,b\,]$ is a $(+)$ *point* for g (respectively, a $(-)$ *point* for g) if $g(x) = \|g\|$ (respectively, $g(x) = -\|g\|$). A set of distinct point $a \le x_0 < x_1 < \cdots < x_n \le b$ will be called an *alternating set* for g if the x_i are alternately $(+)$ points and $(-)$ points; that is, if

$$|g(x_i)| = \|g\|, \qquad i = 0,1,\ldots,n,$$

and

$$g(x_i) = -g(x_{i-1}), \qquad i = 1,2,\ldots,n.$$

Using this notation, we will be able to characterize the polynomial of best approximation. Our first result is where all the fighting takes place:

Theorem 4.3. *Let $f \in C[\,a,b\,]$, and suppose that $p = p_n^*$ is a best approximation to f out of $\mathcal{P}_n$. Then, there is an alternating set for $f-p$ consisting of at least $n+2$ points.*

Proof. If $f \in \mathcal{P}_n$, there's nothing to show. (Why?) Thus, we may suppose that $f \notin \mathcal{P}_n$ and, hence, that $E = E_n(f) = \|f-p\| > 0$.

Now consider the (uniformly) continuous function $\varphi = f - p$. We may partition $[\,a,b\,]$ by way of $a = t_0 < t_1 < \cdots < t_k = b$ into sufficiently small intervals so that

$$|\varphi(x)-\varphi(y)| < E/2 \qquad \text{whenever} \qquad x,y \in [\,t_i,t_{i+1}\,].$$

Here's why we'd want to do such a thing: If $[\,t_i,t_{i+1}\,]$ contains a $(+)$ point for $\varphi = f-p$, then φ is positive on all of $[\,t_i,t_{i+1}\,]$. Indeed,

$$x,y \in [\,t_i,t_{i+1}\,] \;\text{ and }\; \varphi(x) = E \quad\Longrightarrow\quad \varphi(y) > E/2 > 0. \tag{4.1}$$

Similarly, if $[\,t_i,t_{i+1}\,]$ contains a $(-)$ point for φ, then φ is negative on all of $[\,t_i,t_{i+1}\,]$. Consequently, **no interval** $[\,t_i,t_{i+1}\,]$ **can contain both** $(+)$ **points and** $(-)$ **points**.

Call $[t_i, t_{i+1}]$ a (+) *interval* (respectively, a (−) *interval*) if it contains a (+) point (respectively, a (−) point) for $\varphi = f - p$. Notice that **no (+) interval can even touch a (−) interval**. In other words, a (+) interval and a (−) interval must be strictly separated (by some interval containing a zero for φ).

We now relabel the (+) and (−) intervals from left to right, ignoring the "neither" intervals. There's no harm in supposing that the first signed interval is a (+) interval. Thus, we suppose that our relabeled intervals are written

$$\begin{array}{ll} I_1, I_2, \ldots, I_{k_1} & (+) \text{ intervals,} \\ I_{k_1+1}, I_{k_1+2}, \ldots, I_{k_2} & (-) \text{ intervals,} \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ I_{k_{m-1}+1}, I_{k_1+2}, \ldots, I_{k_m} & (-1)^{m-1} \text{ intervals,} \end{array}$$

where I_{k_1} is the last (+) interval before we reach the first (−) interval, I_{k_1+1}. And so on.

For later reference, we let S denote the union of all the signed intervals; that is, $S = \bigcup_{j=1}^{k_m} I_j$, and we let N denote the union of all the "neither" intervals. Thus, S and N are compact sets with $S \cup N = [a, b]$ (note that while S and N aren't quite disjoint, they are at least nonoverlapping—their interiors are disjoint).

Our goal here is to show that $m \geq n + 2$. (So far we only know that $m \geq 2$!) Let's suppose that $m < n + 2$ and see what goes wrong.

Because any (+) interval is strictly separated from any (−) interval, we can find points $z_1, \ldots, z_{m-1} \in N$ such that

$$\begin{array}{c} \max I_{k_1} < z_1 < \min I_{k_1+1} \\ \max I_{k_2} < z_2 < \min I_{k_2+1} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \max I_{k_{m-1}} < z_{m-1} < \min I_{k_{m-1}+1} \end{array}$$

And now we construct the offending polynomial:

$$q(x) = (z_1 - x)(z_2 - x) \cdots (z_{m-1} - x).$$

Notice that $q \in \mathcal{P}_n$ because $m - 1 \leq n$. (Here is the *only* use we'll make of the assumption $m < n + 2$!) We're going to show that $p + \lambda q \in \mathcal{P}_n$ is a better approximation to f than p, for some suitable scalar λ.

We first claim that q and $f - p$ have the same sign. Indeed, q has no zeros in any of the (±) intervals, hence is of constant sign on any such interval. Thus, $q > 0$ on $I_1, \ldots, I_{k_1}$ because each $(z_j - x) > 0$ on these intervals; $q < 0$ on $I_{k_1+1}, \ldots, I_{k_2}$ because here $(z_1 - x) < 0$, while $(z_j - x) > 0$ for $j > 1$; and so on.

We next find λ. Let $e = \max_{x \in N} |f(x) - p(x)|$, where N is the union of all the subintervals $[t_i, t_{i+1}]$ which are neither (+) intervals nor (−) intervals. Then, $e < E$. (Why?) Now choose $\lambda > 0$ so that $\lambda \|q\| < \min\{E - e, E/2\}$. We claim that $p + \lambda q$ is a better approximation to f than p. One case is easy: If $x \in N$, then

$$|f(x) - (p(x) + \lambda q(x))| \;\leq\; |f(x) - p(x)| + \lambda |q(x)| \;\leq\; e + \lambda \|q\| \;<\; E.$$

On the other hand, if $x \notin N$, then x is in either a (+) interval or a (−) interval. In particular, from equation (4.1), we know that $|f(x) - p(x)| > E/2 > \lambda \|q\|$ and, hence, that $f(x) - p(x)$ and $\lambda q(x)$ have the same sign. Thus,

$$\begin{aligned} |f(x) - (p(x) + \lambda q(x))| &= |f(x) - p(x)| - \lambda |q(x)| \\ &\leq E - \lambda \min_{x \in S} |q(x)| \;<\; E, \end{aligned}$$

because q is nonzero on S. This contradiction finishes the proof. (Phew!) $\square$

Remarks 4.4.

1. It should be pointed out that the number $n+2$ here is actually $1 + \dim \mathcal{P}_n$.

2. Notice, too, that if $f - p_n^*$ alternates in sign $n+2$ times, then $f - p_n^*$ must have at least $n+1$ zeros. Thus, p_n^* actually agrees with f (or "interpolates" f) at $n+1$ points.

We're now ready to establish the uniqueness of the polynomial of best approximation. Because the norm in $C[\,a, b\,]$ is not strictly convex, this is a somewhat unexpected (but welcome!) result.

Theorem 4.5. *Let $f \in C[\,a, b\,]$. Then, the polynomial of best approximation to f out of $\mathcal{P}_n$ is unique.*

Proof. Suppose that p, $q \in \mathcal{P}_n$ both satisfy $\|f - p\| = \|f - q\| = E_n(f) = E$. Then, as we've seen, their average $r = (p+q)/2 \in \mathcal{P}_n$ is also best: $\|f - r\| = E$ because $f - r = (f-p)/2 + (f-q)/2$.

By Theorem 4.3, $f - r$ has an alternating set $x_0, x_1, \ldots, x_{n+1}$, containing $n+2$ points. Thus, for each i,

$$(f-p)(x_i) + (f-q)(x_i) = \pm 2E \qquad \text{(alternating)},$$

while

$$-E \le (f-p)(x_i),\ (f-q)(x_i) \le E.$$

But this means that

$$(f-p)(x_i) = (f-q)(x_i) = \pm E \qquad \text{(alternating)}$$

for each i. (Why?) That is, $x_0, x_1, \ldots, x_{n+1}$ is an alternating set for both $f - p$ and $f - q$. In particular, the polynomial $q - p = (f-p) - (f-q)$ has $n+2$ zeros! Because $q - p \in \mathcal{P}_n$, we must have $p = q$. $\square$

Finally, we come full circle:

Theorem 4.6. *Let $f \in C[\,a, b\,]$, and let $p \in \mathcal{P}_n$. If $f - p$ has an alternating set containing $n+2$ (or more) points, then p is the best approximation to f out of $\mathcal{P}_n$.*

Proof. Let $x_0, x_1, \ldots, x_{n+1}$ be an alternating set for $f - p$, and suppose that some $q \in \mathcal{P}_n$ is a better approximation to f than p; that is, $\|f - q\| < \|f - p\|$. In particular, then, we must have

$$|f(x_i) - p(x_i)| = \|f - p\| > \|f - q\| \ge |f(x_i) - q(x_i)|$$

for each $i = 0, 1, \ldots, n+1$. Now the inequality $|a| > |b|$ implies that a and $a - b$ have the same sign (why?), hence $q - p = (f-p) - (f-q)$ alternates in sign $n+2$ times (because $f - p$ does). But then, $q - p$ would have at least $n+1$ zeros. Because $q - p \in \mathcal{P}_n$, we must have $q = p$, which is a contradiction. Thus, p is the best approximation to f out of $\mathcal{P}_n$. $\square$

Example 4.7. (Rivlin [45]) While an alternating set for $f - p_n^*$ is supposed to have at least $n+2$ points, it may well have more than $n+2$ points; thus, alternating sets need not be unique. For example, consider the function $f(x) = \sin 4x$ on $[-\pi, \pi]$. Because there are 8 points where f alternates between ± 1, it follows that $p_0^* = 0$ and that there are $4 \times 4 = 16$ different alternating sets consisting of exactly 2 points (not to mention all those with more than 2 points). In addition, notice that we actually have $p_1^* = \cdots = p_6^* = 0$, but that $p_7^* \neq 0$. (Why?)

Exercise 4.8. Show that $y = x - 1/8$ is the best linear approximation to $y = x^2$ on $[0, 1]$.

Essentially repeating the proof given for Theorem 4.6 yields a lower bound for $E_n(f)$.

Theorem 4.9. (de La Vallée Poussin) *Let $f \in C[a, b]$, and suppose that $q \in \mathcal{P}_n$ is such that $f(x_i) - q(x_i)$ alternates in sign at $n+2$ points $a \le x_0 \le x_1 \le \ldots \le x_{n+1} \le b$. Then*

$$E_n(f) \;\ge\; \min_{i=0,\ldots,n+1} |f(x_i) - q(x_i)|.$$

Proof. If the inequality fails, then the best approximation $p = p_n^*$ would satisfy

$$\max_{0\le i\le n+1} |f(x_i) - p(x_i)| \;\le\; E_n(f) \;<\; \min_{0\le i\le n+1} |f(x_i) - q(x_i)|.$$

Now we could repeat (essentially) the same argument used in the proof of Theorem 4.6 to arrive at a contradiction. The details are left as an **exercise**. □

Even for relatively simple functions, the problem of actually finding the polynomial of best approximation is genuinely difficult (even computationally). We next discuss one such problem, and an important one, that Chebyshev was able to solve.

Problem 1. Find the polynomial $p_{n-1}^* \in \mathcal{P}_{n-1}^*$ of degree at most $n-1$ that best approximates $f(x) = x^n$ on the interval $[-1, 1]$. (This particular choice of interval makes for a tidy solution; we'll discuss the general situation later.)

Because p_{n-1}^* is to minimize $\max_{|x|\le 1} |x^n - p_{n-1}^*(x)|$, our first problem is equivalent to:

Problem 2. Find the monic polynomial of degree n which deviates least from 0 on $[-1, 1]$. In other words, find the monic polynomial of degree n of smallest norm in $C[-1, 1]$.

We'll give two solutions to this problem (which we know has a unique solution, of course). First, let's simplify our notation. We write

$$p(x) = x^n - p_{n-1}^*(x) \quad \text{(the solution)},$$

and

$$M = \|p\| = E_{n-1}(x^n; [-1, 1]).$$

All we know about p is that it has an alternating set $-1 \le x_0 < x_1 < \cdots < x_n \le 1$ containing $(n-1) + 2 = n+1$ points; that is, $|p(x_i)| = M$ and $p(x_{i+1}) = -p(x_i)$ for all i. Using this tiny bit of information, Chebyshev was led to compare the polynomials p^2 and p'. Watch closely!

Step 1. At any x_i in $(-1, 1)$, we must have $p'(x_i) = 0$ (because $p(x_i)$ is a relative extreme value for p). But, p' is a polynomial of degree $n-1$ and so can have at most $n-1$ zeros. Thus, we must have

$$x_i \in (-1, 1) \quad \text{and} \quad p'(x_i) = 0, \quad \text{for} \quad i = 1, \ldots, n-1,$$

(in fact, $x_1, \ldots, x_{n-1}$ are *all* the zeros of $p\,'$) and

$$x_0 = -1, \quad p\,'(x_0) \neq 0, \quad x_{n-1} = 1, \quad p\,'(x_{n-1}) \neq 0.$$

Step 2. Now consider the polynomial $M^2 - p^2 \in \mathcal{P}_{2n}$. We know that $M^2 - (p(x_i))^2 = 0$ for $i = 0, 1, \ldots, n$, and that $M^2 - p^2 \geq 0$ on $[-1, 1]$. Thus, $x_1, \ldots, x_{n-1}$ must be double roots (at least) of $M^2 - p^2$. But this makes for $2(n-1) + 2 = 2n$ roots already, so we must have them all. Hence, $x_1, \ldots, x_{n-1}$ are double roots, x_0 and x_n are simple roots, and these are all the roots of $M^2 - p^2$.

Step 3. Next consider $(p\,')^2 \in \mathcal{P}_{2(n-1)}$. We know that $(p\,')^2$ has a double root at each of $x_1, \ldots, x_{n-1}$ (and no other roots), hence $(1 - x^2)(p\,'(x))^2$ has a double root at each of $x_1, \ldots, x_{n-1}$, and simple roots at x_0 and x_n. Because $(1 - x^2)(p\,'(x))^2 \in \mathcal{P}_{2n}$, we've found all of its roots.

Here's the point to all this "rooting":

Step 4. Because $M^2 - (p(x))^2$ and $(1 - x^2)(p\,'(x))^2$ are polynomials of the same degree with the same roots, they are, up to a constant multiple, the same polynomial! It's easy to see what constant, too: The leading coefficient of p is 1 while the leading coefficient of $p\,'$ is n; thus,

$$M^2 - (p(x))^2 \;=\; \frac{(1 - x^2)(p\,'(x))^2}{n^2}.$$

After tidying up,

$$\frac{p\,'(x)}{\sqrt{M^2 - (p(x))^2}} \;=\; \frac{n}{\sqrt{1 - x^2}}.$$

We really should have an extra $\pm$ here, but we know that $p\,'$ is positive on *some* interval; we'll simply assume that it's positive on $[-1, x_1]$. Now, upon integrating,

$$\arccos\left(\frac{p(x)}{M}\right) \;=\; n \arccos x + C;$$

that is,

$$p(x) = M \cos(n \arccos x + C).$$

But $p(-1) = -M$ (because $p\,'(-1) \geq 0$), so

$$\begin{aligned} \cos(n\pi + C) = -1 \quad &\Longrightarrow \quad C = m\pi \quad (\text{with } n + m \text{ odd}) \\ &\Longrightarrow \quad p(x) = \pm M \cos(n \arccos x). \end{aligned}$$

Look familiar? Because we know that $\cos(n \arccos x)$ is a polynomial of degree n with leading coefficient 2^{n-1} (i.e., the n-th Chebyshev polynomial T_n), the solution to our problem is

$$p(x) \;=\; 2^{-n+1}\, T_n(x).$$

And because $|T_n(x)| \leq 1$ for $|x| \leq 1$ (why?), the minimum norm is $M = 2^{-n+1}$. □

Next we give a "fancy" solution, based on our characterization of best approximations (Theorem 4.6) and a few simple properties of the Chebyshev polynomials.

Theorem 4.10. *For any $n \geq 1$, the formula $p(x) = x^n - 2^{-n+1}\, T_n(x)$ defines a polynomial $p \in \mathcal{P}_{n-1}$ satisfying*

$$2^{-n+1} \;=\; \max_{|x| \leq 1} |x^n - p(x)| \;<\; \max_{|x| \leq 1} |x^n - q(x)|$$

for any other $q \in \mathcal{P}_{n-1}$.

Proof. We know that $2^{-n+1}T_n(x)$ has leading coefficient 1, and so $p \in \mathcal{P}_{n-1}$. Now set $x_k = \cos((n-k)\pi/n)$ for $k = 0, 1, \ldots, n$. Then, $-1 = x_0 < x_1 < \cdots < x_n = 1$ and

$$T_n(x_k) = T_n(\cos((n-k)\pi/n)) = \cos((n-k)\pi) = (-1)^{n-k}.$$

Because $|T_n(x)| = |T_n(\cos\theta)| = |\cos n\theta| \le 1$, for $-1 \le x \le 1$, we've found an alternating set for T_n containing $n+1$ points.

In other words, $x^n - p(x) = 2^{-n+1}T_n(x)$ satisfies $|x^n - p(x)| \le 2^{-n+1}$ and, for each $k = 0, 1, \ldots, n$, has $x_k^n - p(x_k) = 2^{-n+1}T_n(x_k) = (-1)^{n-k}2^{-n+1}$. By our characterization of best approximations (Theorem 4.6), p must be the best approximation to x^n out of $\mathcal{P}_{n-1}$. □

Corollary 4.11. *The monic polynomial of degree exactly n having smallest norm in $C[\,a,b\,]$ is*

$$\frac{(b-a)^n}{2^n 2^{n-1}} \cdot T_n\left(\frac{2x-b-a}{b-a}\right).$$

Proof. **Exercise**. [Hint: If $p(x)$ is a polynomial of degree n with leading coefficient 1, then $\tilde{p}(x) = p((2x-b-a)/(b-a))$ is a polynomial of degree n with leading coefficient $2^n/(b-a)^n$. Moreover, $\max_{a\le x\le b}|p(x)| = \max_{-1\le x\le 1}|\tilde{p}(x)|$.]

Properties of the Chebyshev Polynomials

As we've seen, the Chebyshev polynomial $T_n(x)$ is the (unique, real) polynomial of degree n (having leading coefficient 1 if $n = 0$, and 2^{n-1} if $n \ge 1$) such that $T_n(\cos\theta) = \cos n\theta$ for all θ. The Chebyshev polynomials have dozens of interesting properties and satisfy all sorts of curious equations. We'll catalogue just a few.

C1. $T_n(x) = 2x\,T_{n-1}(x) - T_{n-2}(x)$ for $n \ge 2$.

Proof. It follows from the trig identity $\cos n\theta = 2\cos\theta\,\cos(n-1)\theta - \cos(n-2)\theta$ that $T_n(\cos\theta) = 2\cos\theta\,T_{n-1}(\cos\theta) - T_{n-2}(\cos\theta)$ for all θ; that is, the equation $T_n(x) = 2x\,T_{n-1}(x) - T_{n-2}(x)$ holds for all $-1 \le x \le 1$. But because both sides are polynomials, equality must hold for all x. □

The next two properties are proved in essentially the same way:

C2. $T_m(x) + T_n(x) = \frac{1}{2}\big[\,T_{m+n}(x) + T_{m-n}(x)\,\big]$ for $m > n$.

C3. $T_m(T_n(x)) = T_{mn}(x)$.

C4. $T_n(x) = \frac{1}{2}\Big[\,\big(x + \sqrt{x^2-1}\,\big)^n + \big(x - \sqrt{x^2-1}\,\big)^n\,\Big]$.

Proof. First notice that the expression on the right-hand side is actually a polynomial because, on combining the binomial expansions of $(x+\sqrt{x^2-1}\,)^n$ and $(x-\sqrt{x^2-1}\,)^n$, the odd powers of $\sqrt{x^2-1}$ cancel. Next, for $x = \cos\theta$,

$$\begin{aligned} T_n(x) = T_n(\cos\theta) &= \cos n\theta = \frac{1}{2}(e^{in\theta} + e^{-in\theta}) \\ &= \frac{1}{2}\Big[\,(\cos\theta + i\sin\theta)^n + (\cos\theta - i\sin\theta)^n\,\Big] \\ &= \frac{1}{2}\Big[\,\big(x + i\sqrt{1-x^2}\,\big)^n + \big(x - i\sqrt{1-x^2}\,\big)^n\,\Big] \\ &= \frac{1}{2}\Big[\,\big(x + \sqrt{x^2-1}\,\big)^n + \big(x - \sqrt{x^2-1}\,\big)^n\,\Big]. \end{aligned}$$

We've shown that these two polynomials agree for $|x| \le 1$, hence they must agree for all x (real or complex, for that matter). □

For real x with $|x| \ge 1$, the expression $\frac{1}{2}\big[\,(x + \sqrt{x^2-1}\,)^n + (x - \sqrt{x^2-1}\,)^n\,\big]$ equals $\cosh(n\cosh^{-1} x)$. In other words, we have

C5. $T_n(\cosh x) = \cosh nx$ for all real x.

The next property also follows from property **C4**.

C6. $T_n(x) \le \big(\,|x| + \sqrt{x^2-1}\,\big)^n$ for $|x| \ge 1$.

An approach similar to the proof of property **C4** allows us to write x^n in terms of the Chebyshev polynomials $T_0, T_1, \ldots, T_n$.

C7. For n odd, $2^n x^n = \displaystyle\sum_{k=0}^{[n/2]} \binom{n}{k} 2\,T_{n-2k}(x)$; for n even, $2\,T_0$ should be replaced by T_0.

Proof. For $-1 \le x \le 1$,

$$\begin{aligned} 2^n x^n &= 2^n(\cos\theta)^n = (e^{i\theta} + e^{-i\theta})^n \\ &= e^{in\theta} + \binom{n}{1} e^{i(n-2)\theta} + \binom{n}{2} e^{i(n-4)\theta} + \cdots \\ &\qquad \cdots + \binom{n}{n-2} e^{-i(n-4)\theta} + \binom{n}{n-1} e^{-i(n-2)\theta} + e^{-in\theta} \\ &= 2\cos n\theta + \binom{n}{1} 2\cos(n-2)\theta + \binom{n}{2} 2\cos(n-4)\theta + \cdots \\ &= 2\,T_n(x) + \binom{n}{1} 2\,T_{n-2}(x) + \binom{n}{2} 2\,T_{n-4}(x) + \cdots, \end{aligned}$$

where, if n is even, the last term in this last sum is $\binom{n}{[n/2]} T_0$ (because the central term in the binomial expansion, namely $\binom{n}{[n/2]} = \binom{n}{[n/2]} T_0$, isn't doubled in this case). □

C8. The zeros of T_n are $x_k^{(n)} = \cos((2k-1)\pi/2n)$, $k = 1, \ldots, n$. They're real, simple, and lie in the open interval $(-1, 1)$.

Proof. Just check! But notice, please, that the zeros are listed here in *decreasing* order (because cosine decreases). $\square$

C9. Between two consecutive zeros of T_n, there is precisely one root of T_{n-1}.

Proof. It's not hard to check that

$$\frac{2k-1}{2n} \;<\; \frac{2k-1}{2\,(n-1)} \;<\; \frac{2k+1}{2n},$$

for $k = 1, \ldots, n-1$, which means that $x_k^{(n)} > x_k^{(n-1)} > x_{k+1}^{(n)}$. $\square$

C10. T_n and T_{n-1} have no common zeros.

Proof. Although this is immediate from property **C9**, there's another way to see it: $T_n(x_0) = 0 = T_{n-1}(x_0)$ implies that $T_{n-2}(x_0) = 0$ by property **C1**. Repeating this observation, we would have $T_k(x_0) = 0$ for every $k < n$, including $k = 0$. No good! $T_0(x) = 1$ has no zeros. $\square$

C11. The set $\{x_k^{(n)} : 1 \le k \le n,\ n = 1, 2, \ldots\}$ is dense in $[-1, 1]$.

Proof. Because $\cos x$ is (strictly) monotone on $[\,0, \pi\,]$, it's enough to know that the set $\{(2k-1)\pi/2n\}_{k,n}$ is dense in $[\,0, \pi\,]$, and for this it's enough to know that $\{(2k-1)/2n\}_{k,n}$ is dense in $[\,0, 1\,]$. (Why?) But

$$\frac{2k-1}{2n} = \frac{k}{n} - \frac{1}{2n} \approx \frac{k}{n}$$

for n large; that is, the set $\{(2k-1)/2n\}_{k,n}$ is dense among the rationals in $[\,0, 1\,]$. $\square$

It's interesting to note here that the *distribution* of the roots $\{x_k^{(n)}\}_{k,n}$ can be estimated (see Natanson [41, Vol. I, pp. 48–51]). For large n, the number of roots of T_n that lie in an interval $[\,x, x + \Delta x\,] \subset [-1, 1\,]$ is approximately

$$\frac{n\Delta x}{\pi\sqrt{1-x^2}}.$$

In particular, for n large, the roots of T_n are "thickest" near the endpoints ± 1.

In probabilistic terms, this means that if we assign equal probability to each of the roots $x_0^{(n)}, \ldots, x_n^{(n)}$ (that is, if we think of each root as the position of a point with mass $1/(n+1)$), then the *density* of this probability distribution (or the density of the system of point masses) at a point x is approximately $1/\pi\sqrt{1-x^2}$ for large n. In still other words, this tells us that the probability that a root of T_n lies in the interval $[\,a, b\,]$ is approximately

$$\frac{1}{\pi}\int_a^b \frac{1}{\sqrt{1-x^2}}\,dx.$$

C12. The Chebyshev polynomials are mutually *orthogonal* relative to the weight $w(x) = (1-x^2)^{-1/2}$ on $[-1,1]$.

Proof. For $m \neq n$ the substitution $x = \cos\theta$ yields

$$\int_{-1}^{1} T_n(x)\, T_m(x)\, \frac{dx}{\sqrt{1-x^2}} \;=\; \int_0^\pi \cos m\theta\, \cos n\theta\, d\theta = 0,$$

while for $m = n$ we get

$$\int_{-1}^{1} T_n^2(x)\, \frac{dx}{\sqrt{1-x^2}} \;=\; \int_0^\pi \cos^2 n\theta\, d\theta = \begin{cases} \pi & \text{if } n = 0 \\ \pi/2 & \text{if } n > 0. \end{cases}$$ □

C13. $|T_n'(x)| \leq n^2$ for $-1 \leq x \leq 1$, and $|T_n'(\pm 1)| = n^2$.

Proof. For $-1 < x < 1$ we have

$$\frac{d}{dx} T_n(x) \;=\; \frac{\frac{d}{d\theta}\, T_n(\cos\theta)}{\frac{d}{d\theta}\cos\theta} \;=\; \frac{n\sin n\theta}{\sin\theta}.$$

Thus, $|T_n'(x)| \leq n^2$ because $|\sin n\theta| \leq n|\sin\theta|$ (which can be easily checked by induction, for example). At $x = \pm 1$, we interpret this derivative formula as a limit (as $\theta \to 0$ and $\theta \to \pi$) and find that $|T_n'(\pm 1)| = n^2$. □

As we'll see later, each $p \in \mathcal{P}_n$ satisfies $|p'(x)| \leq \|p\| n^2 = \|p\|\, T_n'(1)$ for $-1 \leq x \leq 1$, and this is, of course, best possible. As it happens, $T_n(x)$ has the largest possible rate of growth outside of $[-1,1]$ among all polynomials of degree n. Specifically:

Theorem 4.12. *Let $p \in \mathcal{P}_n$ and let $\|p\| = \max_{-1\leq x\leq 1} |p(x)|$. Then, for any x_0 with $|x_0| \geq 1$ and any $k = 0, 1, \ldots, n$ we have*

$$|p^{(k)}(x_0)| \leq \|p\|\, |\, T_n^{(k)}(x_0)\, |,$$

where $p^{(k)}$ is the k-th derivative of p.

We'll prove only the case $k = 0$. In other words, we'll check that $|p(x_0)| \leq \|p\|\, |T_n(x_0)|$. The more general case is given in Rivlin [45, Theorem 1.10, p. 31] and uses a similar proof.

Proof. Because all the zeros of T_n lie in $(-1,1)$, we know that $T_n(x_0) \neq 0$. Thus, we may consider the polynomial

$$q(x) \;=\; \frac{p(x_0)}{T_n(x_0)}\, T_n(x) \;-\; p(x) \in \mathcal{P}_n.$$

If the claim were *false*, then

$$\|p\| < \left| \frac{p(x_0)}{T_n(x_0)} \right|.$$

Now at each of the points $y_k = \cos(k\pi/n)$, $k = 0, 1, \ldots, n$, we have $T_n(y_k) = (-1)^k$ and, hence,

$$q(y_k) \;=\; (-1)^k \frac{p(x_0)}{T_n(x_0)} \;-\; p(y_k).$$

Because $|p(y_k)| \le \|p\|$, it follows that q alternates in sign at these $n+1$ points. In particular, q must have at least n zeros in $(-1,1)$. But $q(x_0) = 0$, by design, and $|x_0| \ge 1$. That is, we've found $n+1$ zeros for a polynomial of degree n. So, $q \equiv 0$; that is,

$$p(x) \;=\; \frac{p(x_0)}{T_n(x_0)}\, T_n(x).$$

But then,

$$|p(1)| = \left| \frac{p(x_0)}{T_n(x_0)} \right| > \|p\|,$$

because $T_n(1) = T_n(\cos 0) = 1$, which is a contradiction. $\square$

Corollary 4.13. *Let $p \in \mathcal{P}_n$ and let $\|p\| = \max_{-1\le x\le 1} |p(x)|$. Then, for any x_0 with $|x_0| \ge 1$, we have*

$$|p(x_0)| \le \|p\| \left(|x_0| + \sqrt{x_0^2 - 1}\, \right)^n.$$

Rivlin's proof of Theorem 4.12 in the general case uses the following observation:

C14. For $x \ge 1$ and $k = 0, 1, \ldots, n$, we have $T_n^{(k)}(x) > 0$.

Proof. **Exercise**. [Hint: It follows from Rolle's theorem that $T_n^{(k)}$ is never zero for $x \ge 1$. (Why?) Now just compute $T_n^{(k)}(1)$.]

Chebyshev Polynomials in Practice

The following discussion is cribbed from the book *Chebyshev Polynomials in Numerical Analysis* by L. Fox and I. B. Parker [18].

Example 4.14. As we've seen, the Chebyshev polynomals can be generated by a recurrence relation. By reversing the procedure, we could solve for x^n in terms of $T_0, T_1, \ldots, T_n$. Here are the first few terms in each of these relations:

$$\begin{array}{rcl@{\qquad\qquad}rcl}
T_0(x) &=& 1 & 1 &=& T_0(x)\\
T_1(x) &=& x & x &=& T_1(x)\\
T_2(x) &=& 2x^2 - 1 & x^2 &=& (T_0(x) + T_2(x))/2\\
T_3(x) &=& 4x^3 - 3x & x^3 &=& (3\,T_1(x) + T_3(x))/4\\
T_4(x) &=& 8x^4 - 8x^2 + 1 & x^4 &=& (3\,T_0(x) + 4\,T_2(x) + T_4(x))/8\\
T_5(x) &=& 16x^5 - 20x^3 + 5x & x^5 &=& (10\,T_1(x) + 5\,T_3(x) + T_5(x))/16
\end{array}$$

Note the separation of even and odd terms in each case. Writing ordinary garden variety polynomials in their equivalent Chebyshev form has some distinct advantages for numerical computations. Here's why:

$$1 - x + x^2 - x^3 + x^4 = \frac{15}{6}\, T_0(x) - \frac{7}{4}\, T_1(x) + T_2(x) - \frac{1}{4}\, T_3(x) + \frac{1}{8}\, T_4(x)$$

(after some simplification). Now we see at once that we can get a cubic approximation to $1 - x + x^2 - x^3 + x^4$ on $[-1, 1\,]$ with error at most $1/8$ by simply dropping the T_4 term on the right-hand side (because $|T_4(x)| \le 1$), whereas simply using $1 - x + x^2 - x^3$ as our cubic approximation could cause an error as big as 1. Pretty slick! This gimmick of truncating the equivalent Chebyshev form is called *economization*.

Example 4.15. We next note that a polynomial with small norm on $[-1,1]$ may have annoyingly large coefficients:

$$\begin{aligned}(1-x^2)^{10} &= 1-10x^2+45x^4-120x^6+210x^8-252x^{10}\\ &\qquad +210x^{12}-120x^{14}+45x^{16}-10x^{18}+x^{20}\end{aligned}$$

but in Chebyshev form (look out!):

$$\begin{aligned}(1-x^2)^{10} &= \frac{1}{524{,}288}\Big\{92{,}378\,T_0(x)-167{,}960\,T_2(x)+125{,}970\,T_4(x)\\ &\qquad -77{,}520\,T_6(x)+38{,}760\,T_8(x)-15{,}504\,T_{10}(x)\\ &\qquad +4{,}845\,T_{12}(x)-1{,}140\,T_{14}(x)+190\,T_{16}(x)\\ &\qquad -20\,T_{18(x)}+T_{20}(x)\Big\}\end{aligned}$$

The largest coefficient is now only about 0.3, and the omission of the last three terms produces a maximum error of about 0.0004. Not bad.

Example 4.16. As a last example, consider the Taylor polynomial $e^x = \sum_{k=0}^n x^k/k! + x^{n+1}e^{\xi}/(n+1)!$ (with remainder), where $-1 \le x,\, \xi \le 1$. Taking $n=6$, the truncated series has error no greater than $e/7! \approx 0.0005$. But if we "economize" the first six terms, then:

$$\begin{aligned}\sum_{k=0}^{6} x^k/k &= 1.26606\,T_0(x)+1.13021\,T_1(x)+0.27148\,T_2(x)+0.04427\,T_3(x)\\ &\qquad +0.00547\,T_4(x)+0.00052\,T_5(x)+0.00004\,T_6(x).\end{aligned}$$

The initial approximation already has an error of about 0.0005, so we can certainly drop the T_6 term without any additional error. Even dropping the T_5 term causes an error of no more than 0.001 (or thereabouts). The resulting approximation has a far smaller error than the corresponding truncated Taylor series of degree 4 which is $e/5! \approx 0.023$.

The approach used in our last example has the decided disadvantage that we must first decide where to truncate the Taylor series—which might converge very slowly. A better approach would be to write e^x as a series involving Chebyshev polynomials directly. That is, if possible, we first want to write $e^x = \sum_{k=0}^{\infty} a_k T_k(x)$. If the a_k are absolutely summable, it will be very easy to estimate any truncation error. We'll get some idea on how to go about this when we talk about "least-squares" approximation. As it happens, such a series is easy to find (it's rather like a Fourier series), and its partial sums are remarkably good uniform approximations.

In fact, for continuous f and any $n < 400$, we can never hope for more than one extra decimal place of accuracy by using the best polynomial of degree n in place the the n-th partial sum of this Chebyshev series!

Uniform Approximation by Trig Polynomials

We end this chapter by summarizing (without proofs) the analogues of Theorems 4.3–4.6 for uniform approximation by trig polynomials. Throughout, $f \in C^{2\pi}$ and $\mathcal{T}_n$ denotes the collection of trig polynomials of degree at most n. We will also write

$$E_n^T(f) = \min_{T \in \mathcal{T}_n} \|f - T\|$$

(to distinguish this distance from $E_n(f)$).

T1. f has a best approximation $T^* \in \mathcal{T}_n$.

T2. $f - T^*$ has an alternating set containing $2n+2$ (or more) points in $[0, 2\pi)$. (Note here that $2n + 2 = 1 + \dim \mathcal{T}_n$.)

T3. T^* is unique.

T4. If $T \in \mathcal{T}_n$ is such that $f - T$ has an alternating set containing $2n + 2$ or more points in $[0, 2\pi)$, then $T = T^*$.

The proofs of **T1**–**T4** are very similar to the corresponding results for algebraic polynomials. As you might imagine, **T2** is where all the fighting takes place, and there are a few technical difficulties to cope with. Nevertheless, we'll swallow these facts whole and apply them with a clear conscience to a few examples.

Example 4.17. For $m > n$, the best approximation to $f(x) = A\cos mx + B\sin mx$ out of $\mathcal{T}_n$ is 0!

Proof. We may write $f(x) = R\cos m(x - x_0)$ for some R and x_0. (How?) Now we need only display a sufficiently large alternating set for f (in some interval of length 2π).

Setting $x_k = x_0 + k\pi/m$, $k = 1, 2, \ldots, 2m$, we get $f(x_k) = R\cos k\pi = R(-1)^k$ and $x_k \in (x_0, x_0 + 2\pi]$. Because $m > n$, it follows that $2m \geq 2n + 2$. □

Example 4.18. The best approximation to

$$f(x) = a_0 + \sum_{k=1}^{n+1} \big(a_k \cos kx + b_k \sin kx\big)$$

out of $\mathcal{T}_n$ is

$$T(x) = a_0 + \sum_{k=1}^{n} \big(a_k \cos kx + b_k \sin kx\big),$$

and $E_n^T(f) = \|f - T\| = \sqrt{a_{n+1}^2 + b_{n+1}^2}$.

Proof. By our last example, the best approximation to $f - T$ out of $\mathcal{T}_n$ is 0, hence T must be the best approximation to f. (Why?) The last assertion is easy to check: Because we can always write $A\cos mx + B\sin mx = \sqrt{A^2 + B^2} \cdot \cos m(x - x_0)$, for some x_0, it follows that $\|f - T\| = \sqrt{a_{n+1}^2 + b_{n+1}^2}$. □

Finally, let's make a simple connection between the two types of polynomial approximation:

Theorem 4.19. *Let $f \in C[-1, 1]$ and define $\varphi \in C^{2\pi}$ by $\varphi(\theta) = f(\cos\theta)$. Then,*

$$E_n(f) = \min_{p \in \mathcal{P}_n} \|f - p\| = \min_{T \in \mathcal{T}_n} \|\varphi - T\| \equiv E_n^T(\varphi).$$

Proof. Suppose that $p^*(x) = \sum_{k=0}^n a_k x^k$ is the best approximation to f out of $\mathcal{P}_n$. Then, $\widehat{T}(\theta) = p^*(\cos\theta)$ is in $\mathcal{T}_n$ and, clearly,

$$\max_{-1\le x\le 1} |f(x) - p^*(x)| = \max_{0\le\theta\le 2\pi} |f(\cos\theta) - p^*(\cos\theta)|.$$

Thus, $E_n(f) = \|f - p^*\| = \|\varphi - \widehat{T}\,\| \ge \min_{T\in\mathcal{T}_n} \|\varphi - T\| = E_n^T(\varphi)$.

On the other hand, because φ is *even*, we know that T^*, its best approximation out of $\mathcal{T}_n$, is also even. Thus, $T^*(\theta) = q(\cos\theta)$ for some algebraic polynomial $q \in \mathcal{P}_n$. Consequently, $E_n^T(\varphi) = \|\varphi - T^*\| = \|f - q\| \ge \min_{p\in\mathcal{P}_n} \|f - p\| = E_n(f)$. $\square$

Remarks 4.20.

1. Once we know that $\min_{p\in\mathcal{P}_n} \|f - p\| = \min_{T\in\mathcal{T}_n} \|\varphi - T\|$, it follows that we must also have $T^*(\theta) = p^*(\cos\theta)$.

2. Each *even* $\varphi \in C^{2\pi}$ corresponds to an $f \in C[-1,1\,]$ by setting $f(x) = \varphi(\arccos x)$. The conclusions of Theorem 4.19 and Remark 1 hold in this case, too.

3. Whenever we speak of even trig polynomials, the Chebyshev polynomials are lurking somewhere in the background. Indeed, let $T(\theta)$ be an even trig polynomial, write $x = \cos\theta$, as usual, and consider the following cryptic equation:

$$T(\theta) = \sum_{k=0}^n a_k \cos k\theta = \sum_{k=0}^n a_k T_k(\cos\theta) = p(\cos\theta),$$

   where $p(x) = \sum_{k=0}^n a_k T_k(x) \in \mathcal{P}_n$.

Problems

▷ 1. Prove Corollary 4.2.

▷ 2. Show that the polynomial of degree n having leading coefficient 1 and deviating least from 0 on the interval $[\,a, b\,]$ is given by

$$\frac{(b-a)^n}{2^{2n-1}}\, T_n\!\left(\frac{2x-b-a}{b-a}\right).$$

Is this solution unique? Explain.

▷ 3. Establish the following properties of $T_n(x)$.

(i) $|T_n(x)| > 1$ whenever $|x| > 1$.

(ii) $T_m(x) + T_n(x) = \frac{1}{2}\big[T_{m+n}(x) + T_{m-n}(x)\big]$ for $m > n$.

(iii) $T_m(T_n(x)) = T_{mn}(x)$.

(iv) Show that T_n is a solution to $(1-x^2)y'' - xy' + n^2 y = 0$.

(v) $\operatorname{Re}\left(\sum_{n=0}^{\infty} t^n e^{in\theta}\right) = \sum_{n=0}^{\infty} t^n \cos n\theta = \dfrac{1 - t\cos\theta}{1 - 2t\cos\theta + t^2}$ for $-1 < t < 1$; that is, $\sum_{n=0}^{\infty} t^n T_n(x) = \dfrac{1 - tx}{1 - 2tx + t^2}$ (this is a *generating function* for T_n; it's closely related to the Poisson kernel).

4. Find analogues of properties **C1–C13** and properties (i)–(v) in Problem 3 (if possible) for $U_n(x)$, the Chebyshev polynomials of the second kind.

▷ 5. Show that every $p \in \mathcal{P}_n$ has a unique representation as $p = a_0 + a_1 T_1 + \cdots + a_n T_n$. Find this representation in the case $p(x) = x^n$. [Hint: Using the recurrence formula $2xT_{n-1}(x) = T_n(x) + T_{n-2}(x)$, find representations for $2x^2$, $4x^3$, $8x^4$, etc.]

6. Let $f : X \to Y$ be a continuous map from a metric space X *onto* a metric space Y. If D is dense in X, show that $f(D)$ is dense in Y. Use this result to prove that the zeros of the Chebyshev polynomials are dense in $[-1, 1\,]$.

7. Show that $A\cos mx + B\sin mx = R\cos m(x - x_0) = R\sin m(x - x_1)$ for appropriately chosen values R, x_0, and x_1.

8. If p is a polynomial on $[\,a, b\,]$ of degree n having leading coefficient $a_n > 0$, then $\|p\| \ge a_n(b-a)^n/2^{2n-1}$. If $b - a \ge 4$, then no polynomial of degree exactly n with *integer* coefficients can satisfy $\|p\| < 2$. (Compare this with Problem 11 from Chapter 2.)

9. Given $p \in \mathcal{P}_n$, show that $|p(x)| \le \|p\|\,|T_n(x)|$ for $|x| > 1$.

10. If $p \in \mathcal{P}_n$ with $\|p\| = 1$ on $[-1, 1\,]$, and if $|p(x_i)| = 1$ at $n+1$ distinct point $x_0, \ldots, x_n$ in $[-1, 1\,]$, show that either $p = \pm 1$, or else $p = \pm T_n$. [Hint: One approach is to compare the polynomials $1 - p^2$ and $(1-x^2)(p\,')^2$.]

11. Compute $T_n^{(k)}(1)$ for $k = 0, 1, \ldots, n$, where $T_n^{(k)}$ is the k-th derivative of T_n. For $x \ge 1$ and $k = 0, 1, \ldots, n$, show that $T_n^{(k)}(x) > 0$.

Chapter 5

A Brief Introduction to Interpolation

Lagrange Interpolation

Our goal in this chapter is to prove the following result (as well as discuss its ramifications). In fact, this result is so fundamental that we will present three proofs!

Theorem 5.1. *Let $x_0, x_1, \ldots, x_n$ be distinct points and let $y_0, y_1, \ldots, y_n$ be arbitrary points in $\mathbb{R}$. Then there exists a unique polynomial $p \in \mathcal{P}_n$ satisfying $p(x_i) = y_i$, $i = 0, 1, \ldots, n$.*

First notice that uniqueness is obvious. Indeed, if two polynomials p, $q \in \mathcal{P}_n$ agree at $n+1$ points, then $p \equiv q$. (Why?) The real work comes in proving existence.

First Proof. (Vandermonde's determinant.) We seek $c_0, c_1, \ldots, c_n$ so that $p(x) = \sum_{k=0}^n c_k x^k$ satisfies

$$p(x_i) = \sum_{k=0}^n c_k x_i^k = y_i, \quad i = 0, 1, \ldots, n.$$

That is, we need to solve a system of $n+1$ *linear* equations for the c_i. In matrix form:

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

This equation always has a unique solution because the coefficient matrix has determinant

$$D = \prod_{0 \le j < i \le n} (x_i - x_j) \neq 0.$$

D is called *Vandermonde's determinant* (note that $D > 0$ if $x_0 < x_1 < \cdots < x_n$). Because this fact is of independent interest, we'll sketch a short proof below. $\square$

Lemma 5.2. $D = \prod_{0 \le j < i \le n} (x_i - x_j)$.

Proof. Consider

$$V(x_0, x_1, \ldots, x_{n-1}, x) = \begin{vmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \cdots & x_{n-1}^n \\ 1 & x & x^2 & \cdots & x^n \end{vmatrix}.$$

$V(x_0, x_1, \ldots, x_{n-1}, x)$ is a *polynomial* of degree n in x, and it's 0 whenever $x = x_i$, $i = 0, 1, \ldots, n-1$. Thus, $V(x_0, \ldots, x) = c \prod_{i=0}^{n-1}(x - x_i)$, by comparing roots and degree. However, it's easy to see that the coefficient of x^n in $V(x_0, \ldots, x)$ is $V(x_0, \ldots, x_{n-1})$. Thus, $V(x_0, \ldots, x) = V(x_0, \ldots, x_{n-1}) \prod_{i=0}^{n-1}(x - x_i)$. The result now follows by induction and the obvious case $V(x_0, x_1) = x_1 - x_0$. □

Second Proof. (Lagrange interpolation.) We could define p immediately if we had polynomials $\ell_i(x) \in \mathcal{P}_n$, $i = 0, \ldots, n$, such that $\ell_i(x_j) = \delta_{i,j}$ (where $\delta_{i,j}$ is *Kronecker's delta*; that is, $\delta_{i,j} = 0$ for $i \neq j$, and $\delta_{i,j} = 1$ for $i = j$). Indeed, $p(x) = \sum_{i=0}^{n} y_i \, \ell_i(x)$ would then work as our interpolating polynomial. In short, notice that the polynomials $\{\ell_0, \ell_1, \ldots, \ell_n\}$ would form a (particularly convenient) *basis* for $\mathcal{P}_n$.

We'll give two formulas for $\ell_i(x)$:

(a). Clearly, $\displaystyle \ell_i(x) = \prod_{\substack{j=1,\ldots,n \\ j \neq i}} \frac{x - x_j}{x_i - x_j}$ works.

(b). Start with $W(x) = (x - x_0)(x - x_1) \cdots (x - x_n)$, and notice that the polynomial we need satisfies

$$\ell_i(x) = a_i \cdot \frac{W(x)}{x - x_i}$$

for some $a_i \in \mathbb{R}$. (Why?) But then $1 = \ell_i(x_i) = a_i W'(x_i)$ (again, why?); that is, we must have

$$\ell_i(x) = \frac{W(x)}{(x - x_i)\, W'(x_i)}.$$

[As an aside, note that $W'(x_i)$ is easy to compute: Indeed, $W'(x_i) = \prod_{j \neq i}(x_j - x_i)$.]

Please note that $\ell_i(x)$ is a multiple of the polynomial $\prod_{j \neq i}(x - x_j)$, for $i = 0, \ldots, n$, and that $p(x)$ is then a suitable linear combination of the ℓ_i. □

Third Proof. (Newton's formula.) We seek $p(x)$ of the form

$$p(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \cdots + a_n(x - x_0) \cdots (x - x_{n-1}).$$

(Please note that x_n does not appear on the right-hand side.) This form makes it almost effortless to solve for the a_i by plugging-in the x_i, $i = 0, \ldots, n-1$.

$$y_0 = p(x_0) = a_0$$
$$y_1 = p(x_1) = a_0 + a_1(x_1 - x_0) \implies a_1 = \frac{y_1 - a_0}{x_1 - x_0}.$$

Continuing, we find

$$a_2 = \frac{y_2 - a_0 - a_1(x_2 - x_0)}{(x_2 - x_0)(x_2 - x_1)}$$

$$a_3 = \frac{y_3 - a_0 - a_1(x_3 - x_0) - a_2(x_3 - x_0)(x_3 - x_1)}{(x_3 - x_0)(x_3 - x_1)(x_3 - x_2)}$$

and so on. (Natanson [41, Vol. III] gives another formula for the a_i.) □

Example 5.3. (Cheney [12]) As a quick means of comparing these three solutions, let's find the interpolating polynomial (quadratic) passing through $(1, 2)$, $(2, -1)$, and $(3, 1)$. You're invited to check the following:

(Vandermonde): $p(x) = 10 - \frac{21}{2}\, x + \frac{5}{2}\, x^2$.

(Lagrange): $p(x) = (x - 2)(x - 3) + (x - 1)(x - 3) + \frac{1}{2}(x - 1)(x - 2)$.

(Newton): $p(x) = 2 - 3(x - 1) + \frac{5}{2}(x - 1)(x - 2)$.

As you might have already surmised, Lagrange's method is the easiest to apply by hand, although Newton's formula has much to recommend it too (it's especially well-suited to situations where we introduce additional nodes). We next set up the necessary notation to discuss the finer points of Lagrange's method.

Given $n + 1$ distinct points $a \le x_0 < x_1 < \cdots < x_n \le b$ (sometimes called *nodes*), we first form the polynomials

$$W(x) \;=\; \prod_{i=0}^{n} (x - x_i)$$

and

$$\ell_i(x) \;=\; \prod_{j \ne i} \frac{x - x_j}{x_i - x_j} \;=\; \frac{W(x)}{(x - x_i)\, W'(x_i)}.$$

The *Lagrange interpolation formula* is then

$$L_n(f)(x) \;=\; \sum_{i=0}^{n} f(x_i)\, \ell_i(x).$$

That is, $L_n(f)$ is the unique polynomial in $\mathcal{P}_n$ that agrees with f at the x_i. In particular, notice that we must have $L_n(p) = p$ whenever $p \in \mathcal{P}_n$. In fact, L_n is a *linear projection* from $C[\,a, b\,]$ onto $\mathcal{P}_n$. (Why is $L_n(f)$ linear in f?)

Typically we're given (or construct) *an array of nodes*:

$$X \qquad \left\{ \begin{array}{llll} x_0^{(0)} & & & \\ x_0^{(1)} & x_1^{(1)} & & \\ x_0^{(2)} & x_1^{(2)} & x_2^{(2)} & \\ \vdots & & & \ddots \end{array} \right.$$

and form the corresponding sequence of projections

$$L_n(f)(x) \;=\; \sum_{i=0}^{n} f(x_i^{(n)})\, \ell_i^{(n)}(x).$$

An easy (but admittedly pointless) observation is that for a *given* $f \in C[\,a,b\,]$ we can always find an array X so that $L_n(f) = p_n^*$, the polynomial of best approximation to f out of $\mathcal{P}_n$ (because $f - p_n^*$ has $n+1$ zeros, we may use these for the x_i). Thus, $\|L_n(f) - f\| = E_n(f) \to 0$ in this case. However, the problem of convergence changes character dramatically if we *first choose* X and then consider $L_n(f)$. In general, there's no reason to believe that $L_n(f)$ converges to f. In fact, quite the opposite is true:

Theorem 5.4. (Faber, 1914) *Given any array of nodes X in $[\,a,b\,]$, there is some $f \in C[\,a,b\,]$ for which $\|L_n(f) - f\|$ is unbounded.*

The problem here has little to do with interpolation and everything to do with projections:

Theorem 5.5. (Kharshiladze, Lozinski, 1941) *For each n, let L_n be a continuous, linear projection from $C[\,a,b\,]$ onto $\mathcal{P}_n$. Then there is some $f \in C[\,a,b\,]$ for which $\|L_n(f) - f\|$ is unbounded.*

Evidently, the operators L_n aren't positive (monotone), for otherwise the Bohman-Korovkin theorem (Theorem 2.9) and the fact that L_n is a projection onto $\mathcal{P}_n$ would imply that $L_n(f)$ converges uniformly to f for every $f \in C[\,a,b\,]$.

The proofs of these theorems are long and difficult—we'll save them for another day. (Some of you may recognize the Principle of Uniform Boundedness at work here.) The real point here is that we can't have everything: A positive result about convergence of interpolation will require that we impose some extra conditions on the functions f we want to approximate. As a first step in this direction, we next prove that if f has sufficiently many derivatives, then the error $\|L_n(f) - f\|$ can at least be estimated.

Theorem 5.6. *Suppose that f has $n+1$ continuous derivatives on $[\,a,b\,]$. Let $a \le x_0 < x_1 < \cdots < x_n \le b$, let $W(x) = \prod_{i=0}^n (x - x_i)$, and let $L_n(f) \in \mathcal{P}_n$ be the polynomial that interpolates to f at the x_i. Then*

$$|f(x) - L_n(f)(x)| \;\le\; \frac{1}{(n+1)!}\, \|f^{\,(n+1)}\|\, |W(x)|. \tag{5.1}$$

for every x in $[\,a,b\,]$.

Proof. For convenience, let's write $p = L_n(f)$. We'll prove the Theorem by showing that, given x in $[\,a,b\,]$, there exists a ξ in (a,b) with

$$f(x) - p(x) \;=\; \frac{1}{(n+1)!}\, f^{\,(n+1)}(\xi)\, W(x). \tag{5.2}$$

If x is one of the x_i, then both sides of this formula are 0 and we're done. Otherwise, $W(x) \neq 0$ and we may set $\lambda = [f(x) - p(x)]/W(x)$. Now consider

$$\varphi(t) = f(t) - p(t) - \lambda W(t).$$

Clearly, $\varphi(x_i) = 0$ for each $i = 0, 1, \ldots, n$ and, by our choice of λ, we also have $\varphi(x) = 0$. Here comes Rolle's theorem! Because φ has $n+2$ distinct zeros in $[\,a,b\,]$, we must have $\varphi^{(n+1)}(\xi) = 0$ for some ξ in (a,b). (Why?) Hence,

$$\begin{aligned} 0 = \varphi^{(n+1)}(\xi) \;&=\; f^{\,(n+1)}(\xi) - p^{(n+1)}(\xi) - \lambda W^{(n+1)}(\xi) \\ &=\; f^{\,(n+1)}(\xi) - \left(\frac{f(x) - p(x)}{W(x)}\right) \cdot (n+1)! \end{aligned}$$

because p has degree at most n and W is monic and degree $n+1$. $\square$

Remarks 5.7.

1. Equation (5.2) is called the *Lagrange formula with remainder*. [Compare this result to Taylor's formula with remainder.]

2. The term $f^{(n+1)}(\xi)$ is actually a continuous function of x. That is, $[f(x)-p(x)]/W(x)$ is continuous; its value at an x_i is $[f'(x_i)-p'(x_i)]/W'(x_i)$ (why?) and $W'(x_i) = \prod_{j\neq i}(x_i - x_j) \neq 0$.

3. On *any* interval $[\,a,b\,]$, using *any* nodes, the sequence of Lagrange interpolating polynomials for e^x converge uniformly to e^x. In this case,
$$\|e^x - L_n(e^x)\| \;\leq\; \frac{c}{(n+1)!}\,(b-a)^n \to 0 \quad (\text{as } n\to\infty)$$
where $c = \|e^x\|$ in $C[\,a,b\,]$. A similar result would hold true for any infinitely differentiable function satisfying, say, $\|f^{(n)}\| \leq M^n$ (any entire function, for example).

4. On $[-1,1\,]$, the norm of $\prod_{i=1}^n (x-x_i)$ is minimized by taking $x_i = \cos((2i-1)\pi/2n)$, the zeros of the n-th Chebyshev polynomial T_n. (Why?) As Rivlin points out [45], the zeros of the Chebyshev polynomials are a *nearly optimal* choice for the nodes if good uniform approximation is desired. We'll pursue this observation in greater detail a bit later.

5. In practice, through translation and scaling, interpolation is generally performed on very narrow intervals about 0 of the form $[-\delta,\delta\,]$ where $\delta\approx 2^{-5}$ (or even 2^{-10}). This scaling typically produces interpolating polynomials of smaller degree and, moreover, facilitates error estimation (in terms of the number of accurate bits in a computer calculation). In addition, it is often convenient to require that $f(0)$ be interpolated exactly; in this case, we might seek an interpolating polynomial of the form $p_n(x) = f(0) + x^m p_{n-m}(x)$, where m is the multiplicity of zero as a root of $f(x)-f(0)$, and where p_{n-m} is a polynomial of degree $n-m$ that interpolates to $(f(x)-f(0))/x^m$. For example, in the case of $f(x) = \cos x$, we might seek a polynomial of the form $p_n(x) = 1 + x^2 p_{n-2}(x)$.

The question of convergence of interpolation is actually very closely related to the analogous question for the convergence of Fourier series—and the answer here is nearly the same. We'll have much more to say about this analogy later. For now, let's first note that L_n is continuous (bounded); this will give us our first bit of insight into Faber's negative result.

Lemma 5.8. $\|L_n(f)\| \leq \|f\|\,\big\|\sum_{i=0}^n |\ell_i(x)|\,\big\|$ *for any* $f\in C[\,a,b\,]$.

Proof. **Exercise**.

The expressions $\lambda_n(x) = \sum_{i=0}^n |\ell_i(x)|$ are called the *Lebesgue functions* and their norms $\Lambda_n = \big\|\sum_{i=0}^n |\ell_i(x)|\,\big\|$ are called the *Lebesgue numbers* associated to this process. It's not hard to see that Λ_n is the smallest possible constant that will work in this inequality; in other words, $\|L_n\| = \Lambda_n$. Indeed, if
$$\left\|\sum_{i=0}^n |\ell_i(x)|\right\| = \sum_{i=0}^n |\ell_i(x_0)|,$$

then we can find an $f \in C[\,a,b\,]$ with $\|f\| = 1$ and $f(x_i) = \operatorname{sgn}(\ell_i(x_0))$ for all i. (How?) Then

$$\|L_n(f)\| \ge |L_n(f)(x_0)| = \left|\sum_{i=0}^{n} \operatorname{sgn}(\ell_i(x_0))\,\ell_i(x_0)\right| = \sum_{i=0}^{n} |\ell_i(x_0)| = \Lambda_n\|f\|.$$

As it happens, for any given array of nodes X, we always have $\Lambda_n(X) \ge c\log n$ for some absolute constant c (this is where the hard work comes in; see Rivlin [45] or Natanson [41] for further details), and, in particular, $\Lambda_n(X) \to \infty$ as $n \to \infty$. Given this (and the Principle of Uniform Boundedness; see Appendix G), Faber's Theorem (Theorem 5.4) now follows immediately.

A simple application of the triangle inequality will allow us to bring $E_n(f)$ back into the picture:

Lemma 5.9. (Lebesgue's Theorem) $\|f - L_n(f)\| \le (1 + \Lambda_n)\,E_n(f)$, *for any* $f \in C[\,a,b\,]$.

Proof. Let p^* be the best approximation to f out of $\mathcal{P}_n$. Then, because $L_n(p^*) = p^*$, we have

$$\begin{aligned}\|f - L_n(f)\| &\le \|f - p^*\| + \|L_n(f - p^*)\| \\ &\le (1 + \Lambda_n)\,\|f - p^*\| \;=\; (1 + \Lambda_n)\,E_n(f). \qquad \square\end{aligned}$$

Corollary 5.10. *For any* $f \in C[\,a,b\,]$ *we have*

$$\|f - L_n(f)\| \le (3/2)(1 + \Lambda_n)\,\omega_f(1/\sqrt{n}).$$

It follows from Corollary 5.10 that $L_n(f)$ converges uniformly to f provided that the array of nodes X can be chosen to satisfy $\Lambda_n(X)\omega_f(1/\sqrt{n}) \to 0$ as $n \to \infty$. Consider, however, the following somewhat disheartening fact (see Rivlin [45, Theorem 4.6]): For the array of equally spaced nodes in $[-1,1\,]$, which we will call E, there exist absolute constants k_1 and k_2 such that

$$k_1(3/2)^m \le \Lambda_{2m}(E) \le k_2 2^m e^{2m}$$

for all $m \ge 2$. That is, the Lebesgue numbers for this process grow exponentially fast. It will come as no surprise, then, that there are surprisingly simple functions for which interpolation at equally spaced points fails to converge:

Examples 5.11.

1. (S. Bernstein, 1918) On $[-1,1\,]$, the Lagrange interpolating polynomials to $f(x) = |x|$ based on the system of equally spaced nodes E, satisfy

$$\limsup_{n\to\infty} L_n(f)(x) = +\infty$$

for all $0 < |x| < 1$. In other words, the sequence $(L_n(f)(x))$ *diverges* for all x in $[-1,1\,]$ other than $x = 0, \pm 1$.

2. (C. Runge, 1901) On $[-5,5\,]$, the Lagrange interpolating polynomials to $f(x) = (1 + x^2)^{-1}$ based on the system of equally spaced nodes E satisfy

$$\limsup_{n\to\infty} L_n(f)(x) = +\infty$$

for all $|x| > 3.63$. In this case, the interpolating polynomials exhibit radical behavior near the endpoints of the interval. This phenomenon is similar to the so-called Gibbs phenomenon from Fourier analysis, where the approximating polynomials are sometimes known to "overshoot" their goal.

Nevertheless, as we'll see in the next section, we can find a (nearly optimal) array of nodes T for which the interpolation process will converge for a large class of functions and, moreover, has several practical advantages.

Chebyshev Interpolation

In this section, we focus our attention on the array of nodes in $[-1, 1]$ generated by the zeros of the Chebyshev polynomials, which we will refer to as T (and call simply *the Chebyshev nodes*); these are the points

$$x_k^{(n+1)} = \cos\theta_k^{(n+1)} = \cos\left(\frac{(2k+1)\pi}{2(n+1)}\right), \quad k = 0, \ldots, n, \quad n = 0, 1, 2, \ldots.$$

(Note, specifically, that $(x_k^{(n+1)})_{k=0}^n$ are the zeros of T_{n+1}. Also note that the zeros are listed in *decreasing* order in the formula above; for this reason, some authors write, instead, $x_k^{(n+1)} = -\cos\theta_k^{(n+1)}$.) As usual, when n is clear from context, we will omit the superscript $(n+1)$.

We know that the Lebesgue numbers for any interpolation process grow at least as fast as $\log n$. In the case of interpolation at the Chebyshev nodes T, this is also an upper bound, lending further evidence to our repeated claims that these nodes are nearly optimal (at least for uniform approximation). The following result is due (essentially) to Bernstein [6]; for a proof of the version stated here, see Rivlin [45, Theorem 4.5].

Theorem 5.12. *For the Chebyshev system of nodes T we have $\Lambda_n(T) < (2/\pi)\log n + 4$.*

If we apply Lagrange interpolation to the Chebyshev system of nodes, then

$$W(x) = \prod_{k=0}^{n}(x - x_k) = 2^{-n}T_{n+1}$$

and the Lagrange interpolating polynomial $L_n(f)$, which we will write as $C_n(f)$ in order to distinguish it from the general case, is given by

$$C_n(f)(x) = \sum_{k=0}^{n} f(x_k)\,\frac{T_{n+1}(x)}{(x - x_k)T'_{n+1}(x_k)}.$$

But recall that for $x = \cos\theta$ we have

$$T'_n(x) = \frac{n\sin n\theta}{\sin\theta} = \frac{n\sin n\theta}{\sqrt{1-\cos^2\theta}} = \frac{n\sin n\theta}{\sqrt{1-x^2}}.$$

And so for $x_i = \cos((2i-1)\pi/2n)$, i.e., for $\theta_i = (2i-1)\pi/2n$, it follows that $\sin n\theta_i = \sin((2i-1)\pi/2) = (-1)^{i-1}$; that is,

$$\frac{1}{T'_n(x_i)} = \frac{(-1)^{i-1}\sqrt{1-x_i^2}}{n}.$$

It follows that our interpolation formula may be written as

$$C_n(f)(x) = \sum_{k=0}^{n} f(x_k)\, \frac{(-1)^{k-1} T_{n+1}(x)}{(n+1)(x-x_k)} \sqrt{1-x_k^2}\,. \tag{5.3}$$

Our first result is a refinement of Theorem 5.6 in light of Remark 5.7 (4).

Corollary 5.13. *Suppose that f has $n+1$ continuous derivatives on $[-1,1]$. Let $C_n(f) \in \mathcal{P}_n$ be the polynomial that interpolates f at the zeros of T_{n+1}, the $n+1$-st Chebyshev polynomial. Then*

$$|f(x) - C_n(f)(x)| \;\le\; \frac{1}{(n+1)!}\, \|f^{\,(n+1)}\|\, |2^{-n} T_{n+1}(x)| \;=\; \frac{1}{2^n (n+1)!}\, \|f^{\,(n+1)}\|.$$

Proof. As already noted, the auxiliary polynomial $W(x) = \prod_{i=0}^{n}(x-x_i)$ satisfies $W(x) = 2^{-n}T_{n+1}(x)$ and, of course, $|T_{n+1}(x)| \le 1$. □

While the Chebyshev nodes are not quite optimal for interpolation (see Problem 9), the interpolating polynomial is nevertheless remarkably close to the best approximation. Indeed, according to a result of Powell [42], if we set $\Theta_n(f) = \|f - C_n(f)\|$, then

$$E_n(f) \le \Theta_n(f) \le 4.037\, E_n(f) \text{ for } 1 \le n \le 25,$$

where, as usual, $E_n(f) = \|f - p_n^*\|$ is the minimum error in approximating f by a polynomial of degree at most n. Thus, in practice, for approximations by polynomials of modest degree, Chebyshev interpolation offers an attractive and efficient alternative to finding the best approximation.

Hermite Interpolation

A natural extension of the problem addressed at the beginning of this chapter is to increase the number of conditions on our interpolating polynomial and ask for the polynomial p of least degree that satisfies

$$p^{(m)}(x_k) = y_k^{(m)}, \quad k = 0, \ldots, n, \quad m = 0, 1, \ldots, \alpha_k - 1, \tag{5.4}$$

where the numbers $y_k^{(m)}$ are given in advance. In other words, we specify not only the values of the polynomial at each node x_k, we also specify the values of its first $\alpha_k - 1$ derivatives, where α_k is allowed to vary at each node.

Now the system of equations (5.4) imposes a total of $N = \alpha_0 + \cdots + \alpha_n$ conditions, so it's not at all surprising that we can find a polynomial of degree at most $N-1$ that fulfills them. Moreover, it's not much harder to see that p must be unique.

This problem was first solved by Hermite [27] in 1878. We won't address the general problem (but see Appendix F). Instead, we will settle for displaying the solution in the case where $\alpha_0 = \cdots = \alpha_n = 2$. In other words, the case where we specify only the first derivative y_k' of p at each node x_k. Note that in this case p will have degree at most $2n-1$.

Following the notation used throughout this chapter, we set

$$W(x) = (x-x_0)\cdots(x-x_n), \quad \text{and} \quad \ell_k(x) = \frac{W(x)}{(x-x_k)W'(x_k)}.$$

Now, as is easily verified,

$$\ell_k'(x) = \frac{(x - x_k)W'(x) - W(x)}{(x - x_k)^2 W'(x_k)}.$$

Thus, by l'Hôpital's rule,

$$\ell_k'(x_k) \;=\; \lim_{x \to x_k} \frac{(x - x_k)W''(x) + W'(x) - W'(x)}{2(x - x_k)W'(x_k)} \;=\; \frac{W''(x_k)}{2W'(x_k)}.$$

And now consider the polynomials

$$A_k(x) = \left[1 - \frac{W''(x_k)}{W'(x_k)}(x - x_k)\right] \ell_k(x)^2 \qquad \text{and} \qquad B_k(x) = (x - x_k)\,\ell_k(x)^2. \tag{5.5}$$

Note that A_k and B_k each have degree at most $2(n-1)+1 = 2n-1$. Moreover, it's easy to see that A_k and B_k satisfy

$$\begin{array}{rclcrcl} A_k(x_k) & = & 1 & \qquad & A_k'(x_k) & = & 0 \\ B_k(x_k) & = & 0 & \qquad & B_k'(x_k) & = & 1. \end{array}$$

Consequently, the polynomial

$$p(x) = \sum_{k=0}^{n} y_k A_k(x) + \sum_{k=0}^{n} y_k' B_k(x) \tag{5.6}$$

has degree at most $2n-1$ and satisfies

$$p(x_k) = y_k \quad \text{and} \quad p'(x_k) = y_k'$$

for $k = 0, \ldots, n$.

Exercise 5.14. Show that $\sum_{k=1}^{n} A_k(x) = 1$ for all x. [Hint: (5.6) is exact for all polynomials of degree at most $2n-1$.]

By combining the process just described with interpolation at the Chebyshev nodes, Fejér was able to prove the following result (see Natanson [41, Vol. III] for much more on this result).

Theorem 5.15. *Let $f \in C[-1,1]$ and, for each n, let H_n denote the polynomial of degree at most $2n-1$ that satisfies*

$$H_n(x_k^{(n)}) = f(x_k^{(n)}) \quad \textit{and} \quad H_n'(x_k^{(n)}) = 0,$$

for all k, where

$$x_k^{(n)} = \cos\theta_k^{(n)} = \cos\left(\frac{(2k-1)\pi}{2n}\right), \quad k = 1, \ldots, n$$

are the zeros of the Chebyshev polynomial T_n. Then $H_n \rightrightarrows f$ on $[-1,1]$.

Proof. As usual, we will think of n as fixed and suppress the superscript (n); that is, we will write x_k in place of $x_k^{(n)}$.

For the Chebyshev nodes, recall that we may take $W(x) = T_n(x)$. Next, let's compute $W'(x_k)$, $\ell_k(x)$, and $W''(x_k)$. As we've already seen,

$$T_n'(x_k) = n\frac{\sin n\theta_k}{\sin\theta_k} = \frac{(-1)^{n-1}n}{\sqrt{1-x_k^2}},$$

and, thus,

$$\ell_k(x) = \frac{(-1)^{k-1}T_n(x)}{n(x-x_k)}\sqrt{1-x_k^2}\,.$$

Next,

$$T_n''(x_k) = \frac{n\sin n\theta_k\cos\theta_k - n^2\cos n\theta_k\sin\theta_k}{\sin^3\theta_k} = \frac{(-1)^{n-1}nx_k}{(1-x_k)^{3/2}},$$

and so, in the notation of (5.5), we have

$$\frac{W''(x_k)}{W'(x_k)} = \frac{T_n''(x_k)}{T_n'(x_k)} = \frac{x_k}{1-x_k^2}$$

and hence, after a bit of rewriting,

$$A_k(x) = \left[1 - \frac{W''(x_k)}{W'(x_k)}(x-x_k)\right]\ell_k(x)^2 = \left[\frac{T_n(x)}{n(x-x_k)}\right]^2(1-xx_k).$$

Please note, in particular, that $A_k(x) \geq 0$ for all x in $[-1,1]$, an observation that will play a critical role in the remainder of the proof. Also note that $A_k(x) \leq 2/n^2(x-x_k)^2$ for all $x \neq x_k$ in $[-1,1]$.

Now according to equation (5.6), the polynomial H_n is given by

$$H_n(x) = \sum_{k=1}^n f(x_k)A_k(x)$$

where, as we know, $\sum_{k=1}^n A_k(x) = 1$ for all x in $[-1,1]$. In particular, we have

$$|H_n(x) - f(x)| \leq \sum_{k=1}^n |f(x_k) - f(x)|A_k(x).$$

The rest of the proof follows familiar lines: Given $\varepsilon > 0$, we choose $\delta > 0$ so that $|f(x) - f(y)| < \varepsilon$ whenever $|x-y| < \delta$. For x fixed, let $I \subset \{1,\ldots,n\}$ denote the indices k for which $|x-x_k| < \delta$ and let J denote the indices k for which $|x-x_k| \geq \delta$. Then

$$\sum_{k\in I}|f(x_k)-f(x)|A_k(x) \ < \ \varepsilon\sum_{k=1}^n|f(x_k)-f(x)|A_k(x) \ = \ \varepsilon$$

while

$$\sum_{k\in J}|f(x_k)-f(x)|A_k(x) \ \leq \ 2\|f\|\cdot n\cdot\frac{2}{n^2\delta^2} \ < \ \varepsilon$$

provided that n is sufficiently large. □

The Inequalities of Markov and Bernstein

As it happens, finding the polynomial of degree n that best approximates $f \in C[\,a, b\,]$ can be accomplished through an algorithm that finds the polynomial of best approximation over a set of only $n+2$ (appropriately chosen) points. We won't pursue the algorithm here, but for full details, see the discussion of the *One Point Exchange Algorithm* in Rivlin [45]. On the other hand, one aspect of this algorithm is of great interest and, moreover, is an easy application of the techniques presented in this chapter.

In particular, a key step in the algorithm requires a bound on differentiation over $\mathcal{P}_n$ of the form $\|p'\| \le C_n \|p\|$, where C_n is some constant depending only on n. The fact that such a constant exists is nearly obvious by itself (after all, $\mathcal{P}_n$ is finite-dimensional), but the value of the best constant C_n is of independent interest (and of some historical significance, too).

The inequality we'll prove is due to A. A. Markov from 1889:

Theorem 5.16. (Markov's Inequality) *If $p \in \mathcal{P}_n$, and if $|p(x)| \le 1$ for $|x| \le 1$, then $|p\,'(x)| \le n^2$ for $|x| \le 1$. Moreover, $|p\,'(x)| = n^2$ can only occur at $x = \pm 1$, and only when $p = \pm\, T_n$, the Chebyshev polynomial of degree n.*

Markov's brother, V. A. Markov, later improved on this, in 1916, by showing that $|p^{(k)}(x)| \le T_n^{(k)}(1)$. We've alluded to this fact already (see Rivlin [45, p. 31]), and even more is true. However, we'll settle for the somewhat looser bound given in the theorem.

About 20 years after Markov, in 1912, Bernstein asked for a similar bound for the derivative of a complex polynomial over the unit disk $|z| \le 1$. Now the maximum modulus theorem tells us that we may reduce to the case $|z| = 1$, that is, $z = e^{i\theta}$, and so Bernstein was able to restate the problem in terms of trig polynomials.

Theorem 5.17. (Bernstein's Inequality) *If $S \in \mathcal{T}_n$, and if $|S(\theta)| \le 1$, then $|S\,'(\theta)| \le n$. Equality is only possible for $S(\theta) = \sin n(\theta - \theta_0)$.*

Our plan is to deduce Markov's inequality from Bernstein's inequality by a method of proof due to Pólya and Szegö in 1928. However, we will not prove the assertions about equality in either theorem.

To begin, let's consider the Lagrange interpolation formula in the case where $x_i = \cos((2i-1)\pi/2n)$, $i = 1, \ldots, n$, are the zeros of the Chebyshev polynomial T_n. Recall that we have $-1 < x_n < x_{n-1} < \cdots < x_1 < 1$. (Compare the following result with the calculations used in the proof of Fejér's Theorem 5.15.)

Lemma 5.18. *Each polynomial $p \in \mathcal{P}_{n-1}$ may be written*

$$p(x) \;=\; \frac{1}{n} \sum_{i=1}^{n} p(x_i) \cdot (-1)^{i-1} \sqrt{1 - x_i^2} \cdot \frac{T_n(x)}{x - x_i}$$

where $x_1, \ldots, x_n$ are the zeros of T_n, the n-th Chebyshev polynomial.

Proof. This follows from equation (5.3), with $n+1$ is replaced by n, and the fact that Lagrange interpolation is exact for polynomials of degree $< n$. $\square$

Lemma 5.19. *For any polynomial $p \in \mathcal{P}_{n-1}$, we have*

$$\max_{-1 \le x \le 1} |p(x)| \;\le\; \max_{-1 \le x \le 1} \bigl|\, n \sqrt{1 - x^2}\; p(x) \bigr|.$$

Proof. To save wear and tear, let's write $M = \max_{-1\le x\le 1}\bigl|\, n\sqrt{1-x^2}\, p(x)\bigr|$.

First consider an x in the interval $[\,x_n, x_1\,]$; that is, $|x| \le \cos(\pi/2n) = x_1$. In this case we can estimate $\sqrt{1-x^2}$ from below:

$$\sqrt{1-x^2} \;\ge\; \sqrt{1-x_1^2} \;=\; \sqrt{1-\cos^2\left(\frac{\pi}{2n}\right)} \;=\; \sin\left(\frac{\pi}{2n}\right) \;\ge\; \frac{1}{n},$$

because $\sin\theta \ge 2\theta/\pi$ for $0 \le \theta \le \pi/2$ (from the mean value theorem). Hence, for $|x| \le \cos(\pi/2n)$, we get $|p(x)| \le n\sqrt{1-x^2}\,|p(x)| \le M$.

Now, for x outside the interval $[\,x_n, x_1\,]$, we apply our interpolation formula. In this case, each of the factors $x - x_i$ is of the *same sign.* Thus,

$$\begin{aligned} |p(x)| \;&=\; \frac{1}{n}\left|\sum_{i=1}^n p(x_i)\,\frac{(-1)^{i-1}\sqrt{1-x_i^2}\,T_n(x)}{x-x_i}\right| \\ &\le\; \frac{M}{n^2}\sum_{i=1}^n\left|\frac{T_n(x)}{x-x_i}\right| \;=\; \frac{M}{n^2}\left|\sum_{i=1}^n\frac{T_n(x)}{x-x_i}\right|. \end{aligned}$$

But,

$$\sum_{i=1}^n\frac{T_n(x)}{x-x_i} \;=\; T_n'(x) \qquad \text{(why?)}$$

and we know that $|T_n'(x)| \le n^2$. Thus, $|p(x)| \le M$. $\square$

We next turn our attention to trig polynomials. As usual, given an algebraic polynomial $p \in \mathcal{P}_n$, we will sooner or later consider $S(\theta) = p(\cos\theta)$. In this case, $S'(\theta) = p'(\cos\theta)\sin\theta$ is an odd trig polynomial of degree at most n and $S'(\theta) = p'(\cos\theta)\sin\theta = p'(x)\sqrt{1-x^2}$.

Conversely, if $S \in \mathcal{T}_n$ is an odd trig polynomial, then $S(\theta)/\sin\theta$ is even, and so may be written $S(\theta)/\sin\theta = p(\cos\theta)$ for some algebraic polynomial p of degree at most $n-1$. Thus, from Lemma 5.19,

$$\max_{0\le\theta\le 2\pi}\left|\frac{S(\theta)}{\sin\theta}\right| \;=\; \max_{0\le\theta\le 2\pi}|p(\cos\theta)| \;\le\; n\max_{0\le\theta\le 2\pi}|p(\cos\theta)\sin\theta| \;=\; n\max_{0\le\theta\le 2\pi}|S(\theta)|.$$

This proves

Corollary 5.20. *If $S \in \mathcal{T}_n$ is an odd trig polynomial, then*

$$\max_{0\le\theta\le 2\pi}\left|\frac{S(\theta)}{\sin\theta}\right| \;\le\; n\max_{0\le\theta\le 2\pi}|S(\theta)|.$$

Now we're ready for Bernstein's inequality.

Theorem 5.21. *If $S \in \mathcal{T}_n$, then*

$$\max_{0\le\theta\le 2\pi}|S'(\theta)| \;\le\; n\max_{0\le\theta\le 2\pi}|S(\theta)|.$$

Proof. We first define an auxiliary function $f(\alpha,\theta) = \big[S(\alpha+\theta) - S(\alpha-\theta)\big]/2$. For α fixed, $f(\alpha,\theta)$ is an odd trig polynomial in θ of degree at most n. Consequently,

$$\left|\frac{f(\alpha,\theta)}{\sin\theta}\right| \;\le\; n \max_{0\le\theta\le 2\pi} |f(\alpha,\theta)| \;\le\; n \max_{0\le\theta\le 2\pi} |S(\theta)|.$$

But

$$S'(\alpha) \;=\; \lim_{\theta\to 0} \frac{S(\alpha+\theta) - S(\alpha-\theta)}{2\theta} \;=\; \lim_{\theta\to 0} \frac{f(\alpha,\theta)}{\sin\theta},$$

and hence $|S'(\alpha)| \le n \max_{0\le\theta\le 2\pi} |S(\theta)|$. □

Finally, we prove Markov's inequality.

Theorem 5.22. *If* $p \in \mathcal{P}_n$, *then* $\max_{-1\le x\le 1} |p\,'(x)| \;\le\; n^2 \max_{-1\le x\le 1} |p(x)|$.

Proof. We know that $S(\theta) = p(\cos\theta)$ is a trig polynomial of degree at most n satisfying

$$\max_{-1\le x\le 1} |p(x)| \;=\; \max_{0\le\theta\le 2\pi} |p(\cos\theta)|.$$

Because $S'(\theta) = p\,'(\cos\theta)\sin\theta$ is also trig polynomial of degree at most n, Bernstein's inequality yields

$$\max_{0\le\theta\le 2\pi} |p'(\cos\theta)\sin\theta| \;\le\; n \max_{0\le\theta\le 2\pi} |p(\cos\theta)|.$$

In other words,

$$\max_{-1\le x\le 1} \left|p\,'(x)\sqrt{1-x^2}\,\right| \;\le\; n \max_{-1\le x\le 1} |p(x)|.$$

Because $p\,' \in \mathcal{P}_{n-1}$, the desired inequality now follows easily from Lemma 5.19.

$$\max_{-1\le x\le 1} |p\,'(x)| \;\le\; n \max_{-1\le x\le 1} \left|p\,'(x)\sqrt{1-x^2}\,\right| \;\le\; n^2 \max_{-1\le x\le 1} |p(x)|.$$ □

Problems

1. Let $y_0, y_1, \ldots, y_n \in \mathbb{R}$ be given. Show that the polynomial $p \in \mathcal{P}_n$ satisfying $p(x_i) = y_i$, $i = 0, 1, \ldots, n$, may be written as

$$p(x) \;=\; c \begin{vmatrix} 0 & 1 & x & x^2 & \cdots & x^n \\ y_0 & 1 & x_0 & x_0^2 & \cdots & x_0^n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ y_n & 1 & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix},$$

where c is a certain constant. Find c and prove the formula.

Throughout, $x_0, \ldots, x_n$ are distinct points in some interval $[\,a, b\,]$; $W(x) = \prod_{i=0}^n (x - x_i)$; $\ell_i(x)$, $i = 0, \ldots, n$, denote the fundamental system of polynomials defined by $\ell_i(x) = W(x)/(x - x_i)\, W'(x_i)$; and $L_n(f)(x) = \sum_{i=0}^n f(x_i)\ell_i(x)$ denotes the (unique) polynomial of degree at most n that interpolates to f at the nodes x_i. The associated Lebesgue function is given by $\lambda_n(x) = \sum_{i=0}^n |\ell_i(x)|$ and the Lebesgue number is $\Lambda_n = \|\lambda_n\| = \|L_n\|$.

2. Prove that L_n is a linear projection onto $\mathcal{P}_n$. That is, show that $L_n(\alpha g + \beta h) = \alpha L_n(g) + \beta L_n(h)$, for any g, $h \in C[\,a, b\,]$, α, $\beta \in \mathbb{R}$, and that $L_n(g) = g$ if and only if $g \in \mathcal{P}_n$.
3. Show that $\sum_{i=0}^n \ell_i(x) \equiv 1$. Conclude that $\lambda_n(x) \ge 1$ for all x and n and, hence, $\Lambda_n \ge 1$ for every n.
4. More generally, show that $\sum_{i=0}^n x_i^k \ell_i(x) = x^k$, for $k = 0, 1, \ldots, n$.
5. Show that $\|L_n(f)\| \le \Lambda_n \|f\|$ for all $f \in C[\,a, b\,]$. Show that no smaller number Λ has this property.
6. Show that the error in the Lagrange interpolation formula at a given point x can be written as $(L_n(f) - f)(x) = \sum_{i=0}^n [\, f(x_i) - f(x)\,]\, \ell_i(x)$.
7. Let $f \in C[\,a, b\,]$.
   (a) If, at some point x^* in $[\,a, b\,]$, we have $\lim_{n\to\infty} \lambda_n(x^*)\, E_n(f) = 0$, show that $L_n(f)(x^*) \to f(x^*)$ as $n \to \infty$. [Hint: Examine the proof of Lebesgue's Theorem.]
   (b) If we have $\lim_{n\to\infty} \Lambda_n E_n(f) = 0$, then $L_n(f)$ converges uniformly to f on $[\,a, b\,]$.
8. In the complex plane, show that the polynomial $p_{n-1}(z) = z^{n-1}$ interpolates to the function $f(z) = 1/z$ at the n-th roots of unity, $z_k = e^{2\pi i k/n}$, $k = 0, \ldots, n-1$. Show, too, that $\|f - p_{n-1}\| = \sqrt{2} \not\to 0$ as $n \to \infty$, where $\|\cdot\|$ denotes the sup-norm over the unit disk $\mathbb{T} = \{z : |z| = 1\}$.
9. Let $-1 \le x_0 < x_1 < x_2 \le 1$ and let $\Lambda = \|\sum_{k=0}^2 |\ell_k(x)|\,\|$ denote the corresponding Lebesgue number. Show that the minimum value of Λ is $5/4$ and is attained when $-x_0 = x_2 \ge \frac{2}{3}\sqrt{2}$ and $x_1 = 0$. Meanwhile, if x_0, x_1, and x_2 are chosen to be the roots of T_3, show that the value of Λ is $5/3$. Thus, placing the nodes at the zeros of T_n does not, in general, lead to a minimum value for Λ_n.
10. Show that the Hermite interpolating polynomial of degree at most $2n - 1$; i.e., the solution to the system of equations (5.4), is unique.

Chapter 6

A Brief Introduction to Fourier Series

The *Fourier series* of a 2π-periodic (bounded, integrable) function f is

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} \big(a_k \cos kx + b_k \sin kx\big),$$

where the coefficients are defined by

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)\, \cos kt\, dt \qquad \text{and} \qquad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)\, \sin kt\, dt.$$

Please note that if f is Riemann integrable on $[-\pi, \pi\,]$, then each of these integrals is well-defined and finite; indeed,

$$|a_k| \le \frac{1}{\pi} \int_{-\pi}^{\pi} |f(t)|\, dt$$

and so, for example, we would have $|a_k| \le 2\|f\|$ for $f \in C^{2\pi}$.

We write the partial sums of the series as

$$s_n(f)(x) = \frac{a_0}{2} + \sum_{k=1}^{n} \big(a_k \cos kx + b_k \sin kx\big).$$

Now while $s_n(f)$ need not converge pointwise to f (in fact, it may even diverge at a given point), and while $s_n(f)$ is not typically a good uniform approximation to f, it is still a very natural choice for an approximation to f in the "least squares" sense (which we'll make precise shortly). Said in other words, the Fourier series for f will provide a useful *representation* for f even if it fails to converge pointwise to f.

We begin with a (rather long but entirely elementary) series of observations.

Remarks 6.1.

1. The collection of functions $1, \cos x, \cos 2x, \ldots$, and $\sin x, \sin 2x, \ldots$, are *orthogonal* on $[-\pi, \pi\,]$. That is,

$$\int_{-\pi}^{\pi} \cos mx\, \cos nx\, dx = \int_{-\pi}^{\pi} \sin mx\, \sin nx\, dx = \int_{-\pi}^{\pi} \cos mx\, \sin nx\, dx = 0$$

for any $m \neq n$ (and the last equation even holds for $m = n$),

$$\int_{-\pi}^{\pi} \cos^2 mx\,dx = \int_{-\pi}^{\pi} \sin^2 mx\,dx = \pi$$

for any $m \neq 0$, and, of course, $\int_{-\pi}^{\pi} 1\,dx = 2\pi$.

2. What this means is that if $T(x) = \frac{\alpha_0}{2} + \sum_{k=1}^{n}\big(\alpha_k \cos kx + \beta_k \sin kx\big)$, then

$$\frac{1}{\pi}\int_{-\pi}^{\pi} T(x)\,\cos mx\,dx = \frac{\alpha_m}{\pi}\int_{-\pi}^{\pi} \cos^2 mx\,dx = \alpha_m$$

for $m \neq 0$, while

$$\frac{1}{\pi}\int_{-\pi}^{\pi} T(x)\,dx = \frac{\alpha_0}{2\pi}\int_{-\pi}^{\pi} dx = \alpha_0.$$

That is, if $T \in \mathcal{T}_n$, then T is actually equal to its own Fourier series.

3. The partial sum operator $s_n(f)$ is a *linear projection* from $C^{2\pi}$ onto $\mathcal{T}_n$.

4. If $T(x) = \frac{\alpha_0}{2} + \sum_{k=1}^{n}\big(\alpha_k \cos kx + \beta_k \sin kx\big)$ is a trig polynomial, then

$$\begin{aligned}
\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,T(x)\,dx &= \frac{\alpha_0}{2\pi}\int_{-\pi}^{\pi} f(x)\,dx + \sum_{k=1}^{n}\frac{\alpha_k}{\pi}\int_{-\pi}^{\pi} f(x)\,\cos kx\,dx \\
&\qquad + \sum_{k=1}^{n}\frac{\beta_k}{\pi}\int_{-\pi}^{\pi} f(x)\,\sin kx\,dx \\
&= \frac{\alpha_0 a_0}{2} + \sum_{k=1}^{n}\big(\alpha_k a_k + \beta_k b_k\big),
\end{aligned}$$

where (a_k) and (b_k) are the Fourier coefficients for f. [This should remind you of the *dot product* of the coefficients.]

5. Motivated by Remarks 1, 2, and 4, we define the *inner product* of two elements f, $g \in C^{2\pi}$ by

$$\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,g(x)\,dx.$$

(Be forewarned: Some authors prefer the normalizing factor $1/2\pi$ in place of $1/\pi$ here.) Note that from Remark 4 we have $\langle f, s_n(f)\rangle = \langle s_n(f), s_n(f)\rangle$ for any n. (Why?)

6. If some $f \in C^{2\pi}$ has $a_k = b_k = 0$ for all k, then $f \equiv 0$.

Proof. Indeed, by Remark 4 (or linearity of the integral), this means that

$$\int_{-\pi}^{\pi} f(x)\,T(x)\,dx = 0$$

for any trig polynomial T. But from Weierstrass's second theorem we know that f is the uniform limit of some sequence of trig polynomials (T_n). Thus,

$$\int_{-\pi}^{\pi} f(x)^2\,dx = \lim_{n\to\infty}\int_{-\pi}^{\pi} f(x)\,T_n(x)\,dx = 0.$$

Because f is continuous, this easily implies that $f \equiv 0$. $\square$

7. If f, $g \in C^{2\pi}$ have the same Fourier series, then $f \equiv g$. Hence, the Fourier series for an $f \in C^{2\pi}$ provides a *unique representation* for f (even if the series fails to converge to f).

8. The coefficients $a_0, a_1, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ minimize the expression

$$\varphi(a_0, a_1, \ldots, b_n) = \int_{-\pi}^{\pi} \big[f(x) - s_n(f)(x)\big]^2\,dx.$$

Proof. It's not hard to see, for example, that

$$\frac{\partial\,\varphi}{\partial\,a_k} = \int_{-\pi}^{\pi} 2\big[f(x) - s_n(f)(x)\big] \cos kx\,dx = 0$$

precisely when a_k satisfies

$$\int_{-\pi}^{\pi} f(x)\,\cos kx\,dx = a_k \int_{-\pi}^{\pi} \cos^2 kx\,dx. \qquad \square$$

9. The partial sum $s_n(f)$ is the best approximation to f out of $\mathcal{T}_n$ *relative to the* L_2 *norm*

$$\|f\|_2 \;=\; \sqrt{\langle f, f\rangle} \;=\; \left(\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)^2\,dx\right)^{1/2}.$$

That is,

$$\|f - s_n(f)\|_2 = \min_{T\in\mathcal{T}_n} \|f - T\|_2.$$

Moreover, using Remarks 4 and 5, we have

$$\begin{aligned} \|f - s_n(f)\|_2^2 &= \langle f - s_n(f), f - s_n(f)\rangle \\ &= \langle f, f\rangle - 2\,\langle f, s_n(f)\rangle + \langle s_n(f), s_n(f)\rangle \\ &= \|f\|_2^2 - \|s_n(f)\|_2^2 \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)^2\,dx - \frac{a_0^2}{2} - \sum_{k=1}^{n}(a_k^2 + b_k^2). \end{aligned}$$

[This should remind you of the *Pythagorean theorem.*]

10. It follows from Remark 9 that

$$\frac{1}{\pi}\int_{-\pi}^{\pi} s_n(f)(x)^2\,dx \;=\; \frac{a_0^2}{2} + \sum_{k=1}^{n}\big(a_k^2 + b_k^2\big) \;\le\; \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)^2\,dx.$$

In other symbols, $\|s_n(f)\|_2 \le \|f\|_2$. In particular, the Fourier coefficients of any $f \in C^{2\pi}$ are square summable. (Why?)

11. If $f \in C^{2\pi}$, then its Fourier coefficients (a_n) and (b_n) tend to zero as $n \to \infty$.

12. It follows from Remark 10 and Weierstrass's second theorem that $s_n(f) \to f$ *in the* L_2 *norm* whenever $f \in C^{2\pi}$. Indeed, given $\varepsilon > 0$, choose a trig polynomial T such that $\|f - T\| < \varepsilon$. Then, because $s_n(T) = T$ for large enough n, we have

$$\begin{aligned} \|f - s_n(f)\|_2 &\le \|f - T\|_2 + \|s_n(T - f)\|_2 \\ &\le 2\|f - T\|_2 \\ &\le 2\sqrt{2}\,\|f - T\| < 2\sqrt{2}\,\varepsilon, \end{aligned}$$

where the penultimate inequality follows from the easily verifiable fact that $\|f\|_2 \le \sqrt{2}\,\|f\|$ for any $f \in C^{2\pi}$. (Compare this calculation with Lebesgue's Theorem 5.9.)

By way of comparison, let's give a simple class of functions whose Fourier partial sums provide good *uniform* approximations.

Theorem 6.2. *If $f'' \in C^{2\pi}$, then the Fourier series for f converges absolutely and uniformly to f.*

Proof. First notice that integration by-parts leads to an estimate on the order of growth of the Fourier coefficients of f.

$$\pi a_k = \int_{-\pi}^{\pi} f(x)\,\cos kx\,dx = \int_{-\pi}^{\pi} f(x)\,d\left(\frac{\sin kx}{k}\right) = -\frac{1}{k}\int_{-\pi}^{\pi} f'(x)\,\sin kx\,dx$$

(because f is 2π-periodic). Thus, $|a_k| \le 2\|f'\|/k \to 0$ as $k \to \infty$. Now we integrate by-parts again:

$$-\pi k a_k = \int_{-\pi}^{\pi} f'(x)\,\sin kx\,dx = \int_{-\pi}^{\pi} f'(x)\,d\left(\frac{\cos kx}{k}\right) = \frac{1}{k}\int_{-\pi}^{\pi} f''(x)\,\cos kx\,dx$$

(because f' is 2π-periodic). Thus, $|a_k| \le 2\|f''\|/k^2 \to 0$ as $k \to \infty$. More importantly, this inequality (along with the Weierstrass M-test) implies that the Fourier series for f is both uniformly and absolutely convergent:

$$\left|\frac{a_0}{2} + \sum_{k=1}^{\infty}\big(a_k\cos kx + b_k\sin kx\big)\right| \le \left|\frac{a_0}{2}\right| + \sum_{k=1}^{\infty}\big(|a_k| + |b_k|\big) \le C\sum_{k=1}^{\infty}\frac{1}{k^2}.$$

But why should the series actually converge to f? Well, if we call the sum

$$g(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty}\big(a_k\cos kx + b_k\sin kx\big),$$

then $g \in C^{2\pi}$ (why?) and g has the same Fourier coefficients as f (why?). Hence (by Remarks 6.1 (7), above, $g = f$. □

Our next chore is to find a closed expression for $s_n(f)$. For this we'll need a couple of trig identities; the first two need no explanation.

$$\cos kt\cos kx + \sin kt\sin kx = \cos k(t - x)$$
$$2\cos\alpha\sin\beta = \sin(\alpha+\beta) - \sin(\alpha-\beta)$$
$$\tfrac{1}{2} + \cos\theta + \cos 2\theta + \cdots + \cos n\theta = \frac{\sin\,(n+\frac{1}{2})\,\theta}{2\sin\frac{1}{2}\theta}$$

Here's a short proof for the third:

$$\begin{aligned} \sin \tfrac{1}{2}\theta + \textstyle\sum_{k=1}^n 2\cos k\theta \, \sin \tfrac{1}{2}\theta &= \sin \tfrac{1}{2}\theta + \textstyle\sum_{k=1}^n \big[\sin\,(k+\tfrac{1}{2})\theta - \sin\,(k-\tfrac{1}{2})\theta \,\big] \\ &= \sin\,(n+\tfrac{1}{2})\theta. \end{aligned}$$

The function

$$D_n(t) = \frac{\sin\,(n+\frac{1}{2})\,t}{2\sin\frac{1}{2}\,t}$$

is called *Dirichlet's kernel*. It plays an important role in our next calculation.

We're now ready to re-write our formula for $s_n(f)$.

$$\begin{aligned} s_n(f)(x) &= \tfrac{1}{2}a_0 + \sum_{k=1}^n \big(a_k \cos kx + b_k \sin kx\big) \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left[\tfrac{1}{2} + \sum_{k=1}^n \cos kt \cos kx + \sin kt \sin kx\right] dt \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left[\tfrac{1}{2} + \sum_{k=1}^n \cos k(t-x)\right] dt \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cdot\frac{\sin\,(n+\frac{1}{2})\,(t-x)}{2\sin\frac{1}{2}\,(t-x)}\,dt \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,D_n(t-x)\,dt = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\,D_n(t)\,dt. \end{aligned}$$

It now follows easily that $s_n(f)$ is linear in f (because integration against D_n is linear), that $s_n(f) \in \mathcal{T}_n$ (because $D_n \in \mathcal{T}_n$), and, in fact, that $s_n(\mathcal{T}_m) = \mathcal{T}_{\min(m,n)}$. In other words, s_n is indeed a linear projection onto $\mathcal{T}_n$.

While we know that $s_n(f)$ is a good approximation to f in the L_2 norm, a better understanding of its effectiveness as a uniform approximation will require a better understanding of the Dirichlet kernel D_n. Here are a few pertinent facts:

Lemma 6.3. (a) *D_n is even,*

(b) $\displaystyle \frac{1}{\pi}\int_{-\pi}^{\pi} D_n(t)\,dt = \frac{2}{\pi}\int_0^{\pi} D_n(t)\,dt = 1,$

(c) $|D_n(t)| \le n+\frac{1}{2}$ *and* $D_n(0) = n + \frac{1}{2}$,

(d) $\displaystyle \frac{|\sin\,(n+\frac{1}{2})\,t\,|}{t} \le |D_n(t)| \le \frac{\pi}{2t}$ *for* $0 < t < \pi$,

(e) *If* $\displaystyle \lambda_n = \frac{1}{\pi}\int_{-\pi}^{\pi} |D_n(t)|\,dt,$ *then* $\displaystyle \frac{4}{\pi^2}\log n \le \lambda_n \le 3 + \log n.$

Proof. (a), (b), and (c) are relatively clear from the fact that

$$D_n(t) = \tfrac{1}{2} + \cos t + \cos 2t + \cdots + \cos nt.$$

(Notice, too, that (b) follows from the fact that $s_n(1) = 1$.) For (d) we use a more delicate estimate: Because $2\theta/\pi \le \sin\theta \le \theta$ for $0 < \theta < \pi/2$, it follows that $2t/\pi \le 2\sin(t/2) \le t$ for $0 < t < \pi$. Hence,

$$\frac{\pi}{2t} \;\ge\; \frac{|\sin(n+\frac{1}{2})\,t\,|}{2\sin\frac{1}{2}t} \;\ge\; \frac{|\sin(n+\frac{1}{2})\,t\,|}{t}$$

for $0 < t < \pi$. Next, the upper estimate in (e) is easy:

$$\begin{aligned}
\frac{2}{\pi}\int_0^{\pi} |D_n(t)|\,dt &= \frac{2}{\pi}\int_0^{\pi} \frac{|\sin(n+\frac{1}{2})\,t\,|}{2\sin\frac{1}{2}t}\,dt \\
&\le \frac{2}{\pi}\int_0^{1/n} (n+\tfrac{1}{2})\,dt + \frac{2}{\pi}\int_{1/n}^{\pi} \frac{\pi}{2t}\,dt \\
&= \frac{2n+1}{\pi n} + \log\pi + \log n \\
&< 3 + \log n.
\end{aligned}$$

The lower estimate takes some work:

$$\begin{aligned}
\frac{2}{\pi}\int_0^{\pi} |D_n(t)|\,dt &= \frac{2}{\pi}\int_0^{\pi} \frac{|\sin(n+\frac{1}{2})\,t\,|}{2\sin\frac{1}{2}t}\,dt \\
&\ge \frac{2}{\pi}\int_0^{\pi} \frac{|\sin(n+\frac{1}{2})\,t\,|}{t}\,dt \\
&= \frac{2}{\pi}\int_0^{(n+\frac{1}{2})\pi} \frac{|\sin x\,|}{x}\,dx \\
&\ge \frac{2}{\pi}\int_0^{n\pi} \frac{|\sin x\,|}{x}\,dx \\
&= \frac{2}{\pi}\sum_{k=1}^{n}\int_{(k-1)\pi}^{k\pi} \frac{|\sin x\,|}{x}\,dx \\
&\ge \frac{2}{\pi}\sum_{k=1}^{n}\frac{1}{k\pi}\int_{(k-1)\pi}^{k\pi} |\sin x\,|\,dx \\
&= \frac{4}{\pi^2}\sum_{k=1}^{n}\frac{1}{k} \;\ge\; \frac{4}{\pi^2}\log n,
\end{aligned}$$

because $\sum_{k=1}^{n}\frac{1}{k} \ge \log n$. □

The numbers $\lambda_n = \|D_n\|_1 = \frac{1}{\pi}\int_{-\pi}^{\pi} |D_n(t)|\,dt$ are called the *Lebesgue numbers* associated to this process (compare this to the terminology we used for interpolation). The point here is that λ_n gives the norm of the partial sum operator (projection) on $C^{2\pi}$ and (just as with interpolation) $\lambda_n \to \infty$ as $n \to \infty$. As a matter of no small curiosity, notice that, from Remarks 6.1 (10), the norm of s_n as an operator on L_2 is 1.

Corollary 6.4. *If $f \in C^{2\pi}$, then*

$$|s_n(f)(x)| \le \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x+t)|\,|D_n(t)|\,dt \le \lambda_n\|f\|. \tag{6.1}$$

In particular, $\|s_n(f)\| \le \lambda_n\|f\| \le (3+\log n)\|f\|$.

If we approximate the function $\operatorname{sgn} D_n$ by a continuous function f of norm one, then

$$s_n(f)(0) \approx \frac{1}{\pi}\int_{-\pi}^{\pi} |D_n(t)|\,dt = \lambda_n.$$

Thus, λ_n is the smallest constant that works in equation (6.1). The fact that the partial sum operators are not uniformly bounded on $C^{2\pi}$, along with the uniform boundedness theorem, tells us that there must be some $f \in C^{2\pi}$ for which $\|s_n(f)\|$ is unbounded. But, as we've seen, this has more to do with certain projections than with Fourier series; indeed, a version of the Kharshiladze-Lozinski Theorem is valid in this setting, too (cf. Theorem 5.5).

Theorem 6.5. (Kharshiladze, Lozinski) *For each n, let L_n be a continuous, linear projection from $C^{2\pi}$ onto $\mathcal{T}_n$. Then, there is some $f \in C^{2\pi}$ for which $\|L_n(f) - f\|$ is unbounded.*

Although Corollary 6.4 may not look very useful, it does give us some information about the effectiveness of $s_n(f)$ as a uniform approximation to f. Specifically, we have *Lebesgue's theorem*:

Theorem 6.6. *If $f \in C^{2\pi}$ and if we set $E_n^T(f) = \min_{T\in\mathcal{T}_n} \|f - T\|$, then*

$$E_n^T(f) \;\le\; \|f - s_n(f)\| \;\le\; (4 + \log n)\, E_n^T(f).$$

Proof. Let T^* be the best uniform approximation to f out of $\mathcal{T}_n$. Then, because $s_n(T^*) = T^*$, we get

$$\|f - s_n(f)\| \le \|f - T^*\| + \|s_n(T^* - f)\| \le (4 + \log n)\, \|f - T^*\|. \qquad \square$$

As an application of Lebesgue's theorem, let's speak briefly about "Chebyshev series," a notion that fits neatly between approximation by algebraic polynomials and by trig polynomials.

Theorem 6.7. *Suppose that $f \in C[-1,1\,]$ is twice continously differentiable. Then f may be written as a uniformly and absolutely convergent Chebyshev series; that is, $f(x) = \sum_{k=0}^{\infty} a_k T_k(x)$, where $\sum_{k=0}^{\infty} |a_k| < \infty$.*

Proof. As usual, consider $\varphi(\theta) = f(\cos\theta) \in C^{2\pi}$. Because φ is *even* and twice differentiable, its Fourier series is an absolutely and uniformly convergent cosine series:

$$f(\cos\theta) = \varphi(\theta) = \sum_{k=0}^{\infty} a_k \cos k\theta = \sum_{k=0}^{\infty} a_k T_k(\cos\theta),$$

where $|a_k| \le 2\|\varphi''\|/k^2$. Thus, $f(x) = \sum_{k=0}^{\infty} a_k T_k(x)$. $\square$

If we write $S_n(f)(x) = \sum_{k=0}^{n} a_k T_k(x)$, we get an interesting consequence of this Theorem. First, notice that

$$S_n(f)(\cos\theta) = s_n(\varphi)(\theta).$$

Thus, from Lebesgue's theorem,

$$\begin{aligned} E_n(f) \;\le\; \|f - S_n(f)\|_{C[-1,1\,]} \;&=\; \|\varphi - s_n(\varphi)\|_{C^{2\pi}} \\ &\le\; (4 + \log n)\, E_n^T(\varphi) \;=\; (4 + \log n)\, E_n(f). \end{aligned}$$

For $n < 400$, this reads

$$E_n(f) \;\le\; \|f - S_n(f)\| \;\le\; 10\,E_n(f).$$

That is, for numerical purposes, the error incurred by using $\sum_{k=0}^{n} a_k T_k(x)$ to approximate f is *within one decimal place accuracy* of the best approximation! Notice, too, that $E_n(f)$ would be very easy to estimate in this case because

$$E_n(f) \;\le\; \|f - S_n(f)\| \;=\; \left\| \sum_{k>n} a_k T_k \right\| \;\le\; \sum_{k>n} |a_k| \;\le\; 2\,\|\varphi''\| \sum_{k>n} \frac{1}{k^2}.$$

Lebesgue's theorem should remind you of our "fancy" version of Bernstein's theorem; if we knew that $E_n^T(f)\log n \to 0$ as $n \to \infty$, then we'd know that $s_n(f)$ converged uniformly to f. Our goal, then, is to improve our estimates on $E_n^T(f)$, and the idea behind these improvements is to *replace D_n by a better kernel* (with regard to uniform approximation). Before we pursue anything quite so delicate as an estimate on $E_n^T(f)$, though, let's investigate a simple (and useful) replacement for D_n.

Because the sequence of partial sums (s_n) need not converge to f, we might try looking at their arithmetic means (or Cesàro sums):

$$\sigma_n = \frac{s_0 + s_1 + \cdots + s_{n-1}}{n}.$$

(These averages typically have better convergence properties than the partial sums themselves. Consider σ_n in the (scalar) case $s_n = (-1)^n$, for example.) Specifically, we set

$$\begin{aligned}
\sigma_n(f)(x) &= \frac{1}{n}\Big[s_0(f)(x) + \cdots + s_{n-1}(f)(x)\Big] \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\,\left[\frac{1}{n}\sum_{k=0}^{n-1} D_k(t)\right] dt = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\,K_n(t)\,dt,
\end{aligned}$$

where $K_n = (D_0 + D_1 + \cdots + D_{n-1})/n$ is called *Fejér's kernel.* The same techniques we used earlier can be applied to find a closed form for $\sigma_n(f)$ which, of course, reduces to simplifying $(D_0 + D_1 + \cdots + D_{n-1})/n$. As before, we begin with a trig identity:

$$\begin{aligned}
2\sin\theta \sum_{k=0}^{n-1} \sin\,(2k+1)\theta &= \sum_{k=0}^{n-1}\big[\cos 2k\theta - \cos\,(2k+2)\theta\,\big] \\
&= 1 - \cos 2n\theta = 2\sin^2 n\theta.
\end{aligned}$$

Thus,

$$K_n(t) \;=\; \frac{1}{n}\sum_{k=0}^{n-1} \frac{\sin\,(2k+1)\,t/2}{2\sin\,(t/2)} \;=\; \frac{\sin^2(nt/2)}{2n\sin^2(t/2)}.$$

Please note that K_n is even, *nonnegative*, and $\frac{1}{\pi}\int_{-\pi}^{\pi} K_n(t)\,dt = 1$. Thus, $\sigma_n(f)$ is a positive, linear map from $C^{2\pi}$ onto $\mathcal{T}_n$ (but it's not a projection—why?), satisfying $\|\sigma_n(f)\|_2 \le \|f\|_2$ (why?).

Now the arithmetic mean operator $\sigma_n(f)$ is still a good approximation f in L_2 norm. Indeed,

$$\|f - \sigma_n(f)\|_2 = \frac{1}{n}\left\| \sum_{k=0}^{n-1} (f - s_k(f)) \right\|_2 \le \frac{1}{n}\sum_{k=0}^{n-1} \|f - s_k(f)\|_2 \to 0$$

as $n \to \infty$ (because $\|f - s_k(f)\|_2 \to 0$). But, more to the point, $\sigma_n(f)$ is actually a good *uniform* approximation to f, a fact that we'll call *Fejér's theorem*:

Theorem 6.8. *If $f \in C^{2\pi}$, then $\sigma_n(f)$ converges uniformly to f as $n \to \infty$.*

Note that, because $\sigma_n(f) \in \mathcal{T}_n$, Fejér's theorem implies Weierstrass's second theorem. Curiously, Fejér was only 19 years old when he proved this result (about 1900) while Weierstrass was 75 at the time he proved his approximation theorems.

We'll give two proofs of Fejér's theorem; one with details, one without. But both follow from quite general considerations. First:

Theorem 6.9. *Suppose that $k_n \in C^{2\pi}$ satisfies*

(a) $k_n \ge 0$,

(b) $\displaystyle\frac{1}{\pi}\int_{-\pi}^{\pi} k_n(t)\,dt = 1$, *and*

(c) $\displaystyle\int_{\delta \le |t| \le \pi} k_n(t)\,dt \to 0$ *for every* $\delta > 0$.

Then, $\displaystyle\frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\,k_n(t)\,dt \rightrightarrows f(x)$ *for each* $f \in C^{2\pi}$.

Proof. Let $\varepsilon > 0$. Because f is uniformly continuous, we may choose $\delta > 0$ so that $|f(x) - f(x+t)| < \varepsilon$, for any x, whenever $|t| < \delta$. Next, we use the fact that k_n is nonnegative and integrates to 1 to write

$$\begin{aligned}
\left| f(x) - \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\,k_n(t)\,dt \right| &= \frac{1}{\pi}\left| \int_{-\pi}^{\pi} \big[f(x) - f(x+t)\big]\,k_n(t)\,dt \right| \\
&\le \frac{1}{\pi}\int_{-\pi}^{\pi} \big|f(x) - f(x+t)\big|\,k_n(t)\,dt \\
&\le \frac{\varepsilon}{\pi}\int_{|t|<\delta} k_n(t)\,dt \;+\; \frac{2\|f\|}{\pi}\int_{\delta \le |t| \le \pi} k_n(t)\,dt \\
&< \varepsilon + \varepsilon = 2\varepsilon,
\end{aligned}$$

for n sufficiently large. □

To see that Fejér's kernel satisfies the conditions of the Theorem is easy: In particular, (c) follows from the fact that $K_n(t) \rightrightarrows 0$ on the set $\delta \le |t| \le \pi$. Indeed, because $\sin(t/2)$ increases on $\delta \le t \le \pi$ we have

$$K_n(t) \;=\; \frac{\sin^2(nt/2)}{2n\sin^2(t/2)} \;\le\; \frac{1}{2n\sin^2(\delta/2)} \;\to 0.$$

Our second proof, or sketch, really, is based on a variant of the Bohman-Korovkin theorem for $C^{2\pi}$ due to Korovkin. In this setting, the three "test cases" are

$$f_0(x) = 1, \qquad f_1(x) = \cos x, \qquad \text{and} \qquad f_2(x) = \sin x.$$

Theorem 6.10. *Let (L_n) be a sequence of positive, linear maps on $C^{2\pi}$. If $L_n(f) \rightrightarrows f$ for each of the three functions $f_0(x) = 1$, $f_1(x) = \cos x$, and $f_2(x) = \sin x$, then $L_n(f) \rightrightarrows f$ for every $f \in C^{2\pi}$.*

We won't prove this theorem; rather, we'll check that $\sigma_n(f) \rightrightarrows f$ in each of the three test cases. Because s_n is a projection, this is painfully simple!

$$\begin{aligned}
\sigma_n(f_0) &= \tfrac{1}{n}\,(f_0 + f_0 + \cdots + f_0) = f_0, \\
\sigma_n(f_1) &= \tfrac{1}{n}\,(0 + f_1 + \cdots + f_1) = \tfrac{n-1}{n} \cdot f_1 \rightrightarrows f_1, \\
\sigma_n(f_2) &= \tfrac{1}{n}\,(0 + f_2 + \cdots + f_2) = \tfrac{n-1}{n} \cdot f_2 \rightrightarrows f_2.
\end{aligned}$$

Kernel operators abound in analysis; for example, Landau's proof of the Weierstrass theorem uses the kernel $L_n(x) = c_n(1 - x^2)^n$. And, in the next chapter, we'll encounter *Jackson's kernel*, $J_n(t) = c_n \sin^4 nt/n^3 \sin^4 t$, which is essentially the square of Fejér's kernel. While we will have no need for a general theory of such operators, please note that the key to their utility is the fact that they're nonnegative!

Lastly, a word or two about Fourier series involving complex coefficients. Most modern textbooks consider the case of a 2π-periodic, integrable function $f : \mathbb{R} \to \mathbb{C}$ and define the Fourier series of f by

$$\sum_{k=-\infty}^{\infty} c_k e^{ikt},$$

where now we have only one formula for the c_k:

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\, e^{-ikt}\, dt,$$

but, of course, the c_k may well be complex. This somewhat simpler approach has other advantages; for one, the exponentials e^{ikt} are now an orthonormal set—relative to the normalizing constant $1/2\pi$. And, if we remain consistent with this choice and define the L_2 norm by

$$\|f\|_2 = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)|^2\, dt \right)^{1/2},$$

then we have the simpler estimate $\|f\|_2 \le \|f\|$ for $f \in C^{2\pi}$.

The Dirichlet and Fejer kernels are essentially the same in this case, too, except that we would now write $s_n(f)(x) = \sum_{k=-n}^{n} c_k e^{ikx}$. Given this, the Dirichlet and Fejér kernels can be written

$$\begin{aligned}
D_n(x) = \sum_{k=-n}^{n} e^{ikx} &= 1 + \sum_{k=1}^{n} (e^{ikx} + e^{-ikx}) \\
&= 1 + 2\sum_{k=1}^{n} \cos kx \\
&= \frac{\sin\,(n + \frac{1}{2})\,x}{\sin \frac{1}{2}\,x}
\end{aligned}$$

and

$$\begin{aligned}
K_n(x) &= \frac{1}{n} \sum_{m=0}^{n-1} D_m(x) \\
&= \frac{1}{n} \sum_{m=0}^{n-1} \frac{\sin\,(m + \frac{1}{2})\,x}{\sin \frac{1}{2}\,x}
\end{aligned}$$

$$= \quad \frac{\sin^2(nt/2)}{n\sin^2(t/2)}.$$

In other words, each is twice its real coefficient counterpart. Because the choice of normalizing constant ($1/\pi$ versus $1/2\pi$, and sometimes even $1/\sqrt{\pi}$ or $1/\sqrt{2\pi}$) has a (small) effect on these formulas, you may find some variation in other textbooks.

Problems

∗ 1. Define $f(x) = (\pi - x)^2$ for $0 \le x \le 2\pi$, and extend f to a 2π-periodic continuous function on $\mathbb{R}$ in the obvious way. Check that the Fourier series for f is $\pi^2/3 + 4\sum_{n=1}^{\infty} \cos nx/n^2$. Because this series is uniformly convergent, it actually converges to f. In particular, note that setting $x = 0$ yields the familiar formula $\sum_{n=1}^{\infty} 1/n^2 = \pi^2/6$.

2. (a) Given $n \ge 1$ and $\varepsilon > 0$, show that there is a continuous function $f \in C^{2\pi}$ satisfying $\|f\| = 1$ and $\frac{1}{\pi}\int_{-\pi}^{\pi} |f(t) - \operatorname{sgn} D_n(t)|\, dt < \varepsilon/(n+1)$.

   (b) Show that $s_n(f)(0) \ge \lambda_n - \varepsilon$ and conclude that $\|s_n(f)\| \ge \lambda_n - \varepsilon$.

3. (a) $|D_n(x)| < n + \frac{1}{2} = D_n(0)$ for $0 < |x| < \pi$.

   (b) $2D_n(2x) = U_{2n}(\cos x)$ for $0 \le x \le \pi$.

   (c) $T'_{2n+1}(\cos x) = (2n+1) \cdot 2D_n(2x)$ for $0 \le x \le \pi$ and, hence, $|T'_{2n+1}(t)| < (2n+1)^2 = T'_{2n+1}(\pm 1)$ for $|t| < 1$.

4. (a) If $f,\, k \in C^{2\pi}$, prove that $g(x) = \int_{-\pi}^{\pi} f(x+t)\, k(t)\, dt$ is also in $C^{2\pi}$.

   (b) If we only assume that f is 2π-periodic and Riemann integrable on $[-\pi, \pi\,]$ (but still $k \in C^{2\pi}$), is g still continuous?

   (c) If we simply assume that f and k are 2π-periodic and Riemann integrable on $[-\pi, \pi\,]$, is g still continuous?

5. Let (k_n) be a sequence in $C^{2\pi}$ satisfying the hypotheses of Theorem 6.9. If f is Riemann integrable, show that $\frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\, k_n(t)\, dt \to f(x)$ pointwise, as $n \to \infty$, at each point of continuity of f. In particular, conclude that $\sigma_n(f)(x) \to f(x)$ at each point of continuity of f.

∗ 6. Given $f,\, g \in C^{2\pi}$, we define the *convolution* of f and g, written $f * g$, by

$$(f * g)(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\, g(x-t)\, dt.$$

   (a) Show that $f * g = g * f$ and that $f * g \in C^{2\pi}$.

   (b) If one of f or g is a trig polynomial, show that $f * g$ is again a trig polynomial (of the same degree).

   (c) If one of f or g is continuously differentiable, show that $f * g$ is likewise continuously differentiable and find an integral formula for $(f * g)'(x)$.

7. Show that the complex Fejér kernel can also be written as $K_n(x) = (1/n)\sum_{k=-n}^{n}\big(n - |k|\,\big)\, e^{ikx}$.

Chapter 7

Jackson's Theorems

Direct Theorems

We continue our investigations of the "middle ground" between algebraic and trigonometric approximation by presenting several results due to the great American mathematician Dunham Jackson (from roughly 1911–1912). The first of these results will give us the best possible estimate of $E_n(f)$ in terms of ω_f and n.

Theorem 7.1. *If $f \in C^{2\pi}$, then $E_n^T(f) \le 6\,\omega_f([-\pi,\pi\,]; \frac{1}{n}\,)$.*

Theorem 7.1 should be viewed as an improvement over Bernstein's Theorem 2.6, which stated that $E_n(f) \le \frac{3}{2}\omega_f(\frac{1}{\sqrt{n}})$ for $f \in C[-1,1\,]$. As we'll see, the proof of Theorem 7.1 not only mimics the proof of Bernstein's result, but also uses some of the ideas we talked about in the last chapter. In particular, the proof we'll give involves integration against an "improved" Dirichlet kernel.

Before we dive into the proof, let's list several immediate and important corollaries:

Corollary 7.2. *Weierstrass's Second Theorem.*

Proof. $\omega_f(\frac{1}{n}) \to 0$ for any $f \in C^{2\pi}$. □

Corollary 7.3. (The Dini-Lipschitz Theorem) *If $\omega_f(\frac{1}{n})\log n \to 0$ as $n \to \infty$, then the Fourier series for f converges uniformly to f.*

Proof. From Lebesgue's theorem (Theorem 6.6),

$$\|f - s_n(f)\| \;\le\; (4+\log n)\,E_n^T(f) \;\le\; 6\,(4+\log n)\,\omega_f\left(\frac{1}{n}\right) \to 0. \qquad \square$$

Theorem 7.4. *If $f \in C[-1,1\,]$, then $E_n(f) \le 6\,\omega_f([-1,1\,]; \frac{1}{n}\,)$.*

Proof. Let $\varphi(\theta) = f(\cos\theta)$. Then, as we've seen,

$$E_n(f) \;=\; E_n^T(\varphi) \;\le\; 6\,\omega_\varphi\left([-\pi,\pi\,]; \frac{1}{n}\right) \;\le 6\,\omega_f\left([-1,1\,]; \frac{1}{n}\right),$$

where the last inequality follows from the fact that

$$|\varphi(\alpha) - \varphi(\beta)| \;=\; |f(\cos\alpha) - f(\cos\beta)| \;\le\; \omega_f(|\cos\alpha - \cos\beta\,|) \;\le\; \omega_f(|\alpha-\beta\,|). \qquad \square$$

Corollary 7.5. *If $f \in \text{lip}_K\alpha$ on $[-1,1]$, then $E_n(f) \le 6Kn^{-\alpha}$.* (Recall that Bernstein's theorem gives only $n^{-\alpha/2}$.)

Corollary 7.6. *If $f \in C[-1,1]$ has a bounded derivative, then $E_n(f) \le \frac{6}{n}\,\|f'\|$.*

Corollary 7.7. *If $f \in C[-1,1]$ has a continuous derivative, then $E_n(f) \le \frac{6}{n}\,E_{n-1}(f')$.*

Proof. Let $p^* \in \mathcal{P}_{n-1}$ be the best uniform approximation to f' and consider $p(x) = \int_{-1}^{x} p^*(t)\,dt \in \mathcal{P}_n$. From the previous Corollary,

$$\begin{aligned} E_n(f) &= E_n(f-p) \qquad \text{(Why?)} \\ &\le \frac{6}{n}\,\|f'-p^*\| \;=\; \frac{6}{n}\,E_{n-1}(f').\square \end{aligned}$$

Iterating this last inequality will give the following result:

Corollary 7.8. *If $f \in C[-1,1]$ is k-times continuously differentiable, then*

$$E_n(f) \;\le\; \frac{6^{k+1}}{n(n-1)\cdots(n-k+1)}\;\omega_k\left(\frac{1}{n-k}\right),$$

where ω_k is the modulus of continuity of $f^{(k)}$.

Well, enough corollaries. It's time we proved Theorem 7.1. Now Jackson's approach was to show that

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\,\cdot c_n\left(\frac{\sin nt}{\sin t}\right)^4 dt \;\rightrightarrows\; f(x),$$

where $J_n(t) = c_n(\sin nt/\sin t)^4$ is the improved kernel we alluded to earlier (it's essentially the square of Fejér's kernel). The approach we'll take, due to Korovkin, proves the existence of a suitable kernel without giving a tidy formula for it. On the other hand, it's relatively easy to outline the idea. The key here is that $J_n(t)$ should be an even, nonnegative, trig polynomial of degree n with $\frac{1}{\pi}\int_{-\pi}^{\pi} J_n(t)\,dt = 1$. In other words,

$$J_n(t) \;=\; \frac{1}{2} + \sum_{k=1}^{n} \rho_{k,n}\cos kt \tag{7.1}$$

(why is the first term 1/2?), where $\rho_{1,n},\dots,\rho_{n,n}$ must be chosen so that $J_n(t) \ge 0$. Assuming we can find such $\rho_{k,n}$, here's what we get:

Lemma 7.9. *If $f \in C^{2\pi}$, then*

$$\left|\,f(x) - \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\,J_n(t)\,dt\right| \;\le\; \omega_f\left(\frac{1}{n}\right)\cdot\left[1 + n\pi\sqrt{\frac{1-\rho_{1,n}}{2}}\,\right]. \tag{7.2}$$

Proof. We already know how the first several lines of the proof will go:

$$\begin{aligned} \left|\,f(x) - \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\,J_n(t)\,dt\right| &= \frac{1}{\pi}\left|\int_{-\pi}^{\pi}\big[f(x)-f(x+t)\big]\,J_n(t)\,dt\right| \\ &\le \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)-f(x+t)|\,J_n(t)\,dt \\ &\le \frac{1}{\pi}\int_{-\pi}^{\pi} \omega_f(\,|t|\,)\,J_n(t)\,dt. \end{aligned} \tag{7.3}$$

Next we borrow a trick from Bernstein. We replace $\omega_f(\,|t|\,)$ by

$$\omega_f(\,|t|\,) \;=\; \omega_f\left(n|t|\cdot\frac{1}{n}\right) \;\le\; \big(1+n|t|\,\big)\,\omega_f\left(\frac{1}{n}\right),$$

and so the integral in formula (7.3) is dominated by

$$\omega_f\left(\frac{1}{n}\right)\cdot\frac{1}{\pi}\int_{-\pi}^{\pi}\big(1+n|t|\,\big)\,J_n(t)\,dt \;=\; \omega_f\left(\frac{1}{n}\right)\cdot\left[1+\frac{n}{\pi}\int_{-\pi}^{\pi}|t|\,J_n(t)\,dt\right].$$

All that remains is to estimate $\int_{-\pi}^{\pi}|t|\,J_n(t)\,dt$, and for this we'll appeal to the Cauchy-Schwarz inequality (again, compare this to the proof of Bernstein's theorem).

$$\begin{aligned}
\frac{1}{\pi}\int_{-\pi}^{\pi}|t|\,J_n(t)\,dt &= \frac{1}{\pi}\int_{-\pi}^{\pi}|t|\,J_n(t)^{1/2}J_n(t)^{1/2}\,dt \\
&\le \left(\frac{1}{\pi}\int_{-\pi}^{\pi}|t|^2\,J_n(t)\,dt\right)^{1/2}\left(\frac{1}{\pi}\int_{-\pi}^{\pi}J_n(t)\,dt\right)^{1/2} \\
&= \left(\frac{1}{\pi}\int_{-\pi}^{\pi}|t|^2\,J_n(t)\,dt\right)^{1/2}.
\end{aligned}$$

But,

$$|t|^2 \;\le\; \left[\pi\sin\left(\frac{t}{2}\right)\right]^2 \;=\; \frac{\pi^2}{2}\,(1-\cos t\,).$$

So,

$$\frac{1}{\pi}\int_{-\pi}^{\pi}|t|\,J_n(t)\,dt \;\le\; \left(\frac{\pi^2}{2}\cdot\frac{1}{\pi}\int_{-\pi}^{\pi}(1-\cos t)\,J_n(t)\,dt\right)^{1/2} \;=\; \pi\sqrt{\frac{1-\rho_{1,n}}{2}}. \qquad \square$$

It remains to prove that we can actually find a suitable choice of scalars $\rho_{1,n},\ldots,\rho_{n,n}$. We already know that we need to choose the $\rho_{k,n}$ so that $J_n(t)$ will be nonnegative, but now it's clear that we also want $\rho_{1,n}$ to be very close to 1. To get us started, let's first see why it's easy to generate nonnegative cosine polynomials. Given *real numbers* $c_0,\ldots,c_n$, note that

$$\begin{aligned}
0\le\left|\sum_{k=0}^{n}c_ke^{ikx}\right|^2 &= \left(\sum_{k=0}^{n}c_ke^{ikx}\right)\left(\sum_{j=0}^{n}c_je^{-ijx}\right) \\
&= \sum_{k,j}c_kc_je^{i(k-j)x} \\
&= \sum_{k=0}^{n}c_k^2+\sum_{k>j}c_kc_j\big(e^{i(k-j)x}+e^{i(j-k)x}\big) \\
&= \sum_{k=0}^{n}c_k^2+2\sum_{k>j}c_kc_j\cos(k-j)x \\
&= \sum_{k=0}^{n}c_k^2+2\sum_{k=0}^{n-1}c_kc_{k+1}\cos x+\cdots \\
&\qquad\cdots+2c_0c_n\cos nx.
\end{aligned}$$

In particular, we need to find $c_0, \ldots, c_n$ with

$$\sum_{k=0}^{n} c_k^2 \;=\; \frac{1}{2} \qquad \text{and} \qquad \rho_{1,n} = 2\sum_{k=0}^{n-1} c_k c_{k+1} \;\approx\; 1.$$

What we'll do is find c_k with $\sum_{k=0}^{n-1} c_k c_{k+1} \approx \sum_{k=0}^{n} c_k^2$, and then normalize. But, in fact, we won't actually *find* anything—we'll simply write down a choice of c_k that happens to work! Consider:

$$\begin{aligned}\sum_{k=0}^{n} \sin\left(\frac{k+1}{n+2}\,\pi\right)\sin\left(\frac{k+2}{n+2}\,\pi\right) \;&=\; \sum_{k=0}^{n} \sin\left(\frac{k+1}{n+2}\,\pi\right)\sin\left(\frac{k}{n+2}\,\pi\right)\\ &=\; \frac{1}{2}\sum_{k=0}^{n}\left[\sin\left(\frac{k}{n+2}\,\pi\right)+\sin\left(\frac{k+2}{n+2}\,\pi\right)\right]\sin\left(\frac{k+1}{n+2}\,\pi\right). \qquad (7.4)\end{aligned}$$

By changing the index of summation, it's easy to see that the first two sums are equal and, hence, each is equal to the average of the two. Next we re-write the last sum in (7.4), using the trig identity $\frac{1}{2}\big(\sin A + \sin B\big) = \cos\left(\frac{A-B}{2}\right)\sin\left(\frac{A+B}{2}\right)$, to get

$$\sum_{k=0}^{n} \sin\left(\frac{k+1}{n+2}\,\pi\right)\sin\left(\frac{k+2}{n+2}\,\pi\right) \;=\; \cos\left(\frac{\pi}{n+2}\right)\sum_{k=0}^{n}\sin^2\left(\frac{k+1}{n+2}\,\pi\right).$$

Because $\cos\left(\frac{\pi}{n+2}\right) \approx 1$ for large n, we've done it! If we define $c_k = c\cdot\sin\left(\frac{k+1}{n+2}\,\pi\right)$, where c is chosen so that $\sum_{k=0}^{n} c_k^2 = 1/2$, and if we define $J_n(x)$ using (7.1), where $\rho_{k,n} = c_k$, then $J_n(x) \ge 0$ and $\rho_{1,n} = \cos\left(\frac{\pi}{n+2}\right)$ (why?). The conclusion of Lemma 7.9; that is, the right-hand side of equation (7.2), can now be revised:

$$\sqrt{\frac{1-\rho_{1,n}}{2}} \;=\; \sqrt{\frac{1-\cos\left(\frac{\pi}{n+2}\right)}{2}} \;=\; \sin\left(\frac{\pi}{2n+4}\right) \;\le\; \frac{\pi}{2n},$$

which allows us to conclude that

$$E_n^T(f) \;\le\; \left(1+\frac{\pi^2}{2}\right)\omega_f\left(\frac{1}{n}\right) \;<\; 6\,\omega_f\left(\frac{1}{n}\right).$$

This proves Theorem 7.1. □

Inverse Theorems

Jackson's theorems are what we might call *direct theorems*. If we know something about f, then we can say something about $E_n(f)$. There is also the notion of an *inverse theorem*, meaning that if we know something about $E_n(f)$, we should be able to say something about f. In other words, we would expect an inverse theorem to be, more or less, the converse of some direct theorem. Now inverse theorems are typically much harder to prove than direct theorems, but in order to have some idea of what such theorems might tell us (and to see some of the techniques used in their proofs), we present one of the easier inverse theorems, due to Bernstein. This result yields a (partial) converse to Corollary 7.5.

Theorem 7.10. *If $f \in C^{2\pi}$ satisfies $E_n^T(f) \le A\, n^{-\alpha}$, for some constants A and $0 < \alpha < 1$, then $f \in \text{lip}_K \alpha$ for some constant K.*

Proof. For each n, choose $S_n \in \mathcal{T}_n$ so that $\|f - S_n\| \le A\, n^{-\alpha}$. Then, in particular, (S_n) converges uniformly to f. Now if we set $V_0 = S_1$ and $V_n = S_{2^n} - S_{2^{n-1}}$ for $n \ge 1$, then $V_n \in \mathcal{T}_{2^n}$ and $f = \sum_{n=0}^\infty V_n$. Indeed,

$$\|V_n\| \;\le\; \|S_{2^n} - f\| + \|S_{2^{n-1}} - f\| \;\le\; A\,(2^n)^{-\alpha} + A\,(2^{n-1})^{-\alpha} \;=\; B \cdot 2^{-n\alpha},$$

which is summable; thus, the (telescoping) series $\sum_{n=0}^\infty V_n$ converges uniformly to f. (Why?)

Next we estimate $|f(x) - f(y)|$ using finitely many of the V_n, the precise number to be specified later. Using the mean value theorem and Bernstein's inequality we get

$$\begin{aligned}
|f(x) - f(y)| &\le \sum_{n=0}^{\infty} |V_n(x) - V_n(y)| \\
&\le \sum_{n=0}^{m-1} |V_n(x) - V_n(y)| \;+\; 2\sum_{n=m}^{\infty} \|V_n\| \\
&= \sum_{n=0}^{m-1} |V_n'(\xi_n)|\,|x-y| \;+\; 2\sum_{n=m}^{\infty} \|V_n\| \\
&\le |x-y| \sum_{n=0}^{m-1} 2^n \|V_n\| \;+\; 2\sum_{n=m}^{\infty} \|V_n\| && (7.5) \\
&\le |x-y| \sum_{n=0}^{m-1} B\,2^{n(1-\alpha)} \;+\; 2\sum_{n=m}^{\infty} B\,2^{-n\alpha} \\
&\le C\left[\,|x-y| \cdot 2^{m(1-\alpha)} \;+\; 2^{-m\alpha}\right], && (7.6)
\end{aligned}$$

where we've used, in (7.5), the fact that $V_n \in \mathcal{T}_{2^n}$ and, in (7.6), standard estimates for geometric series. Now we want the right-hand side to be dominated by a constant times $|x-y|^\alpha$. In other words, if we set $|x - y| = \delta$, then we want

$$\delta \cdot 2^{m(1-\alpha)} \;+\; 2^{-m\alpha} \;\le\; D \cdot \delta^\alpha$$

or, equivalently,

$$(2^m\delta)^{(1-\alpha)} \;+\; (2^m\delta)^{-\alpha} \;\le\; D. \tag{7.7}$$

Thus, we should choose m so that $2^m\delta$ is both bounded above and bounded away from zero. For example, if $0 < \delta < 1$, we could choose m so that $1 \le 2^m\delta < 2$. $\square$

In order to better explain the phrase "more or less the converse of some direct theorem," let's see how the previous result falls apart when $\alpha = 1$. Although we might hope that $E_n^T(f) \le A/n$ would imply that $f \in \text{lip}_K 1$, it happens not to be true. The best result in this regard is due to Zygmund, who gave necessary and sufficient conditions on f so that $E_n^T(f) \le A/n$ (and these conditions do not characterize $\text{lip}_K 1$ functions). Instead of pursuing Zygmund's result, we'll settle for simple "surgery" on our previous result, keeping an eye out for what goes wrong. This result is again due to Bernstein.

Theorem 7.11. *If $f \in C^{2\pi}$ satisfies $E_n^T(f) \le A/n$, then $\omega_f(\delta) \le K\delta|\log\delta\,|$ for some constant K and all δ sufficiently small.*

Proof. If we repeat the previous proof, setting $\alpha = 1$, only a few lines change. In particular, the conclusion of that long string of inequalities, ending with (7.6), would now read

$$|f(x) - f(y)| \;\le\; C\left[\,|x-y|\cdot m \;+\; 2^{-m}\,\right] \;=\; C\left[\,m\delta \;+\;\; 2^{-m}\,\right].$$

Clearly, the right-hand side cannot be dominated by a constant times δ, as we might have hoped, for this would force m to be bounded (independent of δ), which in turn bounds δ away from zero. But, if we again think of $2^m\delta$ as the "variable" in this inequality (as suggested by formula (7.7) and the concluding lines of the previous proof), then the term $m\delta$ suggests that the correct order of magnitude of the right-hand side is $\delta|\log\delta\,|$. Thus, we would try to find a constant D so that

$$m\delta \;+\;\; 2^{-m} \;\le\; D\cdot\delta|\log\delta\,|$$

or

$$m(2^m\delta) \;+\; 1 \;\le\; D\cdot(2^m\delta)|\log\delta\,|.$$

Now if we take $0 < \delta < 1/2$, then $\log 2 < -\log\delta = |\log\delta\,|$. Hence, if we again choose $m \ge 1$ so that $1 \le 2^m\delta < 2$, we'll get

$$m\log 2 + \log\delta < \log 2 \quad\Longrightarrow\quad m < \frac{\log 2 - \log\delta}{\log 2} < \frac{2}{\log 2}\,|\log\delta\,|$$

and, finally,

$$m(2^m\delta) + 1 \;\le\; 2m + 1 \;\le\; 3m \;\le\; \frac{6}{\log 2}\,|\log\delta\,| \;\le\; \frac{6}{\log 2}\,(2^m\delta)\,|\log\delta\,|. \qquad \square$$

Chapter 8

Orthogonal Polynomials

Given a positive (except possibly at finitely many points), Riemann integrable *weight function* $w(x)$ on $[\,a,b\,]$, the expression

$$\langle f,g\rangle \;=\; \int_a^b f(x)\,g(x)\,w(x)\,dx$$

defines an inner product on $C[\,a,b\,]$ and

$$\|f\|_2 \;=\; \left(\int_a^b f(x)^2\,w(x)\,dx\right)^{1/2} \;=\; \sqrt{\langle f,f\rangle}$$

defines a *strictly convex* norm on $C[\,a,b\,]$. (See Problem 1 at the end of the chapter.) Thus, given a finite dimensional subspace E of $C[\,a,b\,]$ and an element $f \in C[\,a,b\,]$, there is a unique $g \in E$ such that

$$\|f-g\|_2 \;=\; \min_{h\in E}\|f-h\|_2.$$

We say that g is the *least-squares approximation* to f out of E (relative to w).

Now if we apply the Gram-Schmidt procedure to the sequence $1, x, x^2, \ldots$, we will arrive at a sequence (Q_n) of *orthogonal polynomials* relative to the above inner product. In this special case, however, the Gram-Schmidt procedure simplifies substantially:

Theorem 8.1. *The following procedure defines a sequence (Q_n) of orthogonal polynomials (relative to w). Set:*

$$Q_0(x) = 1, \quad Q_1(x) = x - a_0 = (x-a_0)Q_0(x),$$

and

$$Q_{n+1}(x) = (x-a_n)Q_n(x) - b_nQ_{n-1}(x),$$

for $n \ge 1$, where

$$a_n = \langle\, x\,Q_n, Q_n\,\rangle \big/ \langle\, Q_n, Q_n\,\rangle \quad \textit{and} \quad b_n = \langle\, x\,Q_n, Q_{n-1}\,\rangle \big/ \langle\, Q_{n-1}, Q_{n-1}\,\rangle$$

(and where $x\,Q_n$ is shorthand for the polynomial $x\,Q_n(x)$).

Proof. It's easy to see from these formulas that Q_n is a monic polynomial of degree exactly n. In particular, the Q_n are linearly independent (and nonzero).

Now it's easy to see that Q_0, Q_1, and Q_2 are mutually orthogonal, so let's use induction and check that Q_{n+1} is orthogonal to each Q_k, $k \le n$. First,

$$\langle\, Q_{n+1}, Q_n \,\rangle = \langle\, x\, Q_n, Q_n \,\rangle - a_n \langle\, Q_n, Q_n \,\rangle - b_n \langle\, Q_{n-1}, Q_n \,\rangle = 0$$

and

$$\langle\, Q_{n+1}, Q_{n-1} \,\rangle = \langle\, x\, Q_n, Q_{n-1} \,\rangle - a_n \langle\, Q_n, Q_{n-1} \,\rangle - b_n \langle\, Q_{n-1}, Q_{n-1} \,\rangle = 0,$$

because $\langle\, Q_{n-1}, Q_n \,\rangle = 0$. Next, we take $k < n-1$ and use the recurrence formula twice:

$$\begin{aligned}
\langle\, Q_{n+1}, Q_k \,\rangle &= \langle\, x\, Q_n, Q_k \,\rangle - a_n \langle\, Q_n, Q_k \,\rangle - b_n \langle\, Q_{n-1}, Q_k \,\rangle \\
&= \langle\, x\, Q_n, Q_k \,\rangle = \langle\, Q_n, x\, Q_k \,\rangle \qquad \text{(Why?)} \\
&= \langle\, Q_n, Q_{k+1} + a_k Q_k + b_k Q_{k-1} \,\rangle = 0,
\end{aligned}$$

because $k+1 < n$. □

Remarks 8.2.

1. Using the same trick as above, we have

$$b_n = \langle\, x\, Q_n, Q_{n-1} \,\rangle \big/ \langle\, Q_{n-1}, Q_{n-1} \,\rangle = \langle\, Q_n, Q_n \,\rangle \big/ \langle\, Q_{n-1}, Q_{n-1} \,\rangle > 0.$$

2. Each $p \in \mathcal{P}_n$ can be uniquely written $p = \sum_{i=0}^{n} \alpha_i Q_i$, where $\alpha_i = \langle\, p, Q_i \,\rangle \big/ \langle\, Q_i, Q_i \,\rangle$.

3. If Q is any monic polynomial of degree exactly n, then $Q = Q_n + \sum_{i=0}^{n-1} \alpha_i Q_i$ (why?) and hence

$$\|Q\|_2^2 \;=\; \|Q_n\|_2^2 \;+\; \sum_{i=0}^{n-1} \alpha_i^2 \|Q_i\|_2^2 \;>\; \|Q_n\|_2^2,$$

unless $Q = Q_n$. That is, Q_n has the least $\|\cdot\|_2$ norm of all monic polynomials of degree n.

4. The Q_n are *unique* in the following sense: If (P_n) is another sequence of orthogonal polynomials such that P_n has degree exactly n, then $P_n = \alpha_n Q_n$ for some $\alpha_n \ne 0$. (See Problem 4.) Consequently, there's no harm in referring to the Q_n as *the* sequence of orthogonal polynomials relative to w.

5. For $n \ge 1$ note that $\int_a^b Q_n(t)\, w(t)\, dt = \langle Q_0, Q_n \rangle = 0$.

Examples 8.3.

1. On $[-1, 1]$, the Chebyshev polynomials of the first kind (T_n) are orthogonal relative to the weight $w(x) = 1/\sqrt{1-x^2}$.

$$\begin{aligned}
\int_{-1}^{1} T_m(x)\, T_n(x)\, \frac{dx}{\sqrt{1-x^2}} &= \int_0^{\pi} \cos m\theta\, \cos n\theta\, d\theta \\
&= \begin{cases} 0, & m \ne n \\ \pi, & m = n = 0 \\ \pi/2, & m = n \ne 0. \end{cases}
\end{aligned}$$

Because T_n has degree exactly n, this must be the right choice. Notice, too, that $\frac{1}{\sqrt{2}}\,T_0, T_1, T_2, \ldots$ are *orthonormal* relative to the weight $2/\pi\sqrt{1-x^2}$.

In terms of the inductive procedure given above, we must have $Q_0 = T_0 = 1$ and $Q_n = 2^{-n+1}T_n$ for $n \geq 1$. (Why?) From this it follows that $a_n = 0$, $b_1 = 1/2$, and $b_n = 1/4$ for $n \geq 2$. (Why?) That is, the recurrence formula given in Theorem 8.1 reduces to the familar relationship $T_{n+1}(x) = 2x\,T_n(x) - T_{n-1}(x)$. Curiously, $Q_n = 2^{-n+1}T_n$ minimizes both

$$\max_{-1\leq x\leq 1} |p(x)| \qquad \text{and} \qquad \left(\int_{-1}^{1} p(x)^2 \,\frac{dx}{\sqrt{1-x^2}}\right)^{1/2}$$

over all monic polynomials of degree exactly n.

The Chebyshev polynomials also satisfy $(1-x^2)\,T_n''(x) - x\,T_n'(x) + n^2\,T_n(x) = 0$. Because this is a polynomial identity, it suffices to check it for all $x = \cos\theta$. In this case,

$$T_n'(x) \;=\; \frac{n\sin n\theta}{\sin\theta}$$

and

$$T_n''(x) \;=\; \frac{n^2\cos n\theta\sin\theta - n\sin n\theta\cos\theta}{\sin^2\theta\,(-\sin\theta)}.$$

Hence,

$$\begin{aligned}&(1-x^2)\,T_n''(x) - x\,T_n'(x) + n^2\,T_n(x)\\ &\qquad = \; -n^2\cos n\theta + n\sin n\theta\cot\theta - n\sin n\theta\cot\theta + n^2\cos\theta = 0\end{aligned}$$

2. On $[-1,1\,]$, the Chebyshev polynomials of the second kind (U_n) are orthogonal relative to the weight $w(x) = \sqrt{1-x^2}$. Indeed,

$$\begin{aligned}&\int_{-1}^{1} U_m(x)\,U_n(x)\,(1-x^2)\,\frac{dx}{\sqrt{1-x^2}}\\ &\qquad = \int_0^{\pi} \frac{\sin{(m+1)\theta}}{\sin\theta}\cdot\frac{\sin{(n+1)\theta}}{\sin\theta}\cdot\sin^2\theta\,d\theta = \begin{cases} 0, & m\neq n\\ \pi/2, & m=n.\end{cases}\end{aligned}$$

While we're at it, notice that

$$T_n'(x) \;=\; \frac{n\sin n\theta}{\sin\theta} \;=\; n\,U_{n-1}(x).$$

As a rule, the derivatives of a sequence of orthogonal polynomials are again orthogonal polynomials, but relative to a different weight.

3. On $[-1,1\,]$ with weight $w(x) \equiv 1$, the sequence (P_n) of *Legendre polynomials* are orthogonal, and are typically normalized by $P_n(1) = 1$. The first few Legendre polynomials are $P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = \frac{3}{2}\,x^2 - \frac{1}{2}$, and $P_3(x) = \frac{5}{2}\,x^3 - \frac{3}{2}\,x$. (Check this!) After we've seen a few more examples, we'll come back and give an explicit formula for P_n.

4. All of the examples we've seen so far are special cases of the following: On $[-1,1]$, consider the weight $w(x) = (1-x)^\alpha(1+x)^\beta$, where $\alpha, \beta > -1$. The corresponding orthogonal polynomials $(P_n^{(\alpha,\beta)})$ are called the *Jacobi polynomials* and are typically normalized by requiring that

$$P_n^{(\alpha,\beta)}(1) \;=\; \binom{n+\alpha}{\alpha} \;\equiv\; \frac{(\alpha+1)(\alpha+2)\cdots(\alpha+n)}{n!}.$$

   It follows that $P_n^{(0,0)} = P_n$,

$$P_n^{(-1/2,-1/2)} \;=\; \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2^n n!}\, T_n,$$

   and

$$P_n^{(1/2,1/2)} \;=\; \frac{1\cdot 3\cdot 5\cdots(2n+1)}{2^n (n+1)!}\, U_n.$$

   The polynomials $P_n^{(\alpha,\alpha)}$ are called *ultraspherical* polynomials.

5. There are also several classical examples of orthogonal polynomials on unbounded intervals. In particular,

$(0,\infty)$	$w(x) = e^{-x}$	Laguerre polynomials,
$(0,\infty)$	$w(x) = x^\alpha e^{-x}$	generalized Laguerre polynomials,
$(-\infty,\infty)$	$w(x) = e^{-x^2}$	Hermite polynomials.

Because Q_n is orthogonal to every element of $\mathcal{P}_{n-1}$, a fuller understanding of Q_n will follow from a characterization of the *orthogonal complement* of $\mathcal{P}_{n-1}$. We begin with an easy fact about least-squares approximations in inner product spaces.

Lemma 8.4. *Let E be a finite dimensional subspace of an inner product space X, and let $x \in X \setminus E$. Then, $y^* \in E$ is the least-squares approximation to x out of E (a.k.a. the nearest point to x in E) if and only if $\langle x - y^*,\, y\rangle = 0$ for every $y \in E$; that is, if and only if $(x-y^*) \perp E$.*

Proof. [We've taken E to be finite dimensional so that nearest points will exist; because X is an inner product space, nearest points must also be unique (see Problem 1 for a proof that every inner product norm is strictly convex).]

($\Longleftarrow$) First suppose that $(x-y^*) \perp E$. Then, given any $y \in E$, we have

$$\|x-y\|_2^2 \;=\; \|(x-y^*)+(y^*-y)\|_2^2 \;=\; \|x-y^*\|_2^2 + \|y^*-y\|_2^2,$$

because $y^*-y \in E$ and, hence, $(x-y^*) \perp (y^*-y)$. Thus, $\|x-y\| > \|x-y^*\|$ unless $y = y^*$; that is, y^* is the (unique) nearest point to x in E.

($\Longrightarrow$) Suppose that $x - y^*$ is *not* orthogonal to E. Then there is some $y \in E$ with $\|y\| = 1$ such that $\alpha = \langle x-y^*,\, y\rangle \neq 0$. It now follows that $y^* + \alpha y \in E$ is a better approximation to x than y^* (and $y^* + \alpha y \neq y^*$, of course); that is, y^* is *not* the least-squares approximation to x. To see this, we again compute:

$$\begin{aligned}
\|x-(y^*+\alpha y)\|_2^2 &= \|(x-y^*)-\alpha y\|_2^2 \;=\; \langle (x-y^*)-\alpha y,\, (x-y^*)-\alpha y\rangle \\
&= \|x-y^*\|_2^2 - 2\alpha\,\langle x-y^*,\, y\rangle + \alpha^2 \\
&= \|x-y^*\|_2^2 - \alpha^2 \;<\; \|x-y^*\|_2^2.
\end{aligned}$$

Thus, we must have $\langle x-y^*,\, y\rangle = 0$ for every $y \in E$. □

Lemma 8.5. (Integration by-parts)

$$\int_a^b u^{(n)}v \;=\; \sum_{k=1}^{n}(-1)^{k-1}u^{(n-k)}v^{(k-1)}\Big]_a^b \;+\; (-1)^n\int_a^b uv^{(n)}.$$

Now if v is a polynomial of degree $< n$, then $v^{(n)} = 0$ and we get:

Lemma 8.6. *$f \in C[\,a,b\,]$ satisfies $\displaystyle\int_a^b f(x)\,p(x)\,w(x)\,dx = 0$ for all polynomials $p \in \mathcal{P}_{n-1}$ if and only if there is an n-times differentiable function u on $[\,a,b\,]$ satisfying $fw = u^{(n)}$ and $u^{(k)}(a) = u^{(k)}(b) = 0$ for all $k = 0, 1, \ldots, n-1$.*

Proof. One direction is clear from Lemma 8.5: Given u as above, we would have $\int_a^b fpw = \int_a^b u^{(n)}p = (-1)^n\int_a^b up^{(n)} = 0$.

So, suppose we have that $\int_a^b fpw = 0$ for all $p \in \mathcal{P}_{n-1}$. By integrating fw repeatedly, choosing constants appropriately, we may define a function u satisfying $fw = u^{(n)}$ and $u^{(k)}(a) = 0$ for all $k = 0, 1, \ldots, n-1$. We want to show that the hypotheses on f force $u^{(k)}(b) = 0$ for all $k = 0, 1, \ldots, n-1$.

Now Lemma 8.5 tells us that

$$0 \;=\; \int_a^b fpw \;=\; \sum_{k=1}^{n}(-1)^{k-1}u^{(n-k)}(b)\,p^{(k-1)}(b)$$

for all $p \in \mathcal{P}_{n-1}$. But the numbers $p(b), p'(b), \ldots, p^{(n-1)}(b)$ are completely arbitrary; that is (again by integrating repeatedly, choosing our constants as we please), we can find polynomials p_k of degree $k < n$ such that $p_k^{(k)}(b) \neq 0$ and $p_k^{(j)}(b) = 0$ for $j \neq k$. In fact, $p_k(x) = (x-b)^k$ works just fine! In any case, we must have $u^{(k)}(b) = 0$ for all $k = 0, 1, \ldots, n-1$. $\square$

Rolle's theorem tells us a bit more about the functions orthogonal to $\mathcal{P}_{n-1}$:

Lemma 8.7. *If $w(x) > 0$ in (a,b), and if $f \in C[\,a,b\,]$ is in the orthogonal complement of $\mathcal{P}_{n-1}$ (relative to w); that is, if f satisfies $\int_a^b f(x)\,p(x)\,w(x)\,dx = 0$ for all polynomials $p \in \mathcal{P}_{n-1}$, then f has at least n distinct zeros in the open interval (a,b).*

Proof. Write $fw = u^{(n)}$, where $u^{(k)}(a) = u^{(k)}(b) = 0$ for all $k = 0, 1, \ldots, n-1$. In particular, because $u(a) = u(b) = 0$, Rolle's theorem tells us that u' would have at least one zero in (a,b). But then $u'(a) = u'(c) = u'(b) = 0$, and so u'' must have at least two zeros in (a,b). Continuing, we find that $fw = u^{(n)}$ must have at least n zeros in (a,b). Because $w > 0$, the result follows. $\square$

Corollary 8.8. *Let (Q_n) be the sequence of orthogonal polynomials associated to a given weight w with $w > 0$ in (a,b). Then, the roots of Q_n are real, simple, and lie in (a,b).*

The sheer volume of literature on orthogonal polynomials and other "special functions" is truly staggering. We'll content ourselves with the Legendre and the Chebyshev polynomials. In particular, let's return to the problem of finding an explicit formula for the Legendre polynomials. We could, as Rivlin does, use induction and a few observations that simplify the basic recurrence formula (you're encouraged to read this; see [45, pp. 53–54]). Instead we'll give a simple (but at first sight intimidating) formula that is of use in more general settings than ours.

Lemma 8.6 (with $w \equiv 1$ and $[\,a, b\,] = [-1, 1\,]$) says that if we want to find a *polynomial* f of degree n which is orthogonal to $\mathcal{P}_{n-1}$, then we'll need to take a *polynomial* for u, and this u will have to be divisible by $(x-1)^n(x+1)^n$. (Why?) That is, we must have $P_n(x) = c_n \cdot D^n\big[(x^2-1)^n\big]$, where D denotes differentiation, and where c_n is chosen so that $P_n(1) = 1$.

Lemma 8.9. (Leibniz's formula) $D^n(fg) \;=\; \displaystyle\sum_{k=0}^{n} \binom{n}{k} D^k(f)\, D^{n-k}(g).$

Proof. Induction and the fact that $\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}$. □

Consequently, $Q(x) = D^n\big[(x-1)^n(x+1)^n\big] = \sum_{k=0}^{n}\binom{n}{k} D^k(x-1)^n\, D^{n-k}(x+1)^n$ and it follows that $Q(1) = 2^n n!$ and $Q(-1) = (-1)^n 2^n n!$. This, finally, gives us the formula discovered by Rodrigues in 1814:

$$P_n(x) \;=\; \frac{1}{2^n n!}\, D^n\big[(x^2-1)^n\big]. \tag{8.1}$$

The Rodrigues formula is quite useful (and easily generalizes to the Jacobi polynomials).

Remarks 8.10.

1. By Corollary 8.8, the roots of P_n are real, distinct, and lie in $(-1, 1)$.

2. From the binomial theorem, $(x^2-1)^n = \sum_{k=0}^{n}(-1)^k\binom{n}{k}x^{2n-2k}$. If we apply $\frac{1}{2^n n!}\,D^n$ and simplify, we get another formula for the Legendre polynomials.

$$P_n(x) \;=\; \frac{1}{2^n}\sum_{k=0}^{[n/2]}(-1)^k\binom{n}{k}\binom{2n-2k}{n}\,x^{n-2k}.$$

   In particular, if n is even (odd), then P_n is even (odd). Notice, too, that if we let $\widetilde{P}_n$ denote the polynomial given by the standard construction, then we must have $P_n = 2^{-n}\binom{2n}{n}\widetilde{P}_n$.

3. In terms of our standard recurrence formula, it follows that $a_n = 0$ (because $xP_n(x)^2$ is always odd). It remains to compute b_n. First, integrating by parts,

$$\int_{-1}^{1} P_n(x)^2\,dx \;=\; xP_n(x)^2\Big]_{-1}^{1} \;-\; \int_{-1}^{1} x\cdot 2P_n(x)\,P_n'(x)\,dx,$$

   or $\langle\, P_n, P_n\,\rangle \;=\; 2 - 2\langle\, P_n, xP_n'\,\rangle$. But $xP_n' \;=\; nP_n$ + lower degree terms; hence, $\langle\, P_n, xP_n'\,\rangle \;=\; n\langle\, P_n, P_n\,\rangle$. Thus, $\langle\, P_n, P_n\,\rangle \;=\; 2/(2n+1)$. Using this and the fact that $P_n = 2^{-n}\binom{2n}{n}\widetilde{P}_n$, we'd find that $b_n = n^2/(4n^2-1)$. Thus,

$$\begin{aligned}
P_{n+1} &= 2^{-n-1}\binom{2n+2}{n+1}\widetilde{P}_{n+1}\\
&= 2^{-n-1}\binom{2n+2}{n+1}\left[x\,\widetilde{P}_n \;-\; \frac{n^2}{(4n^2-1)}\,\widetilde{P}_{n-1}\right]\\
&= \frac{2n+1}{n+1}\,x\,P_n \;-\; \frac{n}{n+1}\,P_{n-1}.
\end{aligned}$$

   That is, the Legendre polynomials satisfy the recurrence formula

$$(n+1)\,P_{n+1}(x) \;=\; (2n+1)\,x\,P_n(x) - n\,P_{n-1}(x).$$

4. It follows from the calculations in remark 3, above, that the sequence $\widehat{P}_n = \sqrt{\frac{2n+1}{2}}\, P_n$ is *orthonormal* on $[-1, 1\,]$.

5. The Legendre polynomials satisfy the differential equation $(1-x^2)\, P_n''(x) - 2x\, P_n'(x) + n\,(n+1)\, P_n(x) = 0$. If we set $u = (x^2-1)^n$; that is, if $u^{(n)} = 2^n n! P_n$, note that $(x^2-1)u' = 2nxu$. Now we apply D^{n+1} to both sides of this last equation (using Leibniz's formula) and simplify:
$$\begin{aligned} u^{(n+2)}(x^2-1) + (n+1)\, u^{(n+1)}\, 2x &+ \tfrac{(n+1)n}{2}\, u^{(n)}\, 2 \\ &= 2n\left[u^{(n+1)}\, x + (n+1)\, u^{(n)}\right] \\ \implies\ (1-x^2)\, u^{(n+2)} - 2x\, u^{(n+1)} &+ n\,(n+1)\, u^{(n)} = 0. \end{aligned}$$

6. Through a series of exercises, similar in spirit to remark 5, Rivlin shows that $|P_n(x)| \le 1$ on $[-1, 1\,]$. See [45, pp. 63–64] for details.

Given an orthogonal sequence, it makes sense to consider *generalized Fourier series* relative to the sequence and to find analogues of the Dirichlet kernel, Lebesgue's theorem, and so on. In case of the Legendre polynomials we have the following:

Example 8.11. The Fourier-Legendre series for $f \in C[-1, 1\,]$ is given by $\sum_k \langle\, f, \widehat{P}_k\,\rangle\, \widehat{P}_k$, where
$$\widehat{P}_k = \sqrt{\frac{2k+1}{2}}\, P_k \qquad \text{and} \qquad \langle\, f, \widehat{P}_k\,\rangle = \int_{-1}^{1} f(x)\, \widehat{P}_k(x)\, dx.$$
The partial sum operator $S_n(f) = \sum_{k=0}^{n} \langle\, f, \widehat{P}_k\,\rangle\, \widehat{P}_k$ is a linear projection onto $\mathcal{P}_n$ and may be written as
$$S_n(f)(x) = \int_{-1}^{1} f(t)\, K_n(t,x)\, dt,$$
where $K_n(t,x) = \sum_{k=0}^{n} \widehat{P}_k(t)\, \widehat{P}_k(x)$. (Why?)

Because the polynomials $\widehat{P}_k$ are orthonormal, we have
$$\sum_{k=0}^{n} |\langle\, f, \widehat{P}_k\,\rangle|^2 \ = \ \|S_n(f)\|_2^2 \ \le \ \|f\|_2^2 \ = \ \sum_{k=0}^{\infty} |\langle\, f, \widehat{P}_k\,\rangle|^2,$$
and so the generalized Fourier coefficients $\langle\, f, \widehat{P}_k\,\rangle$ are square summable; in particular, $\langle\, f, \widehat{P}_k\,\rangle \to 0$ as $k \to \infty$. As in the case of Fourier series, the fact that the polynomials (i.e., the span of the $\widehat{P}_k$) are dense in $C[\,a, b\,]$ implies that $S_n(f)$ actually converges to f in the $\|\cdot\|_2$ norm. These same observations remain valid for any sequence of orthogonal polynomials. The real question remains, just as with Fourier series, whether $S_n(f)$ is a good uniform (or even pointwise) approximation to f.

If you're willing to swallow the fact that $|P_n(x)| \le 1$, we get
$$|K_n(t,x)| \ \le \ \sum_{k=0}^{n} \sqrt{\frac{2k+1}{2}}\sqrt{\frac{2k+1}{2}} \ = \ \frac{1}{2}\sum_{k=0}^{n}(2k+1) \ = \ \frac{(n+1)^2}{2}.$$
Hence, $\|S_n(f)\| \le (n+1)^2\|f\|$. That is, the Lebesgue numbers for this process are at most $(n+1)^2$. The analogue of Lebesgue's theorem in this case would then read:
$$\|f - S_n(f)\| \ \le \ Cn^2 E_n(f).$$

Thus, $S_n(f) \rightrightarrows f$ whenever $n^2 E_n(f) \to 0$, and Jackson's theorem tells us when this will happen: *If f is twice continuously differentiable, then the Fourier-Legendre series for f converges uniformly to f on $[-1,1]$.*

The Christoffel-Darboux Identity

It would also be of interest to have a closed form for $K_n(t,x)$. That this is indeed always possible, for any sequence of orthogonal polynomials, is a very important fact.

Using our original notation, let (Q_n) be the sequence of monic orthogonal polynomials corresponding to a given weight w, and let $(\widehat{Q}_n)$ be the *orthonormal* counterpart of (Q_n); in other words, $Q_n = \lambda_n \widehat{Q}_n$, where $\lambda_n = \sqrt{\langle Q_n, Q_n \rangle}$. It will help things here if you recall (from Remarks 8.2 (1)) that $\lambda_n^2 = b_n \lambda_{n-1}^2$.

As with the Legendre polynomials, each $f \in C[a,b]$ is represented by a generalized Fourier series $\sum_k \langle f, \widehat{Q}_k \rangle \widehat{Q}_k$ with partial sum operator

$$S_n(f)(x) = \int_a^b f(t)\, K_n(t,x)\, w(t)\, dt,$$

where $K_n(t,x) = \sum_{k=0}^n \widehat{Q}_k(t)\, \widehat{Q}_k(x)$. As before, S_n is a projection onto $\mathcal{P}_n$; in particular, $S_n(1) = 1$ for every n.

Theorem 8.12. (Christoffel-Darboux) *The kernel $K_n(t,x)$ can be written*

$$\sum_{k=0}^n \widehat{Q}_k(t)\, \widehat{Q}_k(x) \;=\; \lambda_{n+1}\lambda_n^{-1}\, \frac{\widehat{Q}_{n+1}(t)\, \widehat{Q}_n(x) - \widehat{Q}_n(t)\, \widehat{Q}_{n+1}(x)}{t-x}.$$

Proof. We begin with the standard recurrence formulas

$$Q_{n+1}(t) = (t-a_n)\, Q_n(t) - b_n Q_{n-1}(t)$$
$$Q_{n+1}(x) = (x-a_n)\, Q_n(x) - b_n Q_{n-1}(x)$$

(where $b_0 = 0$). Multiplying the first by $Q_n(x)$, the second by $Q_n(t)$, and subtracting:

$$\begin{aligned} &Q_{n+1}(t)\, Q_n(x) - Q_n(t)\, Q_{n+1}(x) \\ &\qquad = \; (t-x)\, Q_n(t)\, Q_n(x) \; + \; b_n \big[\, Q_n(t)\, Q_{n-1}(x) \; - \; Q_n(x)\, Q_{n-1}(t) \,\big] \end{aligned}$$

(and again, $b_0 = 0$). If we divide both sides of this equation by λ_n^2 we get

$$\begin{aligned} &\lambda_n^{-2} \big[\, Q_{n+1}(t)\, Q_n(x) - Q_n(t)\, Q_{n+1}(x) \,\big] \\ &\qquad = \; (t-x)\, \widehat{Q}_n(t)\, \widehat{Q}_n(x) \; + \; \lambda_{n-1}^{-2} \big[\, Q_n(t)\, Q_{n-1}(x) \; - \; Q_n(x)\, Q_{n-1}(t) \,\big]. \end{aligned}$$

Thus, we may repeat the process; arriving finally at

$$\lambda_n^{-2} \big[\, Q_{n+1}(t)\, Q_n(x) - Q_n(t)\, Q_{n+1}(x) \,\big] \;=\; (t-x) \sum_{k=0}^n \widehat{Q}_k(t)\, \widehat{Q}_k(x).$$

The Christoffel-Darboux identity now follows by writing $Q_n = \lambda_n \widehat{Q}_n$, etc. □

And we now have a version of the Dini-Lipschitz theorem (Theorem 7.3).

Theorem 8.13. *Let $f \in C[\,a,b\,]$ and suppose that at some point x_0 in $[\,a,b\,]$ we have*

(i) *f is Lipschitz at x_0; that is, $|f(x_0) - f(x)| \le K|x_0 - x|$ for some constant K and all x in $[\,a,b\,]$; and*

(ii) *the sequence $\big(\widehat{Q}_n(x_0)\big)$ is bounded.*

Then, the series $\sum_k \langle\, f, \widehat{Q}_k \,\rangle\, \widehat{Q}_k(x_0)$ converges to $f(x_0)$.

Proof. First note that the sequence $\lambda_{n+1}\lambda_n^{-1}$ is bounded: Indeed, by Cauchy-Schwarz,

$$\begin{aligned} \lambda_{n+1}^2 &= \langle\, Q_{n+1}, Q_{n+1} \,\rangle \;=\; \langle\, Q_{n+1}, x\, Q_n \,\rangle \\ &\le \|Q_{n+1}\|_2 \cdot \|\, x \,\| \cdot \|Q_n\|_2 \;=\; \max\{|a|, |b|\}\, \lambda_{n+1}\lambda_n. \end{aligned}$$

Thus, $\lambda_{n+1}\lambda_n^{-1} \le c = \max\{|a|, |b|\}$. Now, using the Christoffel-Darboux identity,

$$\begin{aligned} S_n(f)(x_0) - f(x_0) &= \textstyle\int_a^b \big[\, f(t) - f(x_0) \,\big]\, K_n(t, x_0)\, w(t)\, dt \\ &= \lambda_{n+1}\lambda_n^{-1} \int_a^b \frac{f(t) - f(x_0)}{t - x_0} \big[\, \widehat{Q}_{n+1}(t)\, \widehat{Q}_n(x_0) - \widehat{Q}_n(t)\, \widehat{Q}_{n+1}(x_0) \,\big]\, w(t)\, dt \\ &= \lambda_{n+1}\lambda_n^{-1} \big[\, \langle\, h, \widehat{Q}_{n+1} \,\rangle\, \widehat{Q}_n(x_0) - \langle\, h, \widehat{Q}_n \,\rangle\, \widehat{Q}_{n+1}(x_0) \,\big], \end{aligned}$$

where $h(t) = (f(t) - f(x_0))/(t - x_0)$. But h is bounded (and continuous everywhere except, possibly, at x_0) by hypothesis (i), $\lambda_{n+1}\lambda_n^{-1}$ is bounded, and $\widehat{Q}_n(x_0)$ is bounded by hypothesis (ii). All that remains is to notice that the numbers $\langle\, h, \widehat{Q}_n \,\rangle$ are the generalized Fourier coefficients of the bounded, Riemann integrable function h, and so must tend to zero (because, in fact, they're even square summable). $\square$

We end this chapter with a negative result, due to Nikolaev:

Theorem 8.14. *There is no weight w such that every $f \in C[\,a,b\,]$ has a uniformly convergent expansion in terms of orthogonal polynomials. In fact, given any w, there is always some f for which $\|f - S_n(f)\|$ is unbounded.*

Problems

▷ 1. Prove that every inner product norm is strictly convex. Specifically, let $\langle \cdot, \cdot \rangle$ be an inner product on a vector space X, and let $\|x\| = \sqrt{\langle x, x \rangle}$ be the associated norm. Show that:

   (a) $\|x+y\|^2 + \|x-y\|^2 = 2\,(\|x\|^2 + \|y\|^2)$ for all $x,\, y \in X$ (the parallelogram identity).

   (b) If $\|x\| = r = \|y\|$ and if $\|x-y\| = \delta$, then $\left\|\frac{x+y}{2}\right\|^2 = r^2 - (\delta/2)^2$. In particular, $\left\|\frac{x+y}{2}\right\| < r$ whenever $x \neq y$.

The remaining problems follow the notation given on page 79.

▷ 2. (a) Show that the expression $\|f\|_1 = \int_a^b |f(t)| w(t)\, dt$ also defines a norm on $C[\,a,b\,]$.

   (b) Given any f in $C[\,a,b\,]$, show that $\|f\|_1 \le c\|f\|_2$ and $\|f\|_2 \le c\|f\|$, where $c = \left(\int_a^b w(t)\, dt\right)^{1/2}$.

   (c) Conclude that the polynomials are dense in $C[\,a,b\,]$ under all three of the norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|$.

   (d) Show that $C[\,a,b\,]$ is *not* complete under either of the norms $\|\cdot\|_1$ or $\|\cdot\|_2$.

3. Check that Q_n is a monic polynomial of degree exactly n.

4. If (P_n) is another sequence of orthogonal polynomials such that P_n has degree exactly n, for each n, show that $P_n = \alpha_n Q_n$ for some $\alpha_n \neq 0$. In particular, if P_n is a monic polynomial, then $P_n = Q_n$. [Hint: Choose α_n so that $P_n - \alpha_n Q_n \in \mathcal{P}_{n-1}$ and note that $(P_n - \alpha_n Q_n) \perp \mathcal{P}_{n-1}$. Conclude that $P_n - \alpha_n Q_n = 0$.]

5. Given $w > 0$, $f \in C[\,a,b\,]$, and $n \ge 1$, show that $p^* \in \mathcal{P}_{n-1}$ is the least-squares approximation to f out of $\mathcal{P}_{n-1}$ (with respect to w) if and only if $\langle f - p^*\,,\, p \rangle = 0$ for every $p \in \mathcal{P}_{n-1}$; that is, if and only if $(f - p^*) \perp \mathcal{P}_{n-1}$.

6. In the notation of Problem 5, show that $f - p^*$ has at least n distinct zeros in (a, b).

7. If $w > 0$, show that the least-squares approximation to $f(x) = x^n$ out of $\mathcal{P}_{n-1}$ (relative to w) is $q_{n-1}^*(x) = x^n - Q_n(x)$.

▷ 8. Given $f \in C[\,a,b\,]$, let p_n^* denote the best *uniform* approximation to f out of $\mathcal{P}_n$ and let q_n^* denote the least-squares approximation to f out of $\mathcal{P}_n$. Show that $\|f - q_n^*\|_2 \le \|f - p_n^*\|_2$ and conclude that $\|f - q_n^*\|_2 \to 0$ as $n \to \infty$.

9. Show that the Chebyshev polynomials of the first kind, (T_n), and of the second kind, (U_n), satisfy the identities

$$T_n(x) = U_n(x) - x\,U_{n-1}(x)$$

and

$$(1 - x^2)\,U_{n-1}(x) = x\,T_n(x) - T_{n+1}(x).$$

10. Show that the Chebyshev polynomials of the second kind, (U_n), satisfy the recurrence relation

$$U_{n+1}(x) = 2x\,U_n(x) - U_{n-1}(x), \quad n \ge 1,$$

where $U_0(x) = 1$ and $U_1(x) = 2x$. [Compare this with the recurrence relation satisfied by the T_n.]

Chapter 9

Gaussian Quadrature

Introduction

Numerical integration, or *quadrature*, is the process of approximating the value of a definite integral $\int_a^b f(x)\,w(x)\,dx$ based only on a finite number of values or "samples" of f (much like a Riemann sum). A *linear quadrature formula* takes the form

$$\int_a^b f(x)\,w(x)\,dx \;\approx\; \sum_{k=1}^{n} A_k f(x_k),$$

where the *nodes* (x_k) and the *weights* (A_k) are at our disposal. (Note that both sides of the formula are linear in f.)

Example 9.1. Consider the quadrature formula

$$I(f) \;=\; \int_{-1}^{1} f(x)\,dx \;\approx\; \frac{1}{n}\sum_{k=-n}^{n-1} f\left(\frac{2k+1}{2n}\right) \;=\; I_n(f).$$

If f is continuous, then we clearly have $I_n(f) \to \int_{-1}^{1} f$ as $n \to \infty$. (Why?) But in the particular case $f(x) = x^2$ we have (after some simplification)

$$I_n(f) \;=\; \frac{1}{n}\sum_{k=-n}^{n-1}\left(\frac{2k+1}{2n}\right)^2 \;=\; \frac{1}{2n^3}\sum_{k=0}^{n-1}(2k+1)^2 \;=\; \frac{2}{3} - \frac{1}{6n^2}.$$

That is, $|\,I_n(f) - I(f)\,| = 1/6n^2$. In particular, we would need to take $n \ge 130$ to get $1/6n^2 \le 10^{-5}$, for example, and this would require that we perform over 250 evaluations of f. We'd like a method that converges a bit faster! In other words, there's no shortage of quadrature formulas—we just want *faster* ones.

A reasonable requirement for our proposed quadrature formula is that it be *exact for polynomials* of low degree. As it happens, this is easy to do.

Lemma 9.2. *Given $w(x)$ on $[\,a,b\,]$ and nodes $a \le x_1 < \cdots < x_n \le b$, there exist unique weights $A_1,\ldots,A_n$ such that*

$$\int_a^b p(x)\,w(x)\,dx \;=\; \sum_{i=1}^{n} A_i\,p(x_i)$$

for all polynomials $p \in \mathcal{P}_{n-1}$.

Proof. Let $\ell_1, \ldots, \ell_n$ be the Lagrange interpolating polynomials of degree $n-1$ associated to the nodes $x_1, \ldots, x_n$. Recall that we have $p = \sum_{i=1}^n p(x_i)\,\ell_i$ for all $p \in \mathcal{P}_{n-1}$. Hence,

$$\int_a^b p(x)\,w(x)\,dx \;=\; \sum_{i=1}^n p(x_i) \int_a^b \ell_i(x)\,w(x)\,dx.$$

That is, $A_i = \int_a^b \ell_i(x)\,w(x)\,dx$ works. To see that this is the only choice, suppose that

$$\int_a^b p(x)\,w(x)\,dx \;=\; \sum_{i=1}^n B_i\,p(x_i)$$

is exact for all $p \in \mathcal{P}_{n-1}$, and consider the case $p = \ell_j$:

$$A_j \;=\; \int_a^b \ell_j(x)\,w(x)\,dx \;=\; \sum_{i=1}^n B_i\,\ell_j(x_i) \;=\; B_j. \qquad \square$$

The point here is that integration is linear. In particular, when restricted to $\mathcal{P}_{n-1}$, integration is completely determined by its action on a basis for $\mathcal{P}_{n-1}$—in this setting, by the n values $A_i = I(\ell_i)$, $i = 1, \ldots, n$.

Said another way: Because the points $x_1, \ldots, x_n$ are distinct, the n point evaluations $\delta_i(p) = p(x_i)$ satisfy $\mathcal{P}_{n-1} \cap \left(\bigcap_{i=1}^n \ker \delta_i\right) = \{0\}$, and it follows that every linear, real-valued function on $\mathcal{P}_{n-1}$ must be a linear combination of the δ_i. Here's why: Because the x_i are distinct, $\mathcal{P}_{n-1}$ may be identified with $\mathbb{R}^n$ by way of the vector space *isomorphism* $p \mapsto (p(x_1), \ldots, p(x_n))$. Each linear, real-valued function on $\mathcal{P}_{n-1}$ must, then, correspond to a linear, real-valued function on $\mathbb{R}^n$—and any such map is given by inner product against some vector $(A_1, \ldots, A_n)$. In particular, we must have $I(p) = \sum_{i=1}^n A_i\,p(x_i)$.

In any case, we now have our quadrature formula: For $f \in C[\,a, b\,]$ we define $I_n(f) = \sum_{i=1}^n A_i\,f(x_i)$, where $A_i = \int_a^b \ell_i(x)\,w(x)\,dx$. But notice that the proof of Lemma 9.2 suggests an alternate way to write the formula. Indeed, if $L_{n-1}(f)(x) = \sum_{i=1}^n f(x_i)\ell_i(x)$ is the Lagrange interpolating polynomial for f of degree $n-1$ based on the nodes $x_1, \ldots, x_n$, then

$$\int_a^b (L_{n-1}(f))(x)\,w(x)\,dx \;=\; \sum_{i=1}^n f(x_i) \int_a^b \ell_i(x)\,w(x)\,dx \;=\; \sum_{i=1}^n A_i\,f(x_i).$$

In summary, $I_n(f) = I(L_{n-1}(f)) \approx I(f)$; that is,

$$I_n(f) \;=\; \sum_{i=1}^n A_i\,f(x_i) \;=\; \int_a^b (L_{n-1}(f))(x)\,w(x)\,dx \;\approx\; \int_a^b f(x)\,w(x)\,dx \;=\; I(f),$$

where L_{n-1} is the Lagrange interpolating polynomial of degree $n-1$ based on the nodes $x_1, \ldots, x_n$. This formula is obviously exact for $f \in \mathcal{P}_{n-1}$.

It's easy to give a bound on $|I_n(f)|$ in terms of $\|f\|$; indeed,

$$|I_n(f)| \;\le\; \sum_{i=1}^n |A_i|\,|f(x_i)| \;\le \|f\| \left(\sum_{i=1}^n |A_i|\right).$$

By considering a norm one continuous function f satisfying $f(x_i) = \text{sgn} A_i$ for each $i = 1, \ldots, n$, it's easy to see that $\sum_{i=1}^n |A_i|$ is the smallest constant that works in this inequality. In other words, $\lambda_n = \sum_{i=1}^n |A_i|$, $n = 1, 2, \ldots$, are the Lebesgue numbers for this process. As with all previous settings, we want these numbers to be uniformly bounded.

If $w(x) \equiv 1$ and if f is n-times continuously differentiable, we have an error estimate for our quadrature formula:

$$\left| \int_a^b f - \int_a^b L_{n-1}(f) \right| \le \int_a^b |f - L_{n-1}(f)| \le \frac{1}{n!} \|f^{(n)}\| \int_a^b \prod_{i=1}^n |x - x_i| \, dx$$

(recall Theorem 5.6). As it happens, the integral on the right is minimized when the x_i are taken to be the zeros of the Chebyshev polynomial U_n (see Rivlin [45, page 72]).

The fact that a quadrature formula is exact for polynomials of low degree does not by itself guarantee that the formula is highly accurate. The problem is that $\sum_{i=1}^n A_i f(x_i)$ may be estimating a very small quantity through the cancellation of very large quantities. So, for example, a positive function f may yield a negative value for this expression. This wouldn't happen if the A_i were all positive—and we've already seen how useful positivity can be. Our goal here is to further improve our quadrature formula to have this property. But we have yet to take advantage of the fact that the nodes x_i are at our disposal. We'll let Gauss show us the way!

Theorem 9.3. (Gauss) *Fix a weight $w(x)$ on $[a, b]$, and let (Q_n) be the canonical sequence of orthogonal polynomials relative to w. Given n, let $x_1, \ldots, x_n$ be the zeros of Q_n (these all lie in (a, b)), and choose weights $A_1, \ldots, A_n$ so that the formula $\sum_{i=1}^n A_i f(x_i) \approx \int_a^b f(x) \, w(x) \, dx$ is exact for polynomials of degree less than n. Then, in fact, the formula is exact for all polynomials of degree less than $2n$.*

Proof. Given a polynomial P of degree less than $2n$, we may divide: $P = Q_n R + S$, where R and S are polynomials of degree less than n. Then,

$$\begin{aligned}
\int_a^b P(x) \, w(x) \, dx &= \int_a^b Q_n(x) \, R(x) \, w(x) \, dx + \int_a^b S(x) \, w(x) \, dx \\
&= \int_a^b S(x) \, w(x) \, dx, \quad \text{because deg } R < n \\
&= \sum_{i=1}^n A_i S(x_i), \quad \text{because deg } S < n.
\end{aligned}$$

But $P(x_i) = Q_n(x_i) \, R(x_i) + S(x_i) = S(x_i)$, because $Q_n(x_i) = 0$. Hence, $\int_a^b P(x) \, w(x) \, dx = \sum_{i=1}^n A_i P(x_i)$ for all polynomials P of degree less than $2n$. $\square$

Amazing! But, well, not really: $\mathcal{P}_{2n-1}$ is of dimension $2n$, and we had $2n$ numbers $x_1, \ldots, x_n$ and $A_1, \ldots, A_n$ to choose as we saw fit. Said another way, the division algorithm tells us that $\mathcal{P}_{2n-1} = Q_n \mathcal{P}_{n-1} \oplus \mathcal{P}_{n-1}$. Because $Q_n \mathcal{P}_{n-1} \subset \ker(I_n)$, the action of I_n on $\mathcal{P}_{2n-1}$ is the same as its action on a "copy" of $\mathcal{P}_{n-1}$ (where its known to be exact).

In still other words, because any polynomial that vanishes at all the x_i must be divisible by Q_n (and conversely), we have $Q_n \mathcal{P}_{n-1} = \mathcal{P}_{2n-1} \cap (\bigcap_{i=1}^n \ker \delta_i) = \ker(I_n \mid_{\mathcal{P}_{2n-1}})$. Thus, I_n "factors through" the quotient space $\mathcal{P}_{2n-1} / Q_n \mathcal{P}_{n-1} = \mathcal{P}_{n-1}$.

Also not surprising is that this particular choice of x_i is unique.

Lemma 9.4. *Suppose that $a \le x_1 < \cdots < x_n \le b$ and $A_1, \ldots, A_n$ are given so that the equation $\int_a^b P(x)\,w(x)\,dx = \sum_{i=1}^n A_i P(x_i)$ is satisfied for all polynomials P of degree less than $2n$. Then $x_1, \ldots, x_n$ are the zeros of Q_n.*

Proof. Let $Q(x) = \prod_{i=1}^n (x - x_i)$. Then, for $k < n$, the polynomial $Q \cdot Q_k$ has degree $n + k < 2n$. Hence,

$$\int_a^b Q(x)\,Q_k(x)\,w(x)\,dx \;=\; \sum_{i=1}^n A_i Q(x_i)\,Q_k(x_i) = 0.$$

Because Q is a monic polynomial of degree n which is orthogonal to each Q_k, $k < n$, we must have $Q = Q_n$. Thus, the x_i are actually the zeros of Q_n. □

According to Rivlin, the phrase *Gaussian quadrature* is usually reserved for the specific quadrature formula whereby $\int_{-1}^1 f(x)\,dx$ is approximated by $\int_{-1}^1 (L_{n-1}(f))(x)\,dx$, where $L_{n-1}(f)$ is the Lagrange interpolating polynomial to f using the zeros of the n-th Legendre polynomial as nodes. (What a mouthful!) What is actually being described in our version of Gauss's theorem is *Gaussian-type quadrature*.

Before computers, Gaussian quadrature was little more than a curiosity; the roots of Q_n are typically irrational, and certainly not easy to find by hand. By now, though, it's considered a standard quadrature technique. In any case, we still can't judge the quality of Gauss's method without a bit more information.

Gaussian-type Quadrature

First, let's summarize our rather cumbersome notation.

orthogonal polynomial	zeros	weights	approximate integral
Q_1	$x_1^{(1)}$	$A_1^{(1)}$	I_1
Q_2	$x_1^{(2)}, x_2^{(2)}$	$A_1^{(2)}, A_2^{(2)}$	I_2
Q_3	$x_1^{(3)}, x_2^{(3)}, x_3^{(3)}$	$A_1^{(3)}, A_2^{(3)}, A_3^{(3)}$	I_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Hidden here is the Lagrange interpolation formula $L_{n-1}(f) = \sum_{i=1}^n f(x_i^{(n)})\,\ell_i^{(n-1)}$, where $\ell_i^{(n-1)}$ denote the Lagrange polynomials of degree $n-1$ based on the nodes $x_1^{(n)}, \ldots, x_n^{(n)}$. The n-th quadrature formula is then

$$I_n(f) \;=\; \int_a^b L_{n-1}(f)(x)\,w(x)\,dx \;=\; \sum_{i=1}^n A_i^{(n)} f(x_i^{(n)}) \;\approx\; \int_a^b f(x)\,w(x)\,dx,$$

which is exact for polynomials of degree less than $2n$.

By way of one example, Hermite showed that $A_k^{(n)} = \pi/n$ for the Chebyshev weight $w(x) = (1-x^2)^{-1/2}$ on $[-1,1]$. Remarkably, $A_k^{(n)}$ doesn't depend on k! The quadrature formula in this case reads:

$$\int_{-1}^1 \frac{f(x)\,dx}{\sqrt{1-x^2}} \;\approx\; \frac{\pi}{n}\sum_{k=1}^n f\left(\cos\frac{2k-1}{2n}\,\pi\right).$$

Or, if you prefer,

$$\int_{-1}^{1} f(x)\,dx \;\approx\; \frac{\pi}{n} \sum_{k=1}^{n} f\left(\cos \frac{2k-1}{2n}\,\pi\right) \sin \frac{2k-1}{2n}\,\pi.$$

(Why?) You can find full details in Natanson [41, Vol. III].

The key result, due to Stieltjes, is that I_n is *positive*:

Lemma 9.5. $A_1^{(n)}, \ldots, A_n^{(n)} > 0$ *and* $\sum_{i=1}^n A_i^{(n)} = \int_a^b w(x)\,dx$.

Proof. The second assertion is obvious (because $I_n(1) = I(1)$). For the first, fix $1 \le j \le n$ and notice that $(\ell_j^{(n-1)})^2$ is of degree $2(n-1) < 2n$. Thus,

$$\begin{aligned} 0 \;<\; \langle\, \ell_j^{(n-1)}, \ell_j^{(n-1)} \,\rangle \;&=\; \int_a^b \left[\ell_j^{(n-1)}(x)\right]^2 w(x)\,dx \\ &=\; \sum_{i=1}^{n} A_i^{(n)} \left[\ell_j^{(n-1)}(x_i^{(n)})\right]^2 \;=\; A_j^{(n)}, \end{aligned}$$

because $\ell_j^{(n-1)}(x_i^{(n)}) = \delta_{i,j}$. □

Now our last calculation is quite curious; what we've shown is that

$$A_j^{(n)} \;=\; \int_a^b \ell_j^{(n-1)}(x)\,w(x)\,dx \;=\; \int_a^b \left[\ell_j^{(n-1)}(x)\right]^2 w(x)\,dx.$$

Essentially the same calculation as above also proves

Corollary 9.6. $\langle\, \ell_i^{(n-1)}, \ell_j^{(n-1)} \,\rangle = 0$ *for* $i \ne j$.

Because $A_1^{(n)}, \ldots, A_n^{(n)} > 0$, it follows that $I_n(f)$ is positive; that is, $I_n(f) \ge 0$ whenever $f \ge 0$. The second assertion in Lemma 9.5 tells us that the I_n are *uniformly bounded*:

$$|I_n(f)| \;\le\; \|f\| \sum_{i=1}^{n} A_i^{(n)} \;=\; \|f\| \int_a^b w(x)\,dx,$$

and this is the same bound that holds for $I(f) = \int_a^b f(x)\,w(x)\,dx$ itself. Given all of this, proving that $I_n(f) \to I(f)$ is a piece of cake. The following result is again due to Stieltjes (á la Lebesgue).

Theorem 9.7. *In the above notation,* $|I_n(f) - I(f)| \le 2\left(\int_a^b w(x)\,dx\right) E_{2n-1}(f)$. *In particular,* $I_n(f) \to I(f)$ *for evey* $f \in C[\,a, b\,]$.

Proof. Let p^* be the best uniform approximation to f out of $\mathcal{P}_{2n-1}$. Then, because $I_n(p^*) = I(p^*)$, we have

$$\begin{aligned} |I(f) - I_n(f)| \;&\le\; |I(f - p^*)| \;+\; |I_n(f - p^*)| \\ &\le\; \|f - p^*\| \int_a^b w(x)\,dx \;+\; \|f - p^*\| \sum_{i=1}^{n} A_i^{(n)} \\ &=\; 2\,\|f - p^*\| \int_a^b w(x)\,dx \;=\; 2E_{2n-1}(f) \int_a^b w(x)\,dx. \end{aligned}$$

□

Computational Considerations

You've probably been asking yourself: "How do I find the A_i without integrating?" Well, let's first recall the definition: In the case of Gaussian-type quadrature we have

$$A_i^{(n)} \;=\; \int_a^b \ell_i^{(n-1)}(x)\,w(x)\,dx \;=\; \int_a^b \frac{Q_n(x)}{(x - x_i^{(n)})\,Q_n'(x_i^{(n)})}\,w(x)\,dx$$

(because "W" is the same as Q_n here—the x_i are the zeros of Q_n). Next, consider the function

$$\varphi_n(x) \;=\; \int_a^b \frac{Q_n(t) - Q_n(x)}{t - x}\,w(t)\,dt.$$

Because $t - x$ divides $Q_n(t) - Q_n(x)$, note that φ_n is actually a *polynomial* (of degree at most $n-1$) and that

$$\varphi_n(x_i^{(n)}) \;=\; \int_a^b \frac{Q_n(t)}{t - x_i^{(n)}}\,w(t)\,dt \;=\; A_i^{(n)} Q_n'(x_i^{(n)}).$$

Now $Q_n'(x_i^{(n)})$ is readily available; we just need to compute $\varphi_n(x_i^{(n)})$.

Lemma 9.8. *The φ_n satisfy the same recurrence formula as the Q_n; namely,*

$$\varphi_{n+1}(x) = (x - a_n)\varphi_n(x) - b_n\varphi_{n-1}(x), \quad n \ge 1,$$

but with different starting values

$$\varphi_0(x) \equiv 0, \qquad \text{and} \qquad \varphi_1(x) \equiv \int_a^b w(x)\,dx.$$

Proof. The formulas for φ_0 and φ_1 are obviously correct, because $Q_0(x) \equiv 1$ and $Q_1(x) = x - a_0$. We only need to check the recurrence formula itself.

$$\begin{aligned}
\varphi_{n+1}(x) &= \int_a^b \frac{Q_{n+1}(t) - Q_{n+1}(x)}{t - x}\,w(t)\,dt \\
&= \int_a^b \frac{(t - a_n)\,Q_n(t) - b_n Q_{n-1}(t) - (x - a_n)\,Q_n(x) + b_n Q_{n-1}(x)}{t - x}\,w(t)\,dt \\
&= (x - a_n)\int_a^b \frac{Q_n(t) - Q_n(x)}{t - x}\,w(t)\,dt \;-\; b_n\int_a^b \frac{Q_{n-1}(t) - Q_{n-1}(x)}{t - x}\,w(t)\,dt \\
&= (x - a_n)\,\varphi_n(x) - b_n\,\varphi_{n-1}(x),
\end{aligned}$$

because $\int_a^b Q_n(t)\,w(t)\,dt = 0$. □

Of course, the derivatives Q_n' satisfy a recurrence relation of sorts, too:

$$Q_{n+1}'(x) \;=\; Q_n(x) + (x - a_n)\,Q_n'(x) - b_n\,Q_{n-1}'(x).$$

But $Q_n'(x_i^{(n)})$ can be computed without knowing $Q_n'(x)$. Indeed, $Q_n(x) = \prod_{i=1}^n (x - x_i^{(n)})$, so we have $Q_n'(x_i^{(n)}) = \prod_{j\ne i}(x_i^{(n)} - x_j^{(n)})$.

The weights $A_i^{(n)}$, or *Christoffel numbers*, together with the zeros of Q_n are tabulated in a variety of standard cases. See, for example, Abramowitz and Stegun [1] (this wonderful resource is now available online through the U.S. Department of Commerce). In practice, of course, it's enough to tabulate data for the case $[\,a, b\,] = [-1, 1\,]$.

Applications to Interpolation

Although $L_n(f)$ isn't typically a good uniform approximation to f, if we interpolate at the zeros of an orthogonal polynomial Q_{n+1}, then $L_n(f)$ will be a good approximation in the $\|\cdot\|_1$ or $\|\cdot\|_2$ norm generated by the corresponding weight w. Specifically, by rewording our earlier results, it's easy to get estimates for each of the errors $\int_a^b |f - L_n(f)|\, w$ and $\int_a^b |f - L_n(f)|^2\, w$. We use essentially the same notation as before, except now we take

$$L_n(f) \;=\; \sum_{i=1}^{n+1} f\big(x_i^{(n+1)}\big)\, \ell_i^{(n)},$$

where $x_1^{(n+1)}, \ldots, x_{n+1}^{(n+1)}$ are the roots of Q_{n+1} and $\ell_i^{(n)}$ is of degree n. This leads to a quadrature formula that's exact on polynomials of degree less than $2(n+1)$.

As we've already seen, $\ell_1^{(n)}, \ldots, \ell_{n+1}^{(n)}$ are *orthogonal* and so $\|L_n(f)\|_2$ may be computed exactly.

Lemma 9.9. $\|L_n(f)\|_2 \;\le\; \|f\| \left(\int_a^b w(x)\,dx\right)^{1/2}$.

Proof. Because $L_n(f)^2$ is a polynomial of degree $\le 2n < 2(n+1)$, we have

$$\begin{aligned}
\|L_n(f)\|_2^2 \;&=\; \int_a^b [L_n(f)]^2\, w(x)\,dx \\
&=\; \sum_{j=1}^{n+1} A_j^{(n+1)} \left[\sum_{i=1}^{n+1} f\big(x_i^{(n+1)}\big)\, \ell_i^{(n)}\big(x_j^{(n+1)}\big)\right]^2 \\
&=\; \sum_{j=1}^{n+1} A_j^{(n+1)} \left[f\big(x_j^{(n+1)}\big)\right]^2 \\
&\le\; \|f\|^2 \sum_{j=1}^{n+1} A_j^{(n+1)} \;=\; \|f\|^2 \int_a^b w(x)\,dx. \quad \square
\end{aligned}$$

Please note that we also have $\|f\|_2 \le \|f\| \left(\int_a^b w(x)\,dx\right)^{1/2}$; that is, this same estimate holds for $\|f\|_2$ itself.

As usual, once we have an estimate for the norm of an operator, we also have an analogue of Lebesgue's theorem.

Theorem 9.10. $\|f - L_n(f)\|_2 \;\le\; 2\left(\int_a^b w(x)\,dx\right)^{1/2} E_n(f)$.

Proof. Here we go again! Let p^* be the best uniform approximation to f out of $\mathcal{P}_n$ and use the fact that $L_n(p^*) = p^*$.

$$\begin{aligned}
\|f - L_n(f)\|_2 \;&\le\; \|f - p^*\|_2 + \|L_n(f - p^*)\|_2 \\
&\le\; \|f - p^*\| \left(\int_a^b w(x)\,dx\right)^{1/2} \;+\; \|f - p^*\| \left(\int_a^b w(x)\,dx\right)^{1/2} \\
&=\; 2E_n(f) \left(\int_a^b w(x)\,dx\right)^{1/2}. \qquad \square
\end{aligned}$$

Hence, if we interpolate $f \in C[\,a, b\,]$ at the zeros of (Q_n), then $L_n(f) \to f$ in $\|\cdot\|_2$ norm. The analogous result for the $\|\cdot\|_1$ norm is now easy:

Corollary 9.11. $\int_a^b |f(x) - L_n(f)(x)|\, w(x)\, dx \;\le\; 2\left(\int_a^b w(x)\, dx\right) E_n(f)$.

Proof. We apply the Cauchy-Schwarz inequality:

$$\begin{aligned}
\int_a^b |f(x) - L_n(f)(x)|\, w(x)\, dx &= \int_a^b |f(x) - L_n(f)(x)|\, \sqrt{w(x)}\, \sqrt{w(x)}\, dx \\
&\le \left(\int_a^b |f(x) - L_n(f)(x)|^2\, w(x)\, dx\right)^{1/2} \left(\int_a^b w(x)\, dx\right)^{1/2} \\
&\le 2E_n(f) \int_a^b w(x)\, dx. \qquad \square
\end{aligned}$$

Essentially the same device allows an estimate of $\int_a^b f(x)\, dx$ in terms of $\int_a^b f(x)\, w(x)\, dx$ (which may be easier to compute). As this is an easy calculation, we'll combine both statement and proof:

Corollary 9.12. *If* $\int_a^b w(x)^{-1}\, dx$ *is finite, then*

$$\begin{aligned}
\int_a^b |f(x) - L_n(f)(x)|\, dx &= \int_a^b |f(x) - L_n(f)(x)|\, \sqrt{w(x)}\, \frac{1}{\sqrt{w(x)}}\, dx \\
&\le \left(\int_a^b |f(x) - L_n(f)(x)|^2\, w(x)\, dx\right)^{1/2} \left(\int_a^b \frac{1}{w(x)}\, dx\right)^{1/2} \\
&\le 2E_n(f) \left(\int_a^b w(x)\, dx\right)^{1/2} \left(\int_a^b \frac{1}{w(x)}\, dx\right)^{1/2}. \qquad \square
\end{aligned}$$

In particular, the Chebyshev weight satisfies

$$\int_{-1}^1 \frac{dx}{\sqrt{1-x^2}} = \pi \qquad \text{and} \qquad \int_{-1}^1 \sqrt{1-x^2}\, dx = \frac{\pi}{2}.$$

Thus, interpolation at the zeros of the Chebyshev polynomials (of the first kind) would provide good, *simultaneous* approximation in each of the norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|$.

The Moment Problem

Given a positive, continuous weight function $w(x)$ on $[\,a, b\,]$, the number

$$\mu_k \;=\; \int_a^b x^k\, w(x)\, dx$$

is called the k-th *moment* of w. In physical terms, if we think of $w(x)$ as the density of a thin rod placed on the interval $[\,a, b\,]$, then μ_0 is the mass of the rod, μ_1/μ_0 is its center of mass, μ_2 is its moment of inertia (about 0), and so on. In probabilistic terms, if $\mu_0 = 1$, then w is the probability density function for some random variable, μ_1 is the expected value, or mean, of this random variable, and $\mu_2 - \mu_1^2$ is its variance. The moment problem (or problems, really) concern the inverse procedure. What can be measured in real life are the moments—can the moments be used to find the density function?

> **Questions**: Do the moments determine w? Do different weights have different moment sequences? If we knew the sequence (μ_k), could we recover w? How do we tell if a given sequence (μ_k) is the moment sequence for some positive weight? Do "special" weights give rise to "special" sequences?

Now we've already answered one of these questions: The Weierstrass theorem tells us that different weights have different moment sequences. Said another way, if

$$\int_a^b x^k\, w(x)\, dx = 0 \qquad \text{for all } k = 0, 1, 2, \ldots,$$

then $w \equiv 0$. Indeed, by linearity, this says that $\int_a^b p(x)\, w(x)\, dx = 0$ for all polynomials p which, in turn, tells us that $\int_a^b w(x)^2\, dx = 0$. (Why?) The remaining questions are harder to answer. We'll settle for simply stating a few pertinent results.

Given a sequence of numbers (μ_k), we define the n-th difference sequence $(\Delta^n \mu_k)$ by

$$\begin{aligned} \Delta^0 \mu_k &= \mu_k \\ \Delta^1 \mu_k &= \mu_k - \mu_{k+1} \\ \Delta^n \mu_k &= \Delta^{n-1}\mu_k - \Delta^{n-1}\mu_{k+1}, \quad n \ge 1. \end{aligned}$$

For example, $\Delta^2 \mu_k = \mu_k - 2\mu_{k+1} + \mu_{k+2}$. More generally, induction will show that

$$\Delta^n \mu_k = \sum_{i=0}^{n} (-1)^i \binom{n}{i} \mu_{k+i}.$$

In the case of a weight w on the interval $[\,0,1\,]$, this sum is easy to recognize as an integral. Indeed,

$$\int_0^1 x^k (1-x)^n\, w(x)\, dx = \sum_{i=0}^{n} (-1)^i \binom{n}{i} \int_0^1 x^{k+i}\, w(x)\, dx = \sum_{i=0}^{n} (-1)^i \binom{n}{i} \mu_{k+i}.$$

In particular, if w is nonnegative, then we must have $\Delta^n \mu_k \ge 0$ for every n and k. This observation serves as motivation for

Theorem 9.13. *The following are equivalent:*

(a) (μ_k) *is the moment sequence of some nonnegative weight function* w *on* $[\,0,1\,]$.

(b) $\Delta^n \mu_k \ge 0$ *for every* n *and* k.

(c) $a_0\mu_0 + a_1\mu_1 + \cdots + a_n\mu_n \ge 0$ *whenever* $a_0 + a_1 x + \cdots + a_n x^n \ge 0$ *for all* $0 \le x \le 1$.

The equivalence of (a) and (b) is due to Hausdorff. A real sequence satisfying (b) or (c) is sometimes said to be *positive definite*.

Now dozens of mathematicians worked on various aspects of the moment problem: Chebyshev, Markov, Stieltjes, Cauchy, Riesz, Fréchet, and on and on. And several of them, in particular Cauchy and Stieltjes, noticed the importance of the integral $\int_a^b \frac{w(t)}{x-t}\,dt$ in attacking the problem (compare this to Cauchy's well-known integral formula from complex analysis). It was Stieltjes, however, who gave the first complete solution to such a problem—developing his own integral (by considering $\int_a^b \frac{dW(t)}{x-t}$), his own variety of continued fractions, and planting the seeds for the study of orthogonal polynomials while he was at it! We will attempt to at least sketch a few of these connections.

To begin, let's fix our notation: To simplifiy things, we suppose that we're given a nonnegative weight $w(x)$ on a *symmetric* interval $[-a, a\,]$, and that all of the moments of w are finite. We will otherwise stick to our usual notations for (Q_n), the Gaussian-type quadrature formulas, and so on. Next, we consider *the moment-generating function*:

Lemma 9.14. *If* $x \notin [-a, a\,]$, *then* $\displaystyle\int_{-a}^{a} \frac{w(t)}{x-t}\,dt \;=\; \sum_{k=0}^{\infty} \frac{\mu_k}{x^{k+1}}.$

Proof. $\displaystyle\frac{1}{x-t} = \frac{1}{x} \cdot \frac{1}{1-(t/x)} = \sum_{k=0}^{\infty} \frac{t^k}{x^{k+1}}$, and the sum converges uniformly because $|t/x| \le a/|x| < 1$. Now just multiply by $w(t)$ and integrate. □

By way of an example, consider the Chebyshev weight $w(x) = (1-x^2)^{-1/2}$ on $[-1, 1\,]$. For $x > 1$ we have

$$\begin{aligned}
\int_{-1}^{1} \frac{dt}{(x-t)\sqrt{1-t^2}} &= \frac{\pi}{\sqrt{x^2-1}} \qquad \big(\text{set } t = 2u/(1+u^2)\big) \\
&= \frac{\pi}{x}\left(1 - \frac{1}{x^2}\right)^{-1/2} \\
&= \frac{\pi}{x}\left[1 + \frac{1}{2}\cdot\frac{1}{x^2} + \frac{1\cdot 3}{2\cdot 2}\cdot\frac{1}{2!}\cdot\frac{1}{x^4} + \cdots\right],
\end{aligned}$$

using the binomial formula. Thus, we've found all the moments:

$$\begin{aligned}
\mu_0 &= \int_{-1}^{1} \frac{dt}{\sqrt{1-t^2}} = \pi \\
\mu_{2n-1} &= \int_{-1}^{1} \frac{t^{2n-1}dt}{\sqrt{1-t^2}} = 0 \\
\mu_{2n} &= \int_{-1}^{1} \frac{t^{2n}dt}{\sqrt{1-t^2}} = \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2^n n!}\,\pi.
\end{aligned}$$

Stieltjes proved much more: The integral $\int_{-a}^{a} \frac{w(t)}{x-t}\,dt$ is actually an *analytic* function of x in $\mathbb{C} \setminus [-a, a\,]$. In any case, because $x \notin [-a, a\,]$, we know that $\frac{1}{x-t}$ is continuous on $[-a, a\,]$. In particular, we can apply our quadrature formulas and Stieltjes' theorem (Theorem 9.7) to write

$$\int_{-a}^{a} \frac{w(t)}{x-t}\,dt \;=\; \lim_{n\to\infty} \sum_{i=1}^{n} \frac{A_i^{(n)}}{x - x_i^{(n)}},$$

and these sums are recognizable:

Lemma 9.15. $\displaystyle\sum_{i=1}^{n} \frac{A_i^{(n)}}{x - x_i^{(n)}} \;=\; \frac{\varphi_n(x)}{Q_n(x)}.$

Proof. Because φ_n has degree $< n$ and $\varphi_n(x_i^{(n)}) \neq 0$ for any i, we may appeal to the method of partial-fractions to write

$$\frac{\varphi_n(x)}{Q_n(x)} = \frac{\varphi_n(x)}{(x - x_1^{(n)}) \cdots (x - x_n^{(n)})} = \sum_{i=1}^{n} \frac{c_i}{x - x_i^{(n)}}$$

where c_i is given by

$$c_i = \frac{\varphi_n(x)}{Q_n(x)} (x - x_i^{(n)}) \Big]_{x = x_i^{(n)}} = \frac{\varphi_n(x_i^{(n)})}{Q_n'(x_i^{(n)})} = A_i^{(n)}. \qquad \square$$

Now here's where the continued fractions come in: Stieltjes recognized the fact that

$$\frac{\varphi_{n+1}(x)}{Q_{n+1}(x)} = \cfrac{b_0}{(x - a_0) - \cfrac{b_1}{(x - a_1) - \ddots - \cfrac{b_n}{(x - a_n)}}}$$

(which can be proved by induction), where $b_0 = \int_a^b w(t)\, dt$. More generally, induction will show that the n-th *convergent* of a continued fraction can be written as

$$\frac{A_n}{B_n} = \cfrac{p_1}{q_1 - \cfrac{p_2}{q_2 - \ddots - \frac{p_n}{q_n}}}$$

by means of the *recurrence formulas*

$$\begin{aligned} A_0 &= 0 & \qquad B_0 &= 1 \\ A_1 &= p_1 & B_1 &= q_1 \\ A_n &= q_n A_{n-1} + p_n A_{n-2} & B_n &= q_n B_{n-1} + p_n B_{n-2} \end{aligned}$$

where $n = 2, 3, 4, \ldots$. Please note that A_n and B_n satisfy the same recurrence formula, but with different starting values (as is the case with φ_n and Q_n).

Again using the Chebyshev weight as an example, for $x > 1$ we have

$$\frac{\pi}{\sqrt{x^2 - 1}} = \int_{-1}^{1} \frac{dt}{(x - t)\sqrt{1 - t^2}} = \cfrac{\pi}{x - \cfrac{1/2}{x - \cfrac{1/4}{x - \cfrac{1/4}{\ddots}}}}$$

because $a_n = 0$ for all n, $b_1 = 1/2$, and $b_n = 1/4$ for $n \geq 2$. In other words, we've just found a continued fraction expansion for $(x^2 - 1)^{-1/2}$.

Chapter 10

The Müntz Theorems

For several weeks now we've taken advantage of the fact that the monomials $1, x, x^2, \ldots$ have *dense linear span* in $C[\,0,1\,]$. What, if anything, is so special about these particular powers? How about if we consider polynomials of the form $\sum_{k=0}^{n} a_k x^{k^2}$; are they dense, too? More generally, what can be said about the span of a sequence of monomials (x^{λ_n}), where $\lambda_0 < \lambda_1 < \lambda_2 < \cdots$? Of course, we'll have to assume that $\lambda_0 \ge 0$, but it's not hard to see that we will actually need $\lambda_0 = 0$, for otherwise each of the polynomials $\sum_{k=0}^{n} a_k x^{\lambda_k}$ vanishes at $x = 0$ (and so has distance at least 1 from the constant 1 function, for example). If the λ_n are integers, it's also clear that we'll have to have $\lambda_n \to \infty$ as $n \to \infty$. But what else is needed? The answer comes to us from Müntz in 1914. (You sometimes see the name Otto Szász associated with Müntz's theorem, because Szász proved a similar theorem at nearly the same time (1916).)

Theorem 10.1. *Let $0 \le \lambda_0 < \lambda_1 < \lambda_2 < \cdots$. Then the functions (x^{λ_n}) have dense linear span in $C[\,0,1\,]$ if and only if $\lambda_0 = 0$ and $\sum_{n=1}^{\infty} \lambda_n^{-1} = \infty$.*

What Müntz is trying to tell us here is that the λ_n can't get big too quickly. In particular, the polynomials of the form $\sum_{k=0}^{n} a_k x^{k^2}$ are evidently *not* dense in $C[\,0,1\,]$. On the other hand, the λ_n don't have to be unbounded; indeed, Müntz's theorem implies an earlier result of Bernstein from 1912: If $0 < \alpha_1 < \alpha_2 < \cdots < K$ (some constant), then $1, x^{\alpha_1}, x^{\alpha_2}, \ldots$ have dense linear span in $C[\,0,1\,]$.

Before we give the proof of Müntz's theorem, let's invent a bit of notation: We write

$$X_n \;=\; \left\{ \sum_{k=0}^{n} a_k x^{\lambda_k} : a_0, \ldots, a_n \in \mathbb{R} \right\} \;=\; \operatorname{span}\{\, x^{\lambda_k} : k = 0, \ldots, n \,\}$$

and, given $f \in C[\,0,1\,]$, we write $\operatorname{dist}(f, X_n)$ to denote the distance from f to X_n. Let's also write $X = \bigcup_{n=0}^{\infty} X_n$. That is, X is the linear span of the entire sequence $(x^{\lambda_n})_{n=0}^{\infty}$. The question here is whether X is dense, and we'll address the problem by determining whether $\operatorname{dist}(f, X_n) \to 0$, as $n \to \infty$, for every $f \in C[\,0,1\,]$.

If we can show that each (fixed) power x^m can be uniformly approximated by a linear combination of x^{λ_n}, then the Weierstrass theorem will imply that X is dense in $C[\,0,1\,]$. (How?) Surprisingly, the numbers $\operatorname{dist}(x^m, X_n)$ can be estimated. Our proof won't give the best estimate, but it will show how the condition $\sum_{n=1}^{\infty} \lambda_n^{-1} = \infty$ comes into the picture.

Lemma 10.2. *Let $m > 0$. Then,* $\displaystyle \operatorname{dist}(x^m, X_n) \;\le\; \prod_{k=1}^{n} \left| 1 - \frac{m}{\lambda_k} \right|$.

Proof. We may certainly assume that $m \ne \lambda_n$ for any n. Given this, we inductively define a sequence of functions by setting $P_0(x) = x^m$ and

$$P_n(x) \;=\; (\lambda_n - m)\, x^{\lambda_n} \int_x^1 t^{-1-\lambda_n}\, P_{n-1}(t)\, dt$$

for $n \ge 1$. For example,

$$P_1(x) \;=\; (\lambda_1 - m)\, x^{\lambda_1} \int_x^1 t^{-1-\lambda_1}\, t^m\, dt \;=\; -x^{\lambda_1}\, t^{m-\lambda_1}\Big]_x^1 \;=\; x^m - x^{\lambda_1}.$$

By induction, each P_n is of the form $x^m - \sum_{k=0}^{n} a_k x^{\lambda_k}$ for some scalars (a_k):

$$\begin{aligned} P_n(x) &= (\lambda_n - m)\, x^{\lambda_n} \int_x^1 t^{-1-\lambda_n}\, P_{n-1}(t)\, dt \\ &= (\lambda_n - m)\, x^{\lambda_n} \int_x^1 t^{-1-\lambda_n} \left[t^m - \sum_{k=0}^{n-1} a_k t^{\lambda_k} \right] dt \\ &= x^m - x^{\lambda_n} + (\lambda_n - m) \sum_{k=0}^{n-1} \frac{a_k}{\lambda_n - \lambda_k}\, (x^{\lambda_k} - x^{\lambda_n}). \end{aligned}$$

Finally, $\|P_0\| = 1$ and $\|P_n\| \le \left|1 - \frac{m}{\lambda_n}\right| \|P_{n-1}\|$, because

$$|\lambda_n - m|\, x^{\lambda_n} \int_x^1 t^{-1-\lambda_n}\, dt \;=\; \frac{|\lambda_n - m|}{\lambda_n}\, (1 - x^{\lambda_n}) \;\le\; \left| 1 - \frac{m}{\lambda_n} \right|.$$

Thus,

$$\operatorname{dist}(x^m, X_n) \;\le\; \|P_n\| \;\le\; \prod_{k=1}^{n} \left| 1 - \frac{m}{\lambda_k} \right|. \qquad \square$$

The preceding result is due to v. Golitschek. A slightly better estimate, also due to v. Golitschek (1970), is $\operatorname{dist}(x^m, X_n) \le \prod_{k=1}^{n} \frac{|m-\lambda_k|}{m+\lambda_k}$.

Now a well-known fact about infinite products is that for positive a_k, the product $\prod_{k=1}^{\infty} \left|1 - a_k\right|$ diverges (to 0) if and only if the series $\sum_{k=1}^{\infty} a_k$ diverges (to ∞) if and only if the product $\prod_{k=1}^{\infty} \left|1 + a_k\right|$ diverges (to ∞). In particular, $\prod_{k=1}^{n} \left|1 - \frac{m}{\lambda_k}\right| \to 0$ if and only if $\sum_{k=1}^{n} \frac{1}{\lambda_k} \to \infty$. That is, $\operatorname{dist}(x^m, X_n) \to 0$ if and only if $\sum_{k=1}^{\infty} \frac{1}{\lambda_k} = \infty$. This proves the "backward" direction of Müntz's theorem.

We'll prove the "forward" direction of Müntz's theorem by proving a version of Müntz's theorem for the space $L_2[\,0,1\,]$. For our purposes, $L_2[\,0,1\,]$ denotes the space $C[\,0,1\,]$ endowed with the norm

$$\|f\|_2 \;=\; \left(\int_0^1 |f(x)|^2 dx \right)^{1/2},$$

although our results are equally valid in the "official" space $L_2[\,0,1\,]$ (consisting of square-integrable, Lebegue measurable functions). In the latter case, we no longer need to assume that $\lambda_0 = 0$, but we do need to assume that each $\lambda_n > -1/2$ (in order that $x^{2\lambda_n}$ be integrable on $[\,0,1\,]$).

Remarkably, the distance from f to X_n can be computed *exactly* in the L_2 norm. For this we'll need a bit more notation: Given linearly independent vectors $f_1, \ldots, f_n$ in an inner product space, we call

$$G(f_1, \ldots, f_n) \;=\; \begin{vmatrix} \langle f_1, f_1 \rangle & \cdots & \langle f_1, f_n \rangle \\ \vdots & \ddots & \vdots \\ \langle f_n, f_1 \rangle & \cdots & \langle f_n, f_n \rangle \end{vmatrix} \;=\; \det\big[\, \langle f_i, f_j \rangle \,\big]_{i,j}$$

the *Gram determinant* of the f_k.

Lemma 10.3. (Gram) *Let F be a finite dimensional subspace of an inner product space V, and let $g \in V \setminus F$. Then the distance d from g to F satisfies*

$$d^{\,2} \;=\; \frac{G(g, f_1, \ldots, f_n)}{G(f_1, \ldots, f_n)},$$

where $f_1, \ldots, f_n$ is any basis for F.

Proof. Let $f = \sum_{i=1}^n a_i f_i$ be the best approximation to g out of F. Then, because $g - f$ is orthogonal to F, we have, in particular, $\langle f_j, f \rangle = \langle f_j, g \rangle$ for all j; that is,

$$\sum_{i=1}^n a_i \langle f_j, f_i \rangle \;=\; \langle f_j, g \rangle, \quad j = 1, \ldots, n. \tag{10.1}$$

Because this system of equations always has a unique solution $a_1, \ldots, a_n$, we must have $G(f_1, \ldots, f_n) \neq 0$ (and so the formula in the statement of the Lemma at least makes sense).

Next, notice that

$$d^{\,2} \;=\; \langle g - f, g - f \rangle \;=\; \langle g - f, g \rangle \;=\; \langle g, g \rangle \;-\; \langle g, f \rangle;$$

in other words,

$$d^{\,2} \;+\; \sum_{i=1}^n a_i \langle g, f_i \rangle = \; \langle g, g \rangle. \tag{10.2}$$

Now consider (10.1) and (10.2) as a system of $n+1$ equations in the $n+1$ unknowns $a_1, \ldots, a_n$, and $d^{\,2}$; in matrix form we have

$$\begin{bmatrix} 1 & \langle g, f_1 \rangle & \cdots & \langle g, f_n \rangle \\ 0 & \langle f_1, f_1 \rangle & \cdots & \langle f_1, f_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \langle f_n, f_1 \rangle & \cdots & \langle f_n, f_n \rangle \end{bmatrix} \begin{bmatrix} d^{\,2} \\ a_1 \\ \vdots \\ a_n \end{bmatrix} \;=\; \begin{bmatrix} \langle g, g \rangle \\ \langle f_1, g \rangle \\ \vdots \\ \langle f_n, g \rangle \end{bmatrix}.$$

Solving for $d^{\,2}$ using Cramer's rule gives the desired result; expanding along the first column shows that the matrix of coefficients has determinant $G(f_1, \ldots, f_n)$, while the matrix obtained by replacing the "d column" by the right-hand side has determinant $G(g, f_1, \ldots, f_n)$. $\square$

Note: By Lemma 10.3 and induction, every Gram determinant is positive!

In what follows, we will continue to use X_n to denote the span of $x^{\lambda_0}, \ldots, x^{\lambda_n}$, but we now write $\operatorname{dist}_2(f, X_n)$ to denote the distance from f to X_n in the L_2 norm.

Theorem 10.4. *Let m, $\lambda_k > -1/2$ for $k = 0, 1, 2, \ldots$. Then*

$$\operatorname{dist}_2(x^m, X_n) \;=\; \frac{1}{\sqrt{2m+1}} \prod_{k=0}^{n} \frac{|m - \lambda_k|}{m + \lambda_k + 1}.$$

Proof. The proof is based on a determinant formula due to Cauchy:

$$\prod_{i,j}(a_i + b_j) \begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_n} \\ \vdots & \ddots & \vdots \\ \frac{1}{a_n+b_1} & \cdots & \frac{1}{a_n+b_n} \end{vmatrix} \;=\; \prod_{i>j}(a_i - a_j)(b_i - b_j).$$

If we consider each of the a_i and b_j as variables, then each side of the equation is a *polynomial* in $a_1, \ldots, a_n, b_1, \ldots, b_n$. (Why?) Now the right-hand side clearly vanishes if $a_i = a_j$ or $b_i = b_j$ for some $i \neq j$, but the left-hand side also vanishes in any of these cases. Thus, the right-hand side divides the left-hand side. But both polynomials have degree $n - 1$ in each of the a_i and b_j. (Why?) Thus, the left-hand side is a constant multiple of the right-hand side. To show that the constant must be 1, write the left-hand side as

$$\prod_{i \neq j}(a_i + b_j) \begin{vmatrix} 1 & \frac{a_1+b_1}{a_1+b_2} & \cdots & \frac{a_1+b_1}{a_1+b_n} \\ \frac{a_2+b_2}{a_2+b_1} & 1 & \cdots & \frac{a_2+b_2}{a_2+b_n} \\ \vdots & & \ddots & \vdots \\ \frac{a_n+b_n}{a_n+b_1} & \cdots & \frac{a_n+b_n}{a_n+b_{n-1}} & 1 \end{vmatrix}$$

and now take the limit as $b_1 \to -a_1$, $b_2 \to -a_2$, etc. The expression above tends to $\prod_{i \neq j}(a_i - a_j)$, as does the right-hand side of Cauchy's formula.

Now, $\langle\, x^p, x^q \,\rangle = \int_0^1 x^{p+q}\, dx = \frac{1}{p+q+1}$ for p, $q > -1/2$, so

$$G(x^{\lambda_0}, \ldots, x^{\lambda_n}) \;=\; \det\left(\left[\frac{1}{\lambda_i + \lambda_j + 1}\right]_{i,j}\right) \;=\; \frac{\prod_{i>j}(\lambda_i - \lambda_j)^2}{\prod_{i,j}(\lambda_i + \lambda_j + 1)},$$

with a similar formula holding for $G(x^m, x^{\lambda_0}, \ldots, x^{\lambda_n})$. Substituting these expressions into our distance formula and taking square roots finishes the proof. $\square$

Now we can determine exactly when X is dense in $L_2[\,0,1\,]$. For easier comparison to the $C[\,0,1\,]$ case, we will suppose that the λ_n are nonnegative.

Theorem 10.5. *Let $0 \le \lambda_0 < \lambda_1 < \lambda_2 < \cdots$. Then the functions (x^{λ_n}) have dense linear span in $L_2[\,0,1\,]$ if and only if $\sum_{n=1}^{\infty} \lambda_n^{-1} = \infty$.*

Proof. If $\sum_{n=1}^{\infty} \frac{1}{\lambda_n} < \infty$, then each of the products $\prod_{k=1}^{n} \left|1 - \frac{m}{\lambda_k}\right|$ and $\prod_{k=1}^{n} \left|1 + \frac{(m+1)}{\lambda_k}\right|$ converges to some nonzero limit for any m not equal to any λ_k. Thus, $\operatorname{dist}_2(x^m, X_n) \not\to 0$, as $n \to \infty$, for any $m \neq \lambda_k$, $k = 0, 1, 2, \ldots$. In particular, the functions (x^{λ_n}) cannot have dense linear span in $L_2[\,0,1\,]$.

Conversely, if $\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \infty$, then $\prod_{k=1}^{n} \left|1 - \frac{m}{\lambda_k}\right|$ diverges to 0 while $\prod_{k=1}^{n} \left|1 + \frac{(m+1)}{\lambda_k}\right|$ diverges to $+\infty$. Thus, $\operatorname{dist}_2(x^m, X_n) \to 0$, as $n \to \infty$, for every $m > -1/2$. Because the polynomials are dense in $L_2[\,0,1\,]$, this finishes the proof. $\square$

Finally, we can finish the proof of Müntz's theorem in the case of $C[\,0,1\,]$. Suppose that the functions (x^{λ_n}) have dense linear span in $C[\,0,1\,]$. Then, because $\|f\|_2 \le \|f\|$, it follows that the functions (x^{λ_n}) must also have dense linear span in $L_2[\,0,1\,]$. (Why?) Hence, $\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \infty$. $\square$

Just for good measure, here's a second proof of the "backward" direction for $C[\,0,1\,]$ based on the $L_2[\,0,1\,]$ version. Suppose that $\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \infty$, and let $m \ge 1$. Then,

$$\begin{aligned} \left| x^m - \sum_{k=0}^{n} a_k x^{\lambda_k} \right| &= \left| \frac{1}{m} \int_0^x t^{m-1}\,dt - \sum_{k=0}^{n} \frac{a_k}{\lambda_k} \int_0^x t^{\lambda_k - 1}\,dt \right| \\ &\le \int_0^1 \left| \frac{1}{m}\, t^{m-1} - \sum_{k=0}^{n} \frac{a_k}{\lambda_k}\, t^{\lambda_k - 1} \right| dt \\ &\le \left(\int_0^1 \left| \frac{1}{m}\, t^{m-1} - \sum_{k=0}^{n} \frac{a_k}{\lambda_k}\, t^{\lambda_k - 1} \right|^2 dt \right)^{1/2}. \end{aligned}$$

Now because $\sum_{\lambda_n > 1} \frac{1}{\lambda_n - 1} = \infty$ the functions $(x^{\lambda_k - 1})$ have dense linear span in $L_2[\,0,1\,]$. Thus, we can find a_k so that the right-hand side of this inequality is less than some ε. Because this estimate is independent of x, we've shown that

$$\max_{0 \le x \le 1} \left| x^m - \sum_{k=0}^{n} a_k x^{\lambda_k} \right| < \varepsilon. \qquad \square$$

Corollary 10.6. *Let* $0 = \lambda_0 < \lambda_1 < \lambda_2 < \cdots$ *with* $\sum_{n=1}^{\infty} \lambda_n^{-1} = \infty$, *and let* f *be a continuous function on* $[\,0,\infty)$ *for which* $c = \lim_{t\to\infty} f(t)$ *exists. Then,* f *can be uniformly approximated by finite linear combinations of the exponentials* $(e^{-\lambda_n t})_{n=0}^{\infty}$.

Proof. The function defined by $g(x) = f(-\log x)$, for $0 < x \le 1$, and $g(0) = c$, is continuous on $[\,0,1\,]$. Obviously, $g(e^{-t}) = f(t)$ for each $0 \le t < \infty$. Thus, given $\varepsilon > 0$, we can find n and $a_0, \ldots, a_n$ such that

$$\max_{0 \le x \le 1} \left| g(x) - \sum_{k=0}^{n} a_k x^{\lambda_k} \right| = \max_{0 \le t < \infty} \left| f(t) - \sum_{k=0}^{n} a_k e^{-\lambda_k t} \right| < \varepsilon. \qquad \square$$

Chapter 11

The Stone-Weierstrass Theorem

To begin, an *algebra* is a vector space A on which there is a multiplication $(f, g) \mapsto fg$ (from $A \times A$ into A) satisfying

(i) $(fg)h = f(gh)$, for all $f,\ g,\ h \in A$;

(ii) $f(g + h) = fg + fh$ and $(f + g)h = fg + gh$, for all $f,\ g,\ h \in A$;

(iii) $\alpha(fg) = (\alpha f)g = f(\alpha g)$, for all scalars α and all $f,\ g \in A$.

In other words, an algebra is a *ring* under vector addition and multiplication, together with a compatible scalar multiplication. The algebra is *commutative* if

(iv) $fg = gf$, for all $f,\ g \in A$.

And we say that A has an *identity* element if there is a vector $e \in A$ such that

(v) $fe = ef = f$, for all $f \in A$.

In case A is a normed vector space, we also require that the norm satisfy

(vi) $\|fg\| \le \|f\|\,\|g\|$

(this simplifies things a bit), and in this case we refer to A as a *normed algebra.* If a normed algebra is complete, we refer to it as a *Banach algebra.* Finally, a subset B of an algebra A is called a *subalgebra* of A if B is itself an algebra (under the operations it inherits from A); that is, if B is a (vector) subspace of A which is closed under multiplication.

If A is a normed algebra, then all of the various operations on A (or $A \times A$) are continuous. For example, because

$$\|fg - hk\| \;=\; \|fg - fk + fk - hk\| \;\le\; \|f\|\,\|g - k\| + \|k\|\,\|f - h\|$$

it follows that multiplication is continuous. (How?) In particular, if B is a subspace (or subalgebra) of A, then $\overline{B}$, the closure of B, is also a subspace (or subalgebra) of A.

Examples 11.1.

1. If we define multiplication of vectors "coordinatewise," then $\mathbb{R}^n$ is a commutative Banach algebra with identity (the vector $(1, \ldots, 1)$) when equipped with the norm $\|x\|_\infty = \max_{1 \le i \le n} |x_i|$.

2. It's not hard to identify the subalgebras of $\mathbb{R}^n$ among its subspaces. For example, the subalgebras of $\mathbb{R}^2$ are $\{(x,0) : x \in \mathbb{R}\}$, $\{(0,y) : y \in \mathbb{R}\}$, and $\{(x,x) : x \in \mathbb{R}\}$, along with $\{(0,0)\}$ and $\mathbb{R}^2$.

3. Given a set X, we write $B(X)$ for the space of all bounded, real-valued functions on X. If we endow $B(X)$ with the sup norm, and if we define arithmetic with functions pointwise, then $B(X)$ is a commutative Banach algebra with identity (the constant 1 function). The constant functions in $B(X)$ form a subalgebra isomorphic (in every sense of the word) to $\mathbb{R}$.

4. If X is a metric (or topological) space, then we may consider $C(X)$, the space of all continuous, real-valued functions on X. If we again define arithmetic with functions pointwise, then $C(X)$ is a commutative algebra with identity (the constant 1 function). The bounded, continuous functions on X, written $C_b(X) = C(X) \cap B(X)$, form a closed subalgebra of $B(X)$. If X is compact, then $C_b(X) = C(X)$. In other words, if X is compact, then $C(X)$ is itself a closed subalgebra of $B(X)$ and, in particular, $C(X)$ is a Banach algebra with identity.

5. The polynomials form a dense subalgebra of $C[\,a,b\,]$. The trig polynomials form a dense subalgebra of $C^{2\pi}$. These two sentences summarize Weierstrass's two classical theorems in modern parlance and form the basis for Stone's version of the theorem.

Using this new language, we may restate the classical Weierstrass theorem to read: *If a subalgebra A of $C[\,a,b\,]$ contains the functions $e(x) = 1$ and $f(x) = x$, then A is dense in $C[\,a,b\,]$.* Of course, any subalgebra of $C[\,a,b\,]$ containing 1 and x actually contains all the polynomials; thus, our restatement of Weierstrass's theorem amounts to the observation that any subalgebra containing a dense set is itself dense in $C[\,a,b\,]$.

Our goal in this section is to prove an analogue of this new version of the Weierstrass theorem for subalgebras of $C(X)$, where X is a compact metric space. In particular, we will want to extract the essence of the functions 1 and x from this statement. That is, we seek conditions on a subalgebra A of $C(X)$ that will force A to be dense in $C(X)$. The key role played by 1 and x, in the case of $C[\,a,b\,]$, is that a subalgebra containing these two functions must actually contain a much larger set of functions. But because we can't be assured of anything remotely like polynomials living in the more general $C(X)$ spaces, we might want to change our point of view. What we really need is some requirement on a subalgebra A of $C(X)$ that will allow us to *construct* a wide variety of functions in A. And, if A contains a sufficiently rich variety of functions, it might just be possible to show that A is dense.

Because the two replacement conditions we have in mind make sense in any collection of real-valued functions, we state them in some generality.

Let A be a collection of real-valued functions on some set X. We say that A *separates points in* X if, given $x \neq y \in X$, there is some $f \in A$ such that $f(x) \neq f(y)$. We say that A *vanishes at no point of* X if, given $x \in X$, there is some $f \in A$ such that $f(x) \neq 0$.

Examples 11.2.

1. The single function $f(x) = x$ clearly separates points in $[\,a,b\,]$, and the function $e(x) = 1$ obviously vanishes at no point in $[\,a,b\,]$. Any subalgebra A of $C[\,a,b\,]$ containing these two functions will likewise separate points and vanish at no point in $[\,a,b\,]$.

2. The set E of even functions in $C[-1,1\,]$ fails to separate points in $[-1,1\,]$; indeed, $f(x) = f(-x)$ for any even function. However, because the constant functions are

even, E vanishes at no point of $[-1, 1]$. It's not hard to see that E is a proper closed subalgebra of $C[-1, 1]$. The set of odd functions will separate points (because $f(x) = x$ is odd), but the odd functions all vanish at 0. The set of odd functions is a proper closed subspace of $C[-1, 1]$, although not a subalgebra.

3. The set of all functions $f \in C[-1, 1]$ for which $f(0) = 0$ is a proper closed subalgebra of $C[-1, 1]$. In fact, this set is a maximal (in the sense of containment) proper closed subalgebra of $C[-1, 1]$. Note, however, that this set of functions does separate points in $[-1, 1]$ (again, because it contains $f(x) = x$).

4. It's easy to construct examples of non-trivial closed subalgebras of $C(X)$. Indeed, given any closed subset X_0 of X, the set $A(X_0) = \{f \in C(X) : f \text{ vanishes on } X_0\}$ is a non-empty, proper subalgebra of $C(X)$. It's closed in any reasonable topology on $C(X)$ because it's closed under pointwise limits. Subalgebras of the type $A(X_0)$ are of interest because they're actually *ideals* in the ring $C(X)$. That is, if $f \in C(X)$, and if $g \in A(X_0)$, then $fg \in A(X_0)$.

As these few examples illustrate, neither of our new conditions, taken separately, is enough to force a subalgebra of $C(X)$ to be dense. But both conditions together turn out to be sufficient. In order to better appreciate the utility of these new conditions, let's isolate the key computational tool that they permit within an algebra of functions.

Lemma 11.3. *Let A be an algebra of real-valued functions on some set X, and suppose that A separates points in X and vanishes at no point of X. Then, given $x \neq y \in X$ and a, $b \in \mathbb{R}$, we can find an $f \in A$ with $f(x) = a$ and $f(y) = b$.*

Proof. Given any pair of distinct points $x \neq y \in X$, the set $\widetilde{A} = \{\big(f(x), f(y)\big) : f \in A\}$ is a subalgebra of $\mathbb{R}^2$. If A separates points in X, then $\widetilde{A}$ is evidently neither $\{(0, 0)\}$ nor $\{(x, x) : x \in \mathbb{R}\}$. If A vanishes at no point, then $\{(x, 0) : x \in \mathbb{R}\}$ and $\{(0, y) : y \in \mathbb{R}\}$ are both excluded. Thus $\widetilde{A} = \mathbb{R}^2$. That is, for any a, $b \in \mathbb{R}$, there is some $f \in A$ for which $(f(x), f(y)) = (a, b)$. □

Now we can state Stone's version of the Weierstrass theorem (for compact metric spaces). It should be pointed out that the theorem, as stated, also holds in $C(X)$ when X is a compact Hausdorff topological space (with the same proof), but does not hold for algebras of complex-valued functions over $\mathbb{C}$. More on this later.

Theorem 11.4. (Stone-Weierstrass Theorem, real scalars) *Let X be a compact metric space, and let A be a subalgebra of $C(X)$. If A separates points in X and vanishes at no point of X, then A is dense in $C(X)$.*

What Cheney calls an "embryonic" version of this theorem appeared in 1937, as a small part of a massive 106 page paper! Later versions, appearing in 1948 and 1962, benefitted from the work of the great Japanese mathematician Kakutani and were somewhat more palatable to the general mathematical public. But, no matter which version you consult, you'll find them difficult to read. For more details, I would recommend you first consult Rudin [46], Folland [17], Simmons [51], or my book on real analysis [10].

As a first step in attacking the proof of Stone's theorem, notice that if A satisfies the conditions of the theorem, then so does its closure $\overline{A}$. (Why?) Thus, we may assume that A is actually a *closed* subalgebra of $C(X)$ and prove, instead, that $A = C(X)$. Now the closed subalgebras of $C(X)$ inherit more structure than you might first imagine.

Theorem 11.5. *If A is a subalgebra of $C(X)$, and if $f \in A$, then $|f| \in \overline{A}$. Consequently, $\overline{A}$ is a sublattice of $C(X)$.*

Proof. Let $\varepsilon > 0$, and consider the function $|t|$ on the interval $\big[-\|f\|, \|f\|\,\big]$. By the Weierstrass theorem, there is a polynomial $p(t) = \sum_{k=0}^{n} a_k t^k$ such that $\big|\,|t| - p(t)\,\big| < \varepsilon$ for all $|t| \le \|f\|$. In particular, notice that $|p(0)| = |a_0| < \varepsilon$.

Now, because $|f(x)| \le \|f\|$ for all $x \in X$, it follows that $\big|\,|f(x)| - p(f(x))\,\big| < \varepsilon$ for all $x \in X$. But $p(f(x)) = (p(f))(x)$, where $p(f) = a_0\mathbf{1} + a_1 f + \cdots + a_n f^n$, and the function $g = a_1 f + \cdots + a_n f^n \in A$, because A is an algebra. Thus, $\big|\,|f(x)| - g(x)\,\big| \le |a_0| + \varepsilon < 2\varepsilon$ for all $x \in X$. In other words, for each $\varepsilon > 0$, we can supply an element $g \in A$ such that $\|\,|f| - g\| < 2\varepsilon$. That is, $|f| \in \overline{A}$.

The statement that $\overline{A}$ is a sublattice of $C(X)$ means that if we're given $f,\ g \in \overline{A}$, then $\max\{f, g\} \in \overline{A}$ and $\min\{f, g\} \in \overline{A}$, too. But this is actually just a statement about real numbers. Indeed, because

$$2\max\{a, b\} = a + b + |a - b| \qquad \text{and} \qquad 2\min\{a, b\} = a + b - |a - b|$$

it follows that a *subspace* of $C(X)$ is a sublattice precisely when it contains the absolute values of all its elements. □

The point to our last result is that if we're given a closed subalgebra A of $C(X)$, then A is "closed" in every sense of the word: Sums, products, absolute values, max's, and min's of elements from A, and even limits of sequences of these, are all back in A. This is precisely the sort of freedom we'll need if we hope to show that $A = C(X)$.

Please notice that we could have avoided our appeal to the Weierstrass theorem in this last result. Indeed, we only needed to supply polynomial approximations for the single function $|x|$ on $[-1, 1\,]$, and this can be done directly. For example, we could appeal instead to the binomial theorem, using $|x| = \sqrt{1 - (1 - x^2)}$; the resulting series can be shown to converge uniformly on $[-1, 1\,]$. (But see also Problem 9 from Chapter 2.) By side-stepping the classical Weierstrass theorem, it becomes a corollary to Stone's version (rather than the other way around).

Now we're ready for the proof of the Stone-Weierstrass theorem. As we've already pointed out, we may assume that we're given a closed subalgebra (subspace, and sublattice) A of $C(X)$ and we want to show that $A = C(X)$. We'll break the remainder of the proof into two steps:

Step 1: Given $f \in C(X)$, $x \in X$, and $\varepsilon > 0$, there is an element $g_x \in A$ with $g_x(x) = f(x)$ and $g_x(y) > f(y) - \varepsilon$ for all $y \in X$.

From Lemma 11.3, we know that for each $y \in X$, $y \ne x$, we can find an $h_y \in A$ so that $h_y(x) = f(x)$ and $h_y(y) = f(y)$. Because $h_y - f$ is continuous and vanishes at both x and y, the set $U_y = \{t \in X : h_y(t) > f(t) - \varepsilon\}$ is open and contains both x and y. Thus, the sets $(U_y)_{y \ne x}$ form an open cover for X. Because X is compact, finitely many U_y suffice, say $X = U_{y_1} \cup \cdots \cup U_{y_n}$. Now set $g_x = \max\{h_{y_1}, \ldots, h_{y_n}\}$. Because A is a lattice, we have $g_x \in A$. Note that $g_x(x) = f(x)$ because each h_{y_i} agrees with f at x. And $g_x > f - \varepsilon$ because, given $y \ne x$, we have $y \in U_{y_i}$ for some i, and hence $g_x(y) \ge h_{y_i}(y) > f(y) - \varepsilon$.

Step 2: Given $f \in C(X)$ and $\varepsilon > 0$, there is an $h \in A$ with $\|f - h\| < \varepsilon$.

From Step 1, for each $x \in X$ we can find some $g_x \in A$ such that $g_x(x) = f(x)$ and $g_x(y) > f(y) - \varepsilon$ for all $y \in X$. And now we reverse the process used in Step 1: For each x, the set $V_x = \{y \in X : g_x(y) < f(y) + \varepsilon\}$ is open and contains x. Again, because X is compact, $X = V_{x_1} \cup \cdots V_{x_m}$ for some $x_1, \ldots, x_m$. This time, set $h = \min\{g_{x_1}, \ldots, g_{x_m}\} \in A$. As before, $h(y) > f(y) - \varepsilon$ for all y, because each g_{x_i} does so, and $h(y) < f(y) + \varepsilon$ for all y, because at least one g_{x_i} does so.

The conclusion of Step 2 is that A is dense in $C(X)$; but, because A is closed, this means that $A = C(X)$. $\square$

Corollary 11.6. *If X and Y are compact metric spaces, then the subspace of $C(X \times Y)$ spanned by the functions of the form $f(x, y) = g(x)\,h(y)$, $g \in C(X)$, $h \in C(Y)$, is dense in $C(X \times Y)$.*

Corollary 11.7. *If K is a compact subset of $\mathbb{R}^n$, then the polynomials (in n-variables) are dense in $C(K)$.*

Applications to $C^{2\pi}$

In many texts, the Stone-Weierstrass theorem is used to show that the trig polynomials are dense in $C^{2\pi}$. One approach here might be to identify $C^{2\pi}$ with the closed subalgebra of $C[\,0, 2\pi\,]$ consisting of those functions f satisfying $f(0) = f(2\pi)$. Probably easier, though, is to identify $C^{2\pi}$ with the continuous functions on the unit circle $\mathbb{T} = \{e^{i\theta} : \theta \in \mathbb{R}\} = \{z \in \mathbb{C} : |z| = 1\}$ in the complex plane using the identification

$$f \in C^{2\pi} \quad \longleftrightarrow \quad g \in C(\mathbb{T}), \quad \text{where} \quad g(e^{it}) = f(t).$$

Under this correspondence, the trig polynomials in $C^{2\pi}$ match up with (certain) polynomials in $z = e^{it}$ and $\overline{z} = e^{-it}$. But, as we've seen, even if we start with real-valued trig polynomials, we'll end up with polynomials in z and $\overline{z}$ having complex coefficients.

Given this, it might make more sense to consider the complex-valued continuous functions on $\mathbb{T}$. We'll write $C_{\mathbb{C}}(\mathbb{T})$ to denote the complex-valued continuous functions on $\mathbb{T}$, and $C_{\mathbb{R}}(\mathbb{T})$ to denote the real-valued continuous functions on $\mathbb{T}$. Similarly, $C^{2\pi}_{\mathbb{C}}$ is the space of complex-valued, 2π-periodic functions on $\mathbb{R}$, while $C^{2\pi}_{\mathbb{R}}$ stands for the real-valued, 2π-periodic functions on $\mathbb{R}$. Now, under the identification we made earlier, we have $C_{\mathbb{C}}(\mathbb{T}) = C^{2\pi}_{\mathbb{C}}$ and $C_{\mathbb{R}}(\mathbb{T}) = C^{2\pi}_{\mathbb{R}}$. The complex-valued trig polynomials in $C^{2\pi}_{\mathbb{C}}$ now match up with the full set of polynomials, with complex coefficients, in $z = e^{it}$ and $\overline{z} = e^{-it}$. We'll use the Stone-Weierstrass theorem to show that these polynomials are dense in $C_{\mathbb{C}}(\mathbb{T})$.

Now the polynomials in z obviously separate points in $\mathbb{T}$ and vanish at no point of $\mathbb{T}$. Nevertheless, the polynomials in z alone are not dense in $C_{\mathbb{C}}(\mathbb{T})$. To see this, here's a proof that $f(z) = \overline{z}$ cannot be uniformly approximated by polynomials in z. First, suppose that we're given some polynomial $p(z) = \sum_{k=0}^n c_k z^k$. Then

$$\int_0^{2\pi} \overline{f(e^{it})}\, p(e^{it})\, dt \;=\; \int_0^{2\pi} e^{it}\, p(e^{it})\, dt \;=\; \sum_{k=0}^n c_k \int_0^{2\pi} e^{i(k+1)t}\, dt \;=\; 0,$$

and so

$$2\pi \;=\; \int_0^{2\pi} \overline{f(e^{it})}\, f(e^{it})\, dt \;=\; \int_0^{2\pi} \overline{f(e^{it})}\, \big[f(e^{it}) - p(e^{it})\big]\, dt,$$

because $\overline{f(z)}\, f(z) = |f(z)|^2 = 1$. Now, taking absolute values, we get

$$2\pi \;\le\; \int_0^{2\pi} \big|f(e^{it}) - p(e^{it})\big|\, dt \;\le\; 2\pi \|f - p\|.$$

That is, $\|f - p\| \geq 1$ for any polynomial p.

We might as well proceed in some generality: Given a compact metric space X, we'll write $C_{\mathbb{C}}(X)$ for the set of all continuous, complex-valued functions $f : X \to \mathbb{C}$, and we norm $C_{\mathbb{C}}(X)$ by $\|f\| = \max_{x \in X} |f(x)|$ (where $|f(x)|$ is the modulus of the complex number $f(x)$, of course). $C_{\mathbb{C}}(X)$ is a Banach algebra over $\mathbb{C}$. In order to make it clear which field of scalars is involved, we'll write $C_{\mathbb{R}}(X)$ for the real-valued members of $C_{\mathbb{C}}(X)$. Notice, though, that $C_{\mathbb{R}}(X)$ is nothing more than $C(X)$ with a new name.

More generally, we'll write $A_{\mathbb{C}}$ to denote an algebra, over $\mathbb{C}$, of complex-valued functions and $A_{\mathbb{R}}$ to denote the real-valued members of $A_{\mathbb{C}}$. It's not hard to see that $A_{\mathbb{R}}$ is then an algebra, over $\mathbb{R}$, of real-valued functions.

Now if f is in $C_{\mathbb{C}}(X)$, then so is the function $\overline{f}(x) = \overline{f(x)}$ (the complex-conjugate of $f(x)$). This puts

$$\mathrm{Re} f = \frac{1}{2}\big(f + \overline{f}\,\big) \quad \text{and} \quad \mathrm{Im} f = \frac{1}{2i}\big(f - \overline{f}\,\big),$$

the real and imaginary parts of f, in $C_{\mathbb{R}}(X)$ too. Conversely, if $g,\ h \in C_{\mathbb{R}}(X)$, then $g + ih \in C_{\mathbb{C}}(X)$.

This simple observation gives us a hint as to how we might apply the Stone-Weierstrass theorem to subalgebras of $C_{\mathbb{C}}(X)$. Given a subalgebra $A_{\mathbb{C}}$ of $C_{\mathbb{C}}(X)$, suppose that we could prove that $A_{\mathbb{R}}$ is dense in $C_{\mathbb{R}}(X)$. Then, given any $f \in C_{\mathbb{C}}(X)$, we could approximate $\mathrm{Re} f$ and $\mathrm{Im} f$ by elements $g,\ h \in A_{\mathbb{R}}$. But because $A_{\mathbb{R}} \subset A_{\mathbb{C}}$, this means that $g + ih \in A_{\mathbb{C}}$, and $g + ih$ approximates f. That is, $A_{\mathbb{C}}$ is dense in $C_{\mathbb{C}}(X)$. Great! And what did we really use here? Well, we need $A_{\mathbb{R}}$ to contain the real and imaginary parts of "most" functions in $C_{\mathbb{C}}(X)$. If we insist that $A_{\mathbb{C}}$ separate points and vanish at no point, then $A_{\mathbb{R}}$ will contain "most" of $C_{\mathbb{R}}(X)$. And, to be sure that we get both the real and imaginary parts of each element of $A_{\mathbb{C}}$, we'll insist that $A_{\mathbb{C}}$ contain the conjugates of each of its members: $\overline{f} \in A_{\mathbb{C}}$ whenever $f \in A_{\mathbb{C}}$. That is, we'll require that $A_{\mathbb{C}}$ be *self-conjugate* (or, as some authors say, *self-adjoint*).

Theorem 11.8. (Stone-Weierstrass Theorem, complex scalars) *Let X be a compact metric space, and let $A_{\mathbb{C}}$ be a subalgebra, over $\mathbb{C}$, of $C_{\mathbb{C}}(X)$. If $A_{\mathbb{C}}$ separates points in X, vanishes at no point of X, and is self-conjugate, then $A_{\mathbb{C}}$ is dense in $C_{\mathbb{C}}(X)$.*

Proof. Again, write $A_{\mathbb{R}}$ for the set of real-valued members of $A_{\mathbb{C}}$. Because $A_{\mathbb{C}}$ is self-conjugate, $A_{\mathbb{R}}$ contains the real and imaginary parts of every $f \in A_{\mathbb{C}}$;

$$\mathrm{Re} f = \frac{1}{2}\big(f + \overline{f}\,\big) \in A_{\mathbb{R}} \quad \text{and} \quad \mathrm{Im} f = \frac{1}{2i}\big(f - \overline{f}\,\big) \in A_{\mathbb{R}}.$$

Moreover, $A_{\mathbb{R}}$ is a subalgebra, over $\mathbb{R}$, of $C_{\mathbb{R}}(X)$. In addition, $A_{\mathbb{R}}$ separates points in X and vanishes at no point of X. Indeed, given $x \neq y \in X$ and $f \in A_{\mathbb{C}}$ with $f(x) \neq f(y)$, we must have at least one of $\mathrm{Re} f(x) \neq \mathrm{Re} f(y)$ or $\mathrm{Im} f(x) \neq \mathrm{Im} f(y)$. Similarly, $f(x) \neq 0$ means that at least one of $\mathrm{Re} f(x) \neq 0$ or $\mathrm{Im} f(x) \neq 0$ holds. That is, $A_{\mathbb{R}}$ satisfies the hypotheses of the real-scalar version of the Stone-Weierstrass theorem. Consequently, $A_{\mathbb{R}}$ is dense in $C_{\mathbb{R}}(X)$.

Now, given $f \in C_{\mathbb{C}}(X)$ and $\varepsilon > 0$, take $g,\ h \in A_{\mathbb{R}}$ with $\|g - \mathrm{Re} f\| < \varepsilon/2$ and $\|h - \mathrm{Im} f\| < \varepsilon/2$. Then, $g + ih \in A_{\mathbb{C}}$ and $\|f - (g + ih)\| < \varepsilon$. Thus, $A_{\mathbb{C}}$ is dense in $C_{\mathbb{C}}(X)$. $\square$

Corollary 11.9. *The polynomials, with complex coefficients, in z and $\overline{z}$ are dense in $C_{\mathbb{C}}(\mathbb{T})$. In other words, the complex trig polynomials are dense in $C_{\mathbb{C}}^{2\pi}$.*

Note that it follows from the complex-scalar proof that the real parts of the polynomials in z and $\overline{z}$, that is, the real trig polynomials, are dense in $C_{\mathbb{R}}(\mathbb{T}) = C_{\mathbb{R}}^{2\pi}$.

Corollary 11.10. *The real trig polynomials are dense in* $C^{2\pi}_{\mathbb{R}}$.

Applications to Lipschitz Functions

In most modern Real Analysis courses, the classical Weierstrass theorem is used to prove that $C[\,a,b\,]$ is separable. Likewise, the Stone-Weierstrass theorem can be used to show that $C(X)$ is separable, where X is a compact metric space. While we won't have anything quite so convenient as polynomials at our disposal, we do, at least, have a familiar collection of functions to work with.

Given a metric space $(X,d\,)$, and $0 \le K < \infty$, we'll write $\mathrm{lip}_K(X)$ to denote the collection of all real-valued Lipschitz functions on X with constant at most K; that is, $f : X \to \mathbb{R}$ is in $\mathrm{lip}_K(X)$ if $|f(x) - f(y)| \le K d(x,y)$ for all x, $y \in X$. And we'll write $\mathrm{lip}(X)$ to denote the set of functions that are in $\mathrm{lip}_K(X)$ for some K; in other words, $\mathrm{lip}(X) = \bigcup_{K=1}^{\infty} \mathrm{lip}_K(X)$. It's easy to see that $\mathrm{lip}(X)$ is a subspace of $C(X)$; in fact, if X is compact, then $\mathrm{lip}(X)$ is even a subalgebra of $C(X)$. Indeed, given $f \in \mathrm{lip}_K(X)$ and $g \in \mathrm{lip}_M(X)$, we have

$$\begin{aligned} |f(x)g(x) - f(y)g(y)| &\le |f(x)g(x) - f(y)g(x)| + |f(y)g(x) - f(y)g(y)| \\ &\le K\|g\|\,|x-y| + M\|f\|\,|x-y|. \end{aligned}$$

Lemma 11.11. *If X is a compact metric space, then* $\mathrm{lip}(X)$ *is dense in* $C(X)$.

Proof. Clearly, $\mathrm{lip}(X)$ contains the constant functions and so vanishes at no point of X. To see that $\mathrm{lip}(X)$ separates point in X, we use the fact that the metric d is Lipschitz: Given $x_0 \ne y_0 \in X$, the function $f(x) = d(x,y_0)$ satisfies $f(x_0) > 0 = f(y_0)$; moreover, $f \in \mathrm{lip}_1(X)$ because

$$|f(x) - f(y)| \;=\; |d(x,y_0) - d(y,y_0)| \;\le\; d(x,y).$$

Thus, by the Stone-Weierstrass Theorem, $\mathrm{lip}(X)$ is dense in $C(X)$. □

Theorem 11.12. *If X is a compact metric space, then $C(X)$ is separable.*

Proof. It suffices to show that $\mathrm{lip}(X)$ is separable. (Why?) To see this, first notice that $\mathrm{lip}(X) = \bigcup_{K=1}^{\infty} E_K$, where

$$E_K = \{f \in C(X) : \|f\| \le K \text{ and } f \in \mathrm{lip}_K(X)\}.$$

(Why?) The sets E_K are (uniformly) bounded and equicontinuous. Hence, by the Arzelà-Ascoli theorem, each E_K is compact in $C(X)$. Because compact sets are separable, as are countable unions of compact sets, it follows that $\mathrm{lip}(X)$ is separable. □

As it happens, the converse is also true (which is why this is interesting); see Folland [17] or my book on Banach spaces [11] for more details.

Theorem 11.13. *If $C(X)$ is separable, where X is a compact Hausdorff topological space, then X is metrizable.*

Appendix A

The ℓ_p Norms

For completeness, we supply a few of the missing details concerning the ℓ_p-norms. We begin with a handful of classical inequalities of independent interest. First recall that we have defined a scale of "norms" on $\mathbb{R}^n$ by setting:

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \qquad 1 \le p < \infty,$$

and

$$\|x\|_\infty = \max_{1 \le i \le n} |x_i|,$$

where $x = (x_i)_{i=1}^n \in \mathbb{R}^n$. Please note that the case $p = 2$ gives the usual Euclidean norm on $\mathbb{R}^n$ and that the cases $p = 1$ and $p = \infty$ clearly give rise to legitimate norms on $\mathbb{R}^n$.

Common parlance is to refer to these expressions as ℓ_p-norms and to refer to the space $(\mathbb{R}^n, \|\cdot\|_p)$ as ℓ_p^n. The space of all infinite sequences $x = (x_n)_{n=1}^\infty$ for which the analogous infinite sum (or supremum) $\|x\|_p$ is finite is referred to as ℓ_p. What's more, there is a "continuous" analogue of this scale: We might also consider the norms

$$\|f\|_p = \left(\int_a^b |f(x)|^p \, dx\right)^{1/p}, \qquad 1 \le p < \infty,$$

and

$$\|f\|_\infty = \sup_{a \le x \le b} |f(x)|,$$

where f is in $C[\,a,b\,]$ (or is simply Lebesgue integrable). The subsequent discussion actually covers all of these cases, but we will settle for writing our proofs in the $\mathbb{R}^n$ setting only.

Lemma A.1. (Young's inequality). *Let $1 < p < \infty$, and let $1 < q < \infty$ be defined by $\frac{1}{p} + \frac{1}{q} = 1$; that is, $q = \frac{p}{p-1}$. Then, for any a, $b \ge 0$, we have*

$$ab \le \frac{1}{p}\, a^p + \frac{1}{q}\, b^q.$$

Moreover, equality can only occur if $a^p = b^q$. (We refer to p and q as *conjugate* exponents; note that p satisfies $p = \frac{q}{q-1}$. Please note that the case $p = q = 2$ yields the familiar *arithmetic-geometric mean inequality.*)

Proof. A quick calculation before we begin:

$$q - 1 = \frac{p}{p-1} - 1 = \frac{p-(p-1)}{p-1} = \frac{1}{p-1}.$$

Now we just estimate areas; for this you might find it helpful to draw the graph of $y = x^{p-1}$ (or, equivalently, the graph of $x = y^{q-1}$). Comparing areas we get:

$$ab \le \int_0^a x^{p-1}\,dx + \int_0^b y^{q-1}\,dy = \frac{1}{p}\,a^p + \frac{1}{q}\,b^q.$$

The case for equality also follows easily from the graph of $y = x^{p-1}$ (or $x = y^{q-1}$), because $b = a^{p-1} = a^{p/q}$ means that $a^p = b^q$. $\square$

Corollary A.2. (Hölder's inequality). *Let $1 < p < \infty$, and let $1 < q < \infty$ be defined by $\frac{1}{p} + \frac{1}{q} = 1$. Then, for any $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ in $\mathbb{R}$ we have:*

$$\sum_{i=1}^n |a_i b_i| \le \left(\sum_{i=1}^n |a_i|^p\right)^{1/p} \left(\sum_{i=1}^n |b_i|^q\right)^{1/q}.$$

(Please note that the case $p = q = 2$ yields the familiar *Cauchy-Schwarz inequality.*)

Moreover, equality in Hölder's inequality can only occur if there exist nonnegative scalars α and β such that $\alpha\,|a_i|^p = \beta\,|b_i|^q$ for all $i = 1, \ldots, n$.

Proof. Let $A = \left(\sum_{i=1}^n |a_i|^p\right)^{1/p}$ and let $B = \left(\sum_{i=1}^n |b_i|^q\right)^{1/q}$. We may clearly assume that $A,\, B \ne 0$ (why?), and hence we may divide (and appeal to Young's inequality):

$$\frac{|a_i b_i|}{AB} \le \frac{|a_i|^p}{pA^p} + \frac{|b_i|^q}{qB^q}.$$

Adding, we get:

$$\frac{1}{AB}\sum_{i=1}^n |a_i b_i| \le \frac{1}{pA^p}\sum_{i=1}^n |a_i|^p + \frac{1}{qB^q}\sum_{i=1}^n |b_i|^q = \frac{1}{p} + \frac{1}{q} = 1.$$

That is, $\sum_{i=1}^n |a_i b_i| \le AB$.

The case for equality in Hölder's inequality follows from what we know about Young's inequality: Equality in Hölder's inequality means that either $A = 0$, or $B = 0$, or else $|a_i|^p/pA^p = |b_i|^q/qB^q$ for all $i = 1, \ldots, n$. In short, there must exist nonnegative scalars α and β such that $\alpha\,|a_i|^p = \beta\,|b_i|^q$ for all $i = 1, \ldots, n$. $\square$

Notice, too, that the case $p = 1$ ($q = \infty$) works, and is easy:

$$\sum_{i=1}^n |a_i b_i| \le \left(\sum_{i=1}^n |a_i|\right)\left(\max_{1\le i\le n} |b_i|\right).$$

Exercise A.3. When does equality occur in the case $p = 1$ ($q = \infty$)?

Finally, an application of Hölder's inequality leads to an easy proof that $\|\cdot\|_p$ is actually a norm. It will help matters here if we first make a simple observation: If $1 < p < \infty$ and if $q = \frac{p}{p-1}$, notice that

$$\left\| \left(|a_i|^{p-1} \right)_{i=1}^n \right\|_q = \left(\sum_{i=1}^n |a_i|^p \right)^{(p-1)/p} = \|a\|_p^{p-1}.$$

Lemma A.4. (Minkowski's inequality). *Let $1 < p < \infty$ and let $a = (a_i)_{i=1}^n$, $b = (b_i)_{i=1}^n \in \mathbb{R}^n$. Then, $\|a + b\|_p \leq \|a\|_p + \|b\|_p$.*

Proof. In order to prove the triangle inequality, we once again let q be defined by $\frac{1}{p} + \frac{1}{q} = 1$, and now we use Hölder's inequality to estimate:

$$\begin{aligned} \sum_{i=1}^n |a_i + b_i|^p &= \sum_{i=1}^n |a_i + b_i| \cdot |a_i + b_i|^{p-1} \\ &\leq \sum_{i=1}^n |a_i| \cdot |a_i + b_i|^{p-1} + \sum_{i=1}^n |b_i| \cdot |a_i + b_i|^{p-1} \\ &\leq \|a\|_p \cdot \| \left(|a_i + b_i|^{p-1} \right)_{i=1}^n \|_q + \|y\|_p \cdot \| \left(|a_i + b_i|^{p-1} \right)_{i=1}^n \|_q \\ &= \|a + b\|_p^{p-1} \left(\|a\|_p + \|b\|_p \right). \end{aligned}$$

That is, $\|a + b\|_p^p \leq \|a + b\|_p^{p-1} \left(\|a\|_p + \|b\|_p \right)$, and the triangle inequality follows. □

If $1 < p < \infty$, then equality in Minkowski's inequality can only occur if a and b are parallel; that is, the ℓ_p-norm is *strictly convex* for $1 < p < \infty$. Indeed, if $\|a + b\|_p = \|a\|_p + \|b\|_p$, then either $a = 0$, or $b = 0$, or else a, $b \neq 0$ and we have equality at each stage of our proof. Now equality in the first inequality means that $|a_i + b_i| = |a_i| + |b_i|$, which easily implies that a_i and b_i have the same sign. Next, equality in our application of Hölder's inequality implies that there are nonnegative scalars C and D such that $|a_i|^p = C\,|a_i + b_i|^p$ and $|b_i|^p = D\,|a_i + b_i|^p$ for all $i = 1, \ldots, n$. Thus, $a_i = E\,b_i$ for some scalar E and all $i = 1, \ldots, n$.

Of course, the triangle inequality also holds in either of the cases $p = 1$ or $p = \infty$ (with much simpler proofs).

Exercise A.5. When does equality occur in the triangle inequality in the cases $p = 1$ or $p = \infty$? In particular, show that neither of the norms $\|\cdot\|_1$ or $\|\cdot\|_\infty$ is strictly convex.

Appendix B

Completeness and Compactness

Next, we provide a brief review of completeness and compactness. Such review is doomed to inadequacy; the reader unfamiliar with these concepts would be well served to consult a text on advanced calculus such as [29] or [46].

To begin, we recall that a subset A of normed space X (such as $\mathbb{R}$ or $\mathbb{R}^n$) is said to be *closed* if A is closed under the taking of sequential limits. That is, A is closed if, whenever (a_n) is a sequence from A converging to some point $x \in X$, we always have $x \in A$. It's not hard to see that any closed interval, such as $[\,a, b\,]$ or $[\,a, \infty)$, is, indeed, a closed subset of $\mathbb{R}$ in this sense. There are, however, much more complicated examples of closed sets in $\mathbb{R}$.

A normed space X is said to be *complete* if every Cauchy sequence from X converges (to a point in X). It is a familiar fact from Calculus that $\mathbb{R}$ is complete, as is $\mathbb{R}^n$. In fact, the completeness of $\mathbb{R}$ is often assumed as an axiom (in the form of the least upper bound axiom). There are, however, many examples of normed spaces which are not complete; that is, there are examples of normed spaces in which Cauchy sequences need not converge.

We say that a subset A of a normed space X is complete if every Cauchy sequence from A converges *to a point in A*. Please note here that we require not only that Cauchy sequences from A converge, but also that the limit be back in A. As you might imagine, the completeness of A depends on properties of both A and the containing space X.

First note that a complete subset is necessarily also closed. Indeed, because every convergent sequence is also Cauchy, it follows that a complete subset is closed.

Exercise B.1. If A is a complete subset of a normed space X, show that A is also closed.

If the containing space X is itself complete, then it's easy to tell which of its subsets are complete. Indeed, because every Cauchy sequence in X converges (somewhere), all we need to know is whether the subset is closed.

Exercise B.2. Let A be a subset of a complete normed space X. Show that A is complete if and only if A is a closed subset of X. In particular, please note that every closed subset of $\mathbb{R}$ (or $\mathbb{R}^n$) is complete.

Virtually all of the normed spaces encountered in these notes are complete. In some cases, this fact is very easy to check; in others, it can be a bit of a challenge. A few additional tools could prove useful.

To get us started, here is an easy and often used reduction: In order to prove completeness, it's not necessary to show that every Cauchy sequence converges; rather, it suffices to show that every Cauchy sequence has a convergent subsequence.

Exercise B.3. If (x_n) is a Cauchy sequence and if a subsequence (x_{n_k}) converges to a point $x \in X$, show that (x_n) itself converges to x.

Exercise B.4. Given a Cauchy sequence (x_n) in a normed space X, there exists a subsequence (x_{n_k}) satisfying $\sum_k \|x_{n_{k+1}} - x_{n_k}\| < \infty$. (A sequence with summable increments is occasionally referred to as a *fast* Cauchy sequence.) Conclude that X is complete if and only if every fast Cauchy sequence converges.

Next, here is a very general criterion for checking whether a given norm is complete, due to Banach from around 1922. It's quite useful and often very easy to apply.

Theorem B.5. *A normed linear space $(X, \|\cdot\|)$ is complete if and only if every absolutely summable series is summable; that is, if and only if, given a sequence in X for which $\sum_{n=1}^{\infty} \|x_n\| < \infty$, there exists an element $x \in X$ such that $\left\| x - \sum_{n=1}^{N} x_n \right\| \to 0$ as $N \to \infty$.*

Proof. One direction is easy. Suppose that X is complete and that $\sum_{n=1}^{\infty} \|x_n\| < \infty$. Let $s_N = \sum_{n=1}^{N} x_n$. Then, (s_N) is a Cauchy sequence. Indeed, given $M < N$, we have

$$\|s_N - s_M\| = \left\| \sum_{n=M+1}^{N} x_n \right\| \le \sum_{n=M+1}^{N} \|x_n\| \to 0$$

as M, $N \to \infty$. (Why?) Thus, (s_N) converges in X (by assumption).

On the other hand, suppose that every absolutely summable series is summable, and let (x_n) be a Cauchy sequence in X. By passing to a subsequence, if necessary, we may suppose that $\|x_n - x_{n+1}\| < 2^{-n}$ for all n. (How?) But then,

$$\sum_{n=1}^{N} (x_n - x_{n+1}) = x_1 - x_{N+1}$$

converges in X because $\sum_{n=1}^{\infty} \|x_n - x_{n+1}\| < \infty$. Thus, (x_n) converges. $\square$

We next recall that a subset A of a normed space X is said to be *compact* if every sequence from A has a subsequence which converges *to a point in A*. Again, because we have insisted that certain limits remain in A, it's not hard to see that compact sets are necessarily also closed.

Exercise B.6. If A is a compact subset of a normed space X, show that A is also closed.

Moreover, because a Cauchy sequence with a convergent subsequence must itself converge, it follows that every compact set is complete.

Exercise B.7. If A is a compact subset of a normed space X, show that A is also complete.

Because the compactness of a subset A has something to do with *every* sequence in A, it's not hard to believe that it is a more stringent property than the others we've considered so far. In particular, it's not hard to see that a compact set must be bounded.

Exercise B.8. If A is a compact subset of a normed space X, show that A is also bounded. [Hint: If not, then A would contain a sequence (a_n) with $\|a_n\| \to \infty$.]

Now it is generally not so easy to describe the compact subsets of a particular normed space X, however, it is quite easy to describe the compact subsets of $\mathbb{R}$ (or $\mathbb{R}^n$). This well-known result goes by many names; we will refer to it as *the Heine-Borel theorem.*

Theorem B.9. *A subset A of $\mathbb{R}$ (or $\mathbb{R}^n$) is compact if and only if A is both closed and bounded.*

Proof. One direction of the proof is easy: As we've already seen, compact sets in $\mathbb{R}$ are necessarily closed and bounded. For the other direction, notice that if A is a bounded subset of $\mathbb{R}$, then it follows from the Bolzano-Weierstrass theorem that every sequence from A has a subsequence which converges in $\mathbb{R}$. If A is also a closed set, then this limit must, in fact, be back in A. Thus, every sequence in A has a subsequence converging to a point in A. □

It is important to note here that the Heine-Borel theorem *does not* typically hold in the more general setting of metric spaces. By way of an example, consider the set $B = \{x \in \ell_1 : \|x\|_1 \le 1\}$ in the normed linear space ℓ_1 (consisting of all absolutely summable sequences). It's easy to see that B is bounded and not too hard to see that B is closed. However, B is not compact. To see this, consider the sequence (of sequences!) defined by

$$e_n = (\overbrace{0, \ldots, 0}^{n-1 \text{ zeros}}, 1, 0, 0, \ldots) \quad n = 1, 2, \ldots$$

(that is, there is a single nonzero entry, a 1, in the n-th coordinate). It's easy to see that $\|e_n\|_1 = 1$ and, hence, $e_n \in B$ for every n. But (e_n) has no convergent subsequence because $\|e_n - e_m\|_1 = 2$ for any $n \neq m$. Thus, B is not compact.

It is also worth noting that there are other characterizations of compactness, including a few that will carry over successfully to even the very general setting of topological spaces. One such characterization is given below, without proof. (It is offered here as a theorem, but it is more typically given as a definition, from which other characterizations are derived.) For further details, see [46], [10], or [51].

Theorem B.10. *A topological space X is compact if and only if the following condition holds: Given any collection $\mathcal{U}$ of open sets in X with the property that $X = \bigcup\{V : V \in \mathcal{U}\}$, there exist finitely many sets $U_1, \ldots, U_n$ in $\mathcal{U}$ such that $X = \bigcup_{k=1}^{n} U_k$.*

Finally, let's speak briefly about of continuous functions defined on compact sets. Specifically, we will consider a continuous function, real-valued function f defined on a compact interval $[\,a, b\,]$ in $\mathbb{R}$, but much of what we have to say will carry over to the more general setting of a continuous function defined on a compact metric space (taking values in another metric space). The key fact we need is that a continuous function defined on a compact set is *bounded.*

Theorem B.11. *Let $f : [\,a, b\,] \to \mathbb{R}$ be continuous. Then f is bounded. Moreover, f attains both its maximum and minimum values on $[\,a, b\,]$; that is, there exist points x_1, x_2 in $[\,a, b\,]$ such that*

$$f(x_1) = \max_{a \le x \le b} f(x) \qquad \textit{and} \qquad f(x_2) = \min_{a \le x \le b} f(x).$$

Proof. If, to the contrary, f is not bounded, then we can find a sequence of points (x_n) in $[\,a, b\,]$ satisfying $|f(x_n)| > n$. It follows that no subsequence of $\big(f(x_n)\big)$ could possibly converge, for convergent sequences are bounded. This leads to a contradiction: By the Heine-Borel theorem, (x_n) must have a convergent subsequence and, by the continuity of f, the corresponding subsequence of $\big(f(x_n)\big)$ would necessarily converge. This contradiction proves that f is bounded.

We next prove that f attains its maximum value (and leave the other case as an exercise). From the first part of the proof, we know that $M \equiv \sup_{a \le x \le b} f(x)$ is a finite, real number. Thus, we can find a sequence of values $\big(f(x_n)\big)$ converging to M. But, by passing to a subsequence, if necessary, we may suppose that (x_n) converges to some point x_1 in $[\,a, b\,]$. Clearly, x_1 satisfies $f(x_1) = M$. $\square$

Appendix C

Pointwise and Uniform Convergence

We next offer a brief review of pointwise and uniform convergence. We begin with an elementary example:

Example C.1.

(a) For each $n = 1, 2, 3, \ldots$, consider the function $f_n(x) = e^x + \frac{x}{n}$ for $x \in \mathbb{R}$. Note that for each (fixed) x the sequence $(f_n(x))_{n=1}^{\infty}$ converges to $f(x) = e^x$ because

$$|f_n(x) - f(x)| = \frac{|x|}{n} \to 0 \quad \text{as} \quad n \to \infty.$$

In this case we say that the sequence of functions (f_n) converges *pointwise* to the function f on $\mathbb{R}$. But notice, too, that the *rate* of convergence depends on x. In particular, in order to get $|f_n(x) - f(x)| < 1/2$ we would need to take $n > 2|x|$. Thus, at $x = 2$, the inequality is satisfied for all $n > 4$, while at $x = 1000$, the inequality is satisfied only for $n > 2000$. In short, the rate of convergence is *not uniform* in x.

(b) Consider the same sequence of functions as above, but now let's suppose that we restrict that values of x to the interval $[-5, 5\,]$. Of course, we still have that $f_n(x) \to f(x)$ for each (fixed) x in $[-5, 5\,]$; in other words, we still have that (f_n) converges pointwise to f on $[-5, 5\,]$. But notice that the rate of convergence is now uniform over x in $[-5, 5\,]$. To see this, just rewrite the initial calculation:

$$|f_n(x) - f(x)| = \frac{|x|}{n} \le \frac{5}{n} \quad \text{for} \quad x \in [-5, 5\,],$$

and notice that the upper bound $5/n$ tends to 0, as $n \to \infty$, *independent of the choice of x*. In this case, we say that (f_n) converges *uniformly* to f on $[-5, 5\,]$. The point here is that the notion of uniform convergence depends on the underlying domain as well as on the sequence of functions at hand.

With this example in mind, we now offer formal definitions of pointwise and uniform convergence. In both cases we consider a sequence of functions $f_n : X \to \mathbb{R}$, $n = 1, 2, 3, \ldots$, each defined on the same underlying set X, and another function $f : X \to \mathbb{R}$ (the candidate for the limit).

We say that (f_n) *converges pointwise to f on X* if, for each $x \in X$, we have $f_n(x) \to f(x)$ as $n \to \infty$; thus, for each $x \in X$ and each $\varepsilon > 0$, we can find an integer N (which depends on ε and which may also depend on x) such that $|f_n(x) - f(x)| < \varepsilon$ whenever $n > N$. A convenient shorthand for pointwise convergence is: $f_n \to f$ on X or, if X is understood, simply $f_n \to f$.

We say that (f_n) *converges uniformly to f on X* if, for each $\varepsilon > 0$, we can find an integer N (which depends on ε but *not* on x) such that $|f_n(x) - f(x)| < \varepsilon$ for each $x \in X$, provided that $n > N$. Please notice that the phrase "for each $x \in X$" now occurs well after the phrase "for each $\varepsilon > 0$" and, in particular, that the rate of convergence N does not depend on x. It should be reasonably clear that uniform convergence implies pointwise convergence; in other words, uniform convergence is "stronger" than pointwise convergence. For this reason, we sometimes use the shorthand: $f_n \rightrightarrows f$ on X or, if X is understood, simply $f_n \rightrightarrows f$.

The definition of uniform convergence can be simplified by "hiding" one of the quantifiers under different notation; indeed, note that the phrase "$|f_n(x) - f(x)| < \varepsilon$ for any $x \in X$" is (essentially) equivalent to the phrase "$\sup_{x\in X} |f_n(x) - f(x)| < \varepsilon$." Thus, our definition may be reworded as follows: (f_n) converges uniformly to f on X if, given $\varepsilon > 0$, there is an integer N such that $\sup_{x\in X} |f_n(x) - f(x)| < \varepsilon$ for all $n > N$.

The notion of uniform convergence exists for one very good reason: Continuity is preserved under uniform limits. This fact is well worth stating.

Exercise C.2. Let X be a subset of $\mathbb{R}$, let $f,\ f_n : X \to \mathbb{R}$ for $n = 1, 2, 3, \ldots$, and let $x_0 \in X$. If each f_n is continuous at x_0, and if $f_n \rightrightarrows f$ on X, then f is continuous at x_0. In particular, if each f_n is continuous on all of X, then so is f. Give an example showing that this result may fail if we only assume that $f_n \to f$ on X.

Finally, let's consolidate all of our findings into a useful and concrete conclusion. Given a compact interval $[\,a, b\,]$ in $\mathbb{R}$, we denote the vector space of continuous, real-valued functions $f : [\,a, b\,] \to \mathbb{R}$ by $C[\,a, b\,]$. Because $[\,a, b\,]$ is compact, we know that every element of $C[\,a, b\,]$ is bounded. Thus, the expression

$$\|f\| = \sup_{a\le x\le b} |f(x)|, \tag{C.1}$$

which is often called the *sup-norm*, is well-defined and finite for every f in $C[\,a, b\,]$. In fact, it's easy to see that (C.1) defines a norm on $C[\,a, b\,]$.

Exercise C.3. Show that (C.1) defines a norm on $C[\,a, b\,]$.

In light of our earlier discussion, it follows that convergence in the sup-norm in $C[\,a, b\,]$ is equivalent to uniform convergence; that is, a sequence (f_n) in $C[\,a, b\,]$ converges to an element f in $C[\,a, b\,]$ under the sup-norm if and only if $f_n \rightrightarrows f$ on $[\,a, b\,]$. For this reason, (C.1) is often called the *uniform norm.* Whatever we decide to call it, it's considered the norm of choice on $C[\,a, b\,]$, in part because $C[\,a, b\,]$ so happens to be complete under this norm.

Theorem C.4. *$C[\,a, b\,]$ is complete under the norm (C.1).*

Proof. Let (f_n) be a sequence in $C[\,a,b\,]$ that is Cauchy under the sup-norm; that is, for every $\varepsilon > 0$, there exists an index N such that

$$\sup_{a \le x \le b} |f_m(x) - f_n(x)| = \|f_m - f_n\| < \varepsilon$$

whenever m, $n \ge N$. In particular, for any fixed x in $[\,a,b\,]$ note that we have $|f_m(x) - f_n(x)| \le \|f_m - f_n\|$. It follows that the sequence $\big(f(x_n)\big)$ is Cauchy in $\mathbb{R}$. (Why?) Thus,

$$f(x) = \lim_{n\to\infty} f_n(x)$$

is a well-defined, real-valued function on $[\,a,b\,]$. We will show that f is in $C[\,a,b\,]$ and that $f_n \rightrightarrows f$. (But, in fact, we need only prove the second assertion, as the first will then follow.)

Given $\varepsilon > 0$, choose N such that $\|f_m - f_n\| < \varepsilon$ whenever m, $n \ge N$. Then, given any x in $[\,a,b,\,]$, we have

$$|f(x) - f_n(x)| = \lim_{m\to\infty} |f_m(x) - f_n(x)| \le \varepsilon$$

provided that $n \ge N$. Because this bound is independent of x we've actually shown that $\sup_{a \le x \le b} |f(x) - f_n(x)| \le \varepsilon$ whenever $n \ge N$. In other words, we've shown that $f_n \rightrightarrows f$ on $[\,a,b\,]$. $\square$

Appendix D

Brief Review of Linear Algebra

Sums and Quotients

We discuss sums and quotients of vector spaces. In what follows, there is no harm in assuming that all vector spaces are finite-dimensional.

To begin, given vector spaces X and Y, we write $X \oplus Y$ to denote the *direct sum* of X and Y, which may be viewed as the set of all ordered pairs (x, y), $x \in X$, $y \in Y$, endowed with the operations

$$(u, v) + (x, y) = (u + x, v + y), \quad u,\, x \in X,\ v,\, y \in Y$$

and

$$\alpha(x, y) = (\alpha x, \alpha y), \quad \alpha \in \mathbb{R},\ x \in X,\ y \in Y.$$

It is commonplace to ask whether a given vector space may be written as the sum of "smaller" factors. In particular, given subspaces Y and Z of a vector space X, we might ask whether X is isomorphic to $Y \oplus Z$, which we will paraphrase here by simply asking whether X *equals* $Y \oplus Z$. Such a pair of subspaces is said to be *complementary*.

It's not difficult to see that if each $x \in X$ can be uniquely written as a sum, $x = y + z$, where $y \in Y$ and $z \in Z$, then $X = Y \oplus Z$. Indeed, because every vector x can be so written, we must have $X = \text{span}(Y \cup Z)$ and, by uniqueness of the sum, we must have $Y \cap Z = \{0\}$; it follows that the natural splitting $x \mapsto y + x$ induces a vector space isomorphism (i.e., a linear, one-to-one, onto map) $x \mapsto (y, z)$ between X and $Y \oplus Z$. Moreover, this natural splitting induces linear idempotent maps $P : X \to X$ and $Q : X \to X$ by setting $P(x) = y$ and $Q(x) = z$ whenever $x = y + z$, $y \in Y$, $z \in Z$. That is, P and Q are linear and also satisfy $P(P(x)) = P(x)$ and $Q(Q(x)) = Q(x)$. An idempotent map is often called a *projection*, and so we might say that P and Q are *linear projections*. Please note that $P + Q = I$, the identity map on X, and $\ker P = Z$ while $\ker Q = Y$. In addition, because P and Q are idempotents, note that $\text{range}P = Y$ while $\text{range}Q = Z$. For this reason, we might refer to P and $Q = I - P$ as complementary projections.

Conversely, given a linear projection $P : X \to X$ with range Y and kernel Z, it's not hard to see that we must have $X = Y \oplus Z$. Indeed, in this case the isomorphism is given by $x \mapsto (P(x), x - P(x))$. Rather than prove this directly, it may prove helpful to state this result in other terms. For this, we need the notion of a quotient vector space.

Each subspace M of a finite-dimensional vector space X induces an equivalence relation on X by

$$x \sim y \iff x - y \in M.$$

Standard arguments show that the equivalence classes under this relation are the cosets (translates) $x + M$, $x \in X$. That is,

$$x + M = y + M \iff x - y \in M \iff x \sim y.$$

Equally standard is the induced vector arithmetic

$$(x + M) + (y + M) = (x + y) + M \qquad \text{and} \qquad \alpha(x + M) = (\alpha x) + M,$$

where $x, y \in X$ and $\alpha \in \mathbb{R}$. The collection of cosets (or equivalence classes) is a vector space under these operations; it's denoted by X/M and called the *quotient* of X by M. Please note the the zero vector in X/M is simply M itself.

Associated to the quotient space X/M is the *quotient map* $q(x) = x + M$. It's easy to check that $q : X \to X/M$ is a vector space homomorphism with kernel M. (Why?)

Next we recall *the isomorphism theorem.*

Theorem D.1. *Let $T : X \to Y$ be a linear map between vector spaces, and let $q : X \to X/\ker T$ be the quotient map. Then, there exists a (unique, into) isomorphism $S : X/\ker T \to Y$ satisfying $S(q(x)) = T(x)$ for every $x \in X$.*

Proof. Because q maps onto $X/\ker T$, it's "legal" to define a map $S : X/\ker T \to Y$ by setting $S(q(x)) = T(x)$ for $x \in X$. Please note that S is well-defined because

$$\begin{aligned} T(x) = T(y) \iff T(x - y) = 0 &\iff x - y \in \ker T \\ &\iff q(x - y) = 0 \iff q(x) = q(y). \end{aligned}$$

It's easy to see that S is linear, and so precisely the same argument as above shows that S is one-to-one. $\square$

Corollary D.2. *Let $T : X \to Y$ be a linear map between vector spaces. Then the range of T is isomorphic to $X/\ker T$. Moreover, X is isomorphic to $(\text{range}\, T) \oplus (\ker T)$.*

Corollary D.3. *If $P : X \to X$ is a linear projection on a vector space X with range Y and kernel Z, then X is isomorphic to $Y \oplus Z$.*

Inner Product Spaces

Let X be a vector space over $\mathbb{R}$. An *inner product* on X is a function $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{R}$ satisfying

1. $\langle x, y \rangle = \langle y, x \rangle$ for all $x, y \in X$.

2. $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$ for all $x, y, z \in X$ and all $a, b \in \mathbb{R}$.

3. $\langle x, x \rangle \geq 0$ for all $x \in X$ and $\langle x, x \rangle = 0$ only when $x = 0$.

It follows from conditions 1 and 2 that an inner product must be *bilinear*; that is, linear in each coordinate. Condition 3 is often paraphrased by saying that an inner product is *positive definite*. Thus, in brief, an inner product on X is a positive definite bilinear form.

An inner product on X induces a norm on X by setting

$$\|x\| = \langle\, x, x\,\rangle, \quad x \in X.$$

This is by no means obvious, by the way, and requires some explanation. The first step along this path is *the Cauchy-Schwarz inequality.*

Lemma D.4. *For any $x,\, y \in X$ we have $\big|\,\langle\, x, y\,\rangle\,\big| \leq \|x\|\,\|y\|$.*

Proof. If either of x or y is 0, the inequality clearly holds; thus, we may suppose that $x \neq 0 \neq y$. Now, for any scalar $\alpha \in \mathbb{R}$, consider:

$$\begin{aligned} 0 \;\leq\; \langle\, x - \alpha y,\, x - \alpha y\,\rangle \;&=\; \langle\, x, x\,\rangle - 2\alpha\langle\, x, y\,\rangle + \alpha^2\langle\, y, y\,\rangle &\text{(D.1)}\\ &=\; \|x\|^2 - 2\alpha\langle\, x, y\,\rangle + \alpha^2\|y\|^2. &\text{(D.2)}\end{aligned}$$

Because x and y are fixed, the right-hand side of (D.2) defines a *quadratic* in the real variable α which, according to the left-hand side of (D.1), is of constant sign on $\mathbb{R}$. It follows that the discriminant of this quadratic cannot be positive; that is, we must have

$$4\langle\, x, y\,\rangle^2 - 4\|x\|\,\|y\| \leq 0,$$

which is plainly equivalent to the assertion in the lemma. $\square$

According to the Cauchy-Schwarz inequality,

$$\frac{\big|\,\langle\, x, y\,\rangle\,\big|}{\|x\|\,\|y\|} \;=\; \left|\left\langle\, \frac{x}{\|x\|}, \frac{y}{\|y\|}\,\right\rangle\right| \;\leq\; 1$$

for any pair of nonzero vectors $x,\, y \in X$. That is, the expression $\langle\, x/\|x\|,\, y/\|y\|\,\rangle$ takes values between -1 and 1 and, as such, must be the cosine of some angle. Thus, we may define this expression to be *the cosine of the angle between the vectors x and y*; in other words, we define *the angle between x and y* to be the (unique) angle θ in $[\,0, 2\pi)$ satisfying

$$\cos\theta \;=\; \frac{\big|\,\langle\, x, y\,\rangle\,\big|}{\|x\|\,\|y\|}.$$

In particular, note that $\langle\, x, y\,\rangle = 0$ (for nonzero vectors x and y) if and only if the angle between x and y is $\pi/2$. For this reason, we say that x and y are *orthogonal* (or perpendicular) if $\langle\, x, y\,\rangle = 0$. Of course, the zero vector is orthogonal to every vector and, in fact, because the inner product is positive definite, the converse is true, too: The only vector x satisfying $\langle\, x, y\,\rangle = 0$ for all vectors y is $x = 0$.

Lemma D.5. *The expression $\|x\| = \sqrt{\langle\, x, x\,\rangle}$ defines a norm on X.*

Proof. It's easy to check that $\|x\|$ is a nonnegative expression satisfying $\|x\| = 0$ only for $x = 0$, and it's easy to see that $\|\alpha x\| = |\alpha|\,\|x\|$ for any scalar α. Thus, it only remains to prove the triangle inequality. For this, we appeal to the Cauchy-Schwarz inequality:

$$\begin{aligned} \|x+y\|^2 &= \langle\, x+y,\, x+y\,\rangle \\ &= \|x\|^2 + 2\langle\, x, y\,\rangle + \|y\|^2 \\ &\le \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 \\ &= \big(\,\|x\| + \|y\|\,\big)^2. \qquad \square \end{aligned} \tag{D.3}$$

Equation (D.3), which is entirely analogous to the *law of cosines*, by the way, holds the key to installing Euclidean geometry on an inner product space.

Corollary D.6. (The Pythagorean Theorem) *In an inner product space,* $\|x+y\|^2 = \|x\|^2 + \|y\|^2$ *if and only if* $\langle\, x, y\,\rangle = 0$.

Corollary D.7. (The Parallelogram Identity) *For any* x, y *in an inner product space* X*, we have* $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$.

We say that a set of vectors $\{x_\alpha : \alpha \in A\}$ is *orthogonal* (or, more accurately, mutually orthogonal) if $\langle\, x_\alpha, x_\beta\,\rangle = 0$ for any $\alpha \ne \beta \in A$. If, in addition, each x_α has norm one; that is, if $\langle\, x_\alpha, x_\alpha\,\rangle = 1$ for all $\alpha \in A$, we say that $\{x_\alpha : \alpha \in A\}$ is an *orthonormal* set of vectors. It is easy to see that an orthogonal set of nonzero vectors must be linearly independent. Indeed, if $x_1, \ldots, x_n$ are mutually orthogonal nonzero vectors in an inner product space X, and if we set $y = \sum_{i=1}^n \alpha_i x_i$ for some choice of of scalars $\alpha_1, \ldots, \alpha_n$, then we have

$$\langle\, y, x_j\,\rangle = \sum_{i=1}^n \alpha_i \langle\, x_i, x_j\,\rangle = \alpha_j \langle\, x_j, x_j\,\rangle.$$

Thus,

$$\alpha_i = \frac{\langle\, y, x_i\,\rangle}{\langle\, x_i, x_i\,\rangle}; \quad \text{that is,} \quad y = \sum_{i=1}^n \frac{\langle\, y, x_i\,\rangle}{\langle\, x_i, x_i\,\rangle} x_i.$$

In particular, it follows that $y = 0$ if and only if $\alpha_i = 0$ for all i which occurs if and only if $\langle\, y, x_i\,\rangle = 0$ for all i. Thus, the vectors $x_1, \ldots, x_n$ are linearly independent. Moreover, the coefficients of y, relative to the x_i, can be readily computed.

It is a natural question, then, whether an inner product space has a basis consisting of mutually orthogonal vectors. The answer is Yes and there are a couple of ways to see this. Perhaps easiest is to employ the *Gram-Schmidt process* which provides a technique for *constructing* orthogonal vectors.

Exercise D.8. Let $x_1, \ldots, x_n$ be nonzero and orthogonal and let $x \in X$.

(a) $x \in \operatorname{span}\{x_j : j = 1, \ldots, n\}$ if and only if $x = \sum_{j=1}^n \frac{\langle x, x_j\rangle}{\langle x_j, x_j\rangle}\, x_j$.

(b) For any x, the vector $y = x - \sum_{j=1}^n \frac{\langle x, x_j\rangle}{\langle x_j, x_j\rangle}\, x_j$ is orthogonal to each x_j.

(c) $\sum_{j=1}^n \frac{\langle x, x_j\rangle}{\langle x_j, x_j\rangle}\, x_j$ is the nearest point to x in $\operatorname{span}\{x_j : j = 1, \ldots, n\}$.

(d) $x \in \operatorname{span}\{x_j : j = 1, \ldots, n\}$ if and only if $\|x\|^2 = \sum_{j=1}^n \left| \frac{\langle x, x_j\rangle}{\langle x_j, x_j\rangle} \right|^2$.

Exercise D.9. (The Gram-Schmidt Process) Let (e_n) be a linearly independent *sequence* in X. Then there is an essentially unique orthogonal sequence (x_n) such that

$$\operatorname{span}\{x_1, \ldots, x_k\} = \operatorname{span}\{e_1, \ldots, e_k\}$$

for all k. [Hint: Use induction and the results in Exercise D.8 to define $x_{k+1} = e_{k+1} - v$, where $v \in \operatorname{span}\{x_1, \ldots, x_k\}$.]

Corollary D.10. *Every finite-dimensional inner product space has an orthonormal basis.*

Using techniques entirely similar to those used in proving that every vector space has a (Hamel) basis (see, for example, [8]), it can be shown that every inner product space has an orthonormal basis.

Theorem D.11. *Every inner product space has an orthonormal basis.*

Appendix E

Continuous Linear Transformations

We next discuss continuity for linear transformations (or operators) between normed vector spaces. Throughout this section, we consider a linear map $T : V \to W$ between vector spaces V and W; that is we suppose that T satisfies $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ for all x, $y \in V$, and all scalars α, β. Please note that every linear map T satisfies $T(0) = 0$. If we further suppose that V is endowed with the norm $\|\cdot\|$, and that W is endowed with the norm $|||\cdot|||$, the we may consider the issue of continuity of the map T.

The key result for our purposes is that, for linear maps, continuity—even at a single point—is equivalent to uniform continuity (and then some!).

Theorem E.1. *Let $(V, \|\cdot\|)$ and $(W, |||\cdot|||)$ be normed vector spaces, and let $T : V \to W$ be a linear map. Then, the following are equivalent:*

(i) *T is Lipschitz;*

(ii) *T is uniformly continuous;*

(iii) *T is continuous* (*everywhere*);

(iv) *T is continuous at $0 \in V$;*

(v) *there is a constant $C < \infty$ such that $|||\, T(x)\, ||| \le C\|x\|$ for all $x \in V$.*

Proof. Clearly, (i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (iv). We need to show that (iv) $\Longrightarrow$ (v), and that (v) $\Longrightarrow$ (i) (for example). The second of these is easier, so let's start there.

(v) $\Longrightarrow$ (i): If condition (v) holds for a linear map T, then T is Lipschitz (with constant C) because $|||\, T(x) - T(y)\, ||| = |||\, T(x - y)\, ||| \le C\|x - y\|$ for any x, $y \in V$.

(iv) $\Longrightarrow$ (v): Suppose that T is continuous at 0. Then we may choose a $\delta > 0$ so that $|||\, T(x)\, ||| = |||\, T(x) - T(0)\, ||| \le 1$ whenever $\|x\| = \|x - 0\| \le \delta$. (How?)

Given $0 \ne x \in V$, we may scale by the factor $\delta/\|x\|$ to get $\big\|\, \delta x/\|x\|\, \big\| = \delta$. Hence, $\big|\big|\big|\, T\big(\delta x/\|x\|\big)\, \big|\big|\big| \le 1$. But $T\big(\delta x/\|x\|\big) = (\delta/\|x\|)\, T(x)$, because T is linear, and so we get $|||\, T(x)\, ||| \le (1/\delta)\|x\|$. That is, $C = 1/\delta$ works in condition (v). (Note that because condition (v) is trivial for $x = 0$, we only care about the case $x \ne 0$.) $\square$

A linear map satisfying condition (v) of the Theorem (i.e., a continuous linear map) is often said to be *bounded*. The meaning in this context is slightly different than usual. Here it means that T maps bounded sets to bounded sets. This follows from the fact that T is Lipschitz. Indeed, if $|||\, T(x)\, ||| \le C\|x\|$ for all $x \in V$, then (as we've seen) $|||\, T(x) - T(y)\, ||| \le C\|x - y\|$ for any x, $y \in V$, and hence T maps the ball about x of radius r into the ball about $T(x)$ of radius Cr. In symbols, $T\big(B_r(x)\big) \subset B_{Cr}\big(T(x)\big)$. More generally, T maps a set of diameter d into a set of diameter at most Cd. There's no danger of confusion in our using the word bounded to mean something new here; the ordinary usage of the word (as applied to functions) is uninteresting for linear maps. A nonzero linear map always has an unbounded range. (Why?)

The smallest constant that works in (v) is called the *norm* of the operator T and is usually written $\|T\|$. In symbols,

$$\|T\| \;=\; \sup_{x \neq 0} \frac{|||Tx|||}{\|x\|} \;=\; \sup_{\|x\|=1} |||Tx|||.$$

Thus, T is bounded (continuous) if and only if $\|T\| < \infty$.

The fact that all norms on a finite-dimensional normed space are equivalent provides a final (rather spectacular) corollary.

Corollary E.2. *Let V and W be normed vector spaces with V finite-dimensional. Then, every linear map $T : V \to W$ is continuous.*

Proof. Let $x_1, \ldots, x_n$ be a basis for V and let $\|\sum_{i=1}^n \alpha_i x_i\|_1 = \sum_{i=1}^n |\alpha_i|$, as before. From the Lemma on page 3, we know that there is a constant $B < \infty$ such that $\|x\|_1 \le B\,\|x\|$ for every $x \in V$.

Now if $T : (V, \|\cdot\|\,) \to (W, |||\cdot|||)$ is linear, we get

$$\begin{aligned}
\left|\left|\left|\, T\left(\sum_{i=1}^n \alpha_i x_i\right)\right|\right|\right| &= \left|\left|\left|\, \sum_{i=1}^n \alpha_i T(x_i)\right|\right|\right| \\
&\le \sum_{i=1}^n |\alpha_i|\; |||\, T(x_i)\, ||| \\
&\le \left(\max_{1\le j\le n} |||\, T(x_j)\, |||\right) \sum_{i=1}^n |\alpha_i| \\
&\le B\left(\max_{1\le j\le n} |||\, T(x_j)\, |||\right) \left\|\sum_{i=1}^n \alpha_i x_i\right\|.
\end{aligned}$$

That is, $|||\, T(x)\, ||| \le C\|x\|$, where $C = B \max_{1\le j\le n} |||\, T(x_j)\, |||$ (a constant depending only on T and the choice of basis for V). From our last result, T is continuous (bounded). $\square$

Appendix F

Linear Interpolation

Although we won't need anything quite so fancy, it is of some interest to discuss more general problems of interpolation. We again suppose that we are given distinct points $x_0 < \cdots < x_n$ in $[\,a, b\,]$, but now we suppose that we are given an array of information

$$\begin{array}{ccccc} y_0 & y_0' & y_0'' & \cdots & y_0^{(m_0)} \\ y_1 & y_1' & y_1'' & \cdots & y_1^{(m_1)} \\ \vdots & \vdots & \vdots & & \vdots \\ y_n & y_n' & y_n'' & \cdots & y_n^{(m_n)}, \end{array}$$

where each m_i is a nonnegative integer. Our problem is to find the polynomial p of *least degree* that incorporates all of this data by satisfying

$$\begin{array}{cccc} p(x_0) = y_0 & p\,'(x_0) = y_0' & \cdots & p^{(m_0)}(x_0) = y_0^{(m_0)} \\ p(x_1) = y_1 & p\,'(x_1) = y_1' & \cdots & p^{(m_1)}(x_1) = y_1^{(m_1)} \\ \vdots & \vdots & & \vdots \\ p(x_n) = y_n & p\,'(x_n) = y_n' & \cdots & p^{(m_n)}(x_n) = y_n^{(m_n)}. \end{array}$$

In other words, we specify not only the value of p at each x_i, but also the first m_i derivatives of p at x_i. This is often referred to as the problem of *Hermite interpolation*.

Because the problem has a total of $m_0 + m_1 + \cdots + m_n + n + 1$ "degrees of freedom," it won't come as any surprise that is has a (unique) solution p of degree (at most) $N = m_0 + m_1 + \cdots + m_n + n$. Rather than discuss this particular problem any further, let's instead discuss the general problem of *linear interpolation*.

The notational framework for our problem is an n-dimensional vector space X on which m linear, real-valued functions (or *linear functionals*) $L_1, \ldots, L_m$ are defined. The general problem of linear interpolation asks whether the system of equations

$$L_i(f) = y_i, \quad i = 1, \ldots, m \tag{F.1}$$

has a (unique) solution $f \in X$ for any given set of scalars $y_1, \ldots, y_m \in \mathbb{R}$. Because a linear functional is completely determined by its values on any basis, we would next be led to consider a basis $f_1, \ldots, f_n$ for X, and from here it is a small step to rewrite (F.1) as a *matrix equation.* That is, we seek a solution $f = a_1 f_1 + \cdots + a_n f_n$ satisfying

$$\begin{aligned} a_1 L_1(f_1) + \cdots + a_n L_1(f_n) &= y_1 \\ a_1 L_2(f_1) + \cdots + a_n L_2(f_n) &= y_2 \\ &\vdots \\ a_1 L_m(f_1) + \cdots + a_n L_m(f_n) &= y_m. \end{aligned}$$

If we are to guarantee a solution $a_1, \ldots, a_n$ for each choice of $y_1, \ldots, y_m$, then we'll need to have $m = n$ and, moreover, the matrix $[L_i(f_j)]$ will have to be nonsingular.

Lemma F.1. *Let X be an n-dimensional vector space with basis vectors $f_1, \ldots, f_n$, and let $L_1, \ldots, L_n$ be linear functionals on X. Then, $L_1, \ldots, L_n$ are linearly independent if and only if the matrix $[L_i(f_j)]$ is nonsingular; that is, if and only if* $\det\big(L_i(f_j)\big) \neq 0$.

Proof. If $[L_i(f_j)]$ is singular, then the matrix equation

$$\begin{aligned} c_1 L_1(f_1) + \cdots + c_n L_n(f_1) &= 0 \\ c_1 L_1(f_2) + \cdots + c_n L_n(f_2) &= 0 \\ &\vdots \\ c_1 L_1(f_n) + \cdots + c_n L_n(f_n) &= 0 \end{aligned}$$

has a nontrivial solution $c_1, \ldots, c_n$. Thus, the functional $c_1 L_1 + \cdots + c_n L_n$ satisfies

$$(c_1 L_1 + \cdots + c_n L_n)(f_i) = 0, \quad i = 1, \ldots, n.$$

Because $f_1, \ldots, f_n$ form a basis for X, this means that

$$(c_1 L_1 + \cdots + c_n L_n)(f) = 0$$

for all $f \in X$. That is, $c_1 L_1 + \cdots + c_n L_n = 0$ (the zero functional), and so $L_1, \ldots, L_n$ are linearly dependent.

Conversely, if $L_1, \ldots, L_n$ are linearly dependent, just reverse the steps in the first part of the proof to see that $[L_i(f_j)]$ is singular. □

Theorem F.2. *Let X be an n-dimensional vector space and let $L_1, \ldots, L_n$ be linear functionals on X. Then the interpolation problem*

$$L_i(f) = y_i, \quad i = 1, \ldots, n \tag{F.2}$$

always has a (unique) solution $f \in X$ for any choice of scalars $y_1, \ldots, y_n$ if and only if $L_1, \ldots, L_n$ are linearly independent.

Proof. Let $f_1, \ldots, f_n$ be a basis for X. Then (F.2) is equivalent to the system of equations

$$\begin{aligned} a_1 L_1(f_1) + \cdots + a_n L_1(f_n) &= y_1 \\ a_1 L_2(f_1) + \cdots + a_n L_2(f_n) &= y_2 \\ &\vdots \\ a_1 L_n(f_1) + \cdots + a_n L_n(f_n) &= y_n \end{aligned} \tag{F.3}$$

by taking $f = a_1 f_1 + \cdots + a_n f_n$. Thus, (F.2) always has a solution if and only if (F.3) always has a solution if and only if $[L_i(f_j)]$ is nonsingular if and only if $L_1, \ldots, L_n$ are linearly independent. In any of these cases, note that the solution must be unique. $\square$

In the case of Lagrange interpolation, $X = \mathcal{P}_n$ and L_i is evaluation at x_i; i.e., $L_i(f) = f(x_i)$, which is easily seen to be linear in f. Moreover, $L_0, \ldots, L_n$ are linearly independent provided that $x_0, \ldots, x_n$ are distinct. (Why?)

In the case of Hermite interpolation, the linear functionals are of the form $L_{x,k}(f) = f^{(k)}(x)$, differentiation composed with a point evaluation. If $k \neq m$, then $L_{x,k}$ and $L_{x,m}$ are linearly independent; if $x \neq y$, then $L_{x,k}$ and $L_{y,m}$ are linearly independent for any k and m. (How would you check this?)

Appendix G

The Principle of Uniform Boundedness

Our goal in this section is to give an elementary proof of:

Theorem G.1. (The Uniform Boundedness Theorem) *Let $(T_\alpha)_{\alpha\in A}$ be a family of linear maps from a complete normed linear space X into a normed linear space Y. If the family is pointwise bounded, then it is, in fact, uniformly bounded. That is, if, for each $x \in X$,*

$$\sup_{\alpha\in A} \|T_\alpha(x)\| \;<\; \infty,$$

then, in fact,

$$\sup_{\alpha\in A} \|T_\alpha\| \;=\; \sup_{\alpha\in A}\,\sup_{\|x\|=1} \|T_\alpha(x)\| \;<\; \infty.$$

Proof. Suppose that $\sup_\alpha \|T_\alpha\| = \infty$. We will extract a sequence of operators (T_n) and construct a sequence of vectors (x_n) such that

(a) $\|x_n\| = 4^{-n}$, for all n, and

(b) $\|T_n(x)\| > n$, for all n, where $x = \sum_{n=1}^{\infty} x_n$.

To better understand the proof, consider

$$T_n(x) \;=\; T_n(x_1 + \cdots + x_{n-1}) \;+\; T_n x_n \;+\; T_n(x_{n+1} + \cdots).$$

The first term has norm bounded by $M_{n-1} = \sup_\alpha \|\, T_\alpha(x_1 + \cdots + x_{n-1})\,\|$. We'll choose the central term so that

$$\|T_n(x_n)\| \;\approx\; \|T_n\|\,\|x_n\| \;>>\; M_{n-1}.$$

We'll control the last term by choosing $x_{n+1}, x_{n+2}, \ldots$, to satisfy

$$\left\| \sum_{k=n+1}^{\infty} x_k \right\| \;\le\; \frac{1}{3}\,\|x_n\|.$$

Time for some details...

Suppose that $x_1, \ldots, x_{n-1}$ and $T_1, \ldots, T_{n-1}$ have been chosen. Set

$$M_{n-1} = \sup_{\alpha \in A} \| T_\alpha(x_1 + \cdots + x_{n-1}) \|.$$

Choose T_n so that

$$\|T_n\| > 3 \cdot 4^n \big(M_{n-1} + n\big).$$

Next, choose x_n to satisfy

$$\|x_n\| = 4^{-n} \quad \text{and} \quad \|T_n(x_n)\| > \frac{2}{3} \|T_n\| \, \|x_n\|.$$

This completes the inductive step.

It now follows from our construction that

$$\|T_n(x_n)\| \ > \ \frac{2}{3} \|T_n\| \, \|x_n\| \ > \ 2\big(M_{n-1} + n\big),$$

and

$$\begin{aligned} \| T_n(x_{n+1} + \cdots) \| \quad &\leq \quad \|T_n\| \sum_{k=n+1}^{\infty} 4^{-k} \\ &= \quad \|T_n\| \cdot \frac{1}{3} \cdot 4^{-n} \\ &= \quad \frac{1}{3} \|T_n\| \, \|x_n\| \\ &< \quad \frac{1}{2} \|T_n(x_n)\|. \end{aligned}$$

Thus,

$$\begin{aligned} \|T_n(x)\| \quad &\geq \quad \|T_n(x_n)\| \ - \ \| T_n(x_1 + \cdots + x_{n-1}) \| \ - \ \| T_n(x_{n+1} + \cdots) \| \\ &> \quad \frac{1}{2} \|T_n(x_n)\| \ - \ \| T_n(x_1 + \cdots + x_{n-1}) \| \\ &> \quad \big(M_{n-1} + n\big) \ - \ M_{n-1} \ = \ n. \end{aligned}$$

□

Corollary G.2. (The Banach-Steinhaus Theorem) *Suppose that (T_n) is a sequence of bounded linear operators mapping a complete normed linear space X into a normed linear space Y and suppose that*

$$T(x) = \lim_{n \to \infty} T_n(x)$$

exists in Y for each $x \in X$. Then T is a bounded linear operator.

Proof. It's obvious that T is a well-defined linear map. All that remains is to prove that T is continuous. But, because the sequence $(T_n(x))_{n=1}^\infty$ converges for each $x \in X$, it must also be bounded; that is, we must have

$$\sup_n \|T_n(x)\| < \infty \text{ for each } x \in X.$$

Thus, according to the Uniform Boundedness Theorem, we also have $C = \sup_n \|T_n\| < \infty$ and this constant will serve as a bound for $\|T\|$. Indeed, $\|T_n(x)\| \leq C\|x\|$ for every $x \in X$ and so, by the continuity of the norm in Y,

$$\|T(x)\| = \lim_{n \to \infty} \|T_n(x)\| \leq C\|x\|$$

for every $x \in X$. □

The original proof of Theorem G.1, due to Banach and Steinhaus in 1927 [3], is lost to us, I'm sorry to report. As the story goes, Saks, the referee of their paper, suggested an alternate proof using the Baire category theorem, which is the proof most commonly given these days; it is a staple in any modern introductory course on functional analysis.

I am told by Joe Diestel that their original manuscript is thought to have been lost during the war. We'll probably never know their original method of proof, but it's a fair guess that their proof was very similar to the one given above. This is not based on idle conjecture: For one, the technique of the proof (often called a "gliding hump" argument) was quite well-known to Banach and Steinhuas and had already surfaced in their earlier work. More importantly, the technique was well-known to many authors at the time; in particular, this is essentially the same proof given by Hausdorff in 1932.

What I find most curious is the fact that this proof resurfaces every few years (in the *Monthly*, for example) under the label "a non-topological proof of the uniform boundedness theorem." See, for example, Gál [19] and Hennefeld [26]. Apparently, the proof using Baire's theorem (itself an elusive result) is memorable while the gliding hump proof (based solely on first principles) is not.

Appendix H

Approximation on Finite Sets

In this chapter we consider a question of computational interest: Because best approximations are often very hard to find, how might we *approximate* the best approximation? One answer to this question lies in approximations over *finite* sets. Here's the plan:

(1) Fix a finite subset X_m of $[\,a,b\,]$ consisting of m distinct points $a \le x_1 < \cdots < x_m \le b$, and find the best approximation to f out of $\mathcal{P}_n$ *considered as a subspace of* $C(X_m)$. In other words, if we call the best approximation $p_n^*(X_m)$, then

$$\max_{1\le i\le m} |f(x_i) - p_n^*(X_m)(x_i)| = \min_{p\in\mathcal{P}_n} \max_{1\le i\le m} |f(x_i) - p(x_i)| \equiv E_n(f;X_m).$$

(2) Argue that this process converges (in some sense) to the best approximation on all of $[\,a,b\,]$ provided that X_m "gets big" as $m \to \infty$. In actual practice, there's no need to worry about $p_n^*(X_m)$ converging to p_n^* (the best approximation on all of $[\,a,b\,]$); rather, we will argue that $E_n(f;X_m) \to E_n(f)$ and appeal to "abstract nonsense."

(3) Find an efficient *strategy* for carrying out items (1) and (2).

Remarks H.1.

1. If $m \le n+1$, then $E_n(f;X_m) = 0$. That is, we can always find a polynomial $p \in \mathcal{P}_n$ that agrees with f at $n+1$ (or fewer) points. (How?) Of course, p won't be unique if $m < n+1$. (Why?) In any case, we might as well assume that $m \ge n+2$. In fact, as we'll see, the case $m = n+2$ is all that we really need to worry about.

2. If $X \subset Y \subset [\,a,b\,]$, then $E_n(f;X) \le E_n(f;Y) \le E_n(f)$. Indeed, if $p \in \mathcal{P}_n$ is the best approximation on Y, then

$$E_n(f;X) \le \max_{x\in X} |f(x) - p(x)| \le \max_{x\in Y} |f(x) - p(x)| = E_n(f;Y).$$

   Consequently, we expect $E_n(f;X_m)$ to *increase* to $E_n(f)$ as X_m "gets big."

Now if we were to repeat our earlier work on characterizing best approximations, restricting ourselves to X_m everywhere, here's what we'd get:

Theorem H.2. *Let $m \ge n+2$. Then,*

(i) *$p \in \mathcal{P}_n$ is a best approximation to f on X_m if and only if $f - p$ has an alternating set containing $n+2$ points* out of X_m*; that is, $f - p = \pm E_n(f; X_m)$, alternately, on X_m.*

(ii) *$p_n^*(X_m)$ is unique.*

Next let's see how this reduces our study to the case $m = n + 2$.

Theorem H.3. *Fix n, $m \geq n+2$, and $f \in C[\,a, b\,]$.*

(i) *If $p_n^* \in \mathcal{P}_n$ is best on all of $[\,a, b\,]$, then there is a subset X_{n+2}^* of $[\,a, b\,]$, containing $n+2$ points, such that $p_n^* = p_n^*(X_{n+2}^*)$. Moreover, $E_n(f; X_{n+2}) \leq E_n(f) = E_n(f; X_{n+2}^*)$ for any other subset X_{n+2} of $[\,a, b\,]$, with equality if and only if $p_n^*(X_{n+2}) = p_n^*$.*

(ii) *If $p_n^*(X_m) \in \mathcal{P}_n$ is best on X_m, then there is a subset X_{n+2}^* of X_m such that $p_n^*(X_m) = p_n^*(X_{n+2}^*)$ and $E_n(f; X_m) = E_n(f; X_{n+2}^*)$. For any other $X_{n+2} \subset X_m$ we have $E_n(f; X_{n+2}) \leq E_n(f; X_{n+2}^*) = E_n(f; X_m)$, with equality if and only if $p_n^*(X_{n+2}) = p_n^*(X_m)$.*

Proof. (i): Let X_{n+2}^* be an alternating set for $f - p_n^*$ over $[\,a, b\,]$ containing exactly $n+2$ points. Then, X_{n+2}^* is also an alternating set for $f - p_n^*$ over X_{n+2}^*. That is, for $x \in X_{n+2}^*$,

$$\pm(f(x) - p_n^*(x)) = E_n(f) = \max_{y \in X_{n+2}^*} |f(y) - p_n^*(y)|.$$

So, by uniqueness of best approximations on X_{n+2}^*, we must have $p_n^* = p_n^*(X_{n+2}^*)$ and $E_n(f) = E_n(f; X_{n+2}^*)$. The second assertion follows from a similar argument using the uniqueness of p_n^* on $[\,a, b\,]$.

(ii): This is just (i) with $[\,a, b\,]$ replaced everywhere by X_m. □

Here's the point: Through some as yet undisclosed method, we choose X_m with $m \geq n+2$ (in fact, $m >> n+2$) such that $E_n(f; X_m) \leq E_n(f) \leq E_n(f; X_m) + \varepsilon$, and then we search for the "best" $X_{n+2} \subset X_m$, meaning the *largest* value of $E_n(f; X_{n+2})$. We then take $p_n^*(X_{n+2})$ as an approximation for p_n^*. As we'll see momentarily, $p_n^*(X_{n+2})$ can be computed directly and explicitly.

Now suppose that the elements of X_{n+2} are $a \leq x_0 < x_1 < \cdots < x_{n+1} \leq b$, let $p = p_n^*(X_{n+2})$ be $p(x) = a_0 + a_1 x + \cdots + a_n x^n$, and let

$$E = E_n(f; X_{n+2}) = \max_{0 \leq i \leq n+1} |f(x_i) - p(x_i)|.$$

In order to compute p and E, we use the fact that $f(x_i) - p(x_i) = \pm E$, alternately, and write (for instance)

$$\begin{aligned} f(x_0) &= E + p(x_0) \\ f(x_1) &= -E + p(x_1) \\ &\vdots \\ f(x_{n+1}) &= (-1)^{n+1} E + p(x_{n+1}) \end{aligned} \tag{H.1}$$

(where the "E column" might, instead, read $-E$, E, ..., $(-1)^n E$). That is, in order to find p and E, we need to solve a system of $n+2$ *linear* equations in the $n+2$ unknowns E, $a_0, \ldots, a_n$. The determinant of this system is (up to sign)

$$\begin{vmatrix} 1 & 1 & x_0 & \cdots & x_0^n \\ -1 & 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & & \ddots & \vdots \\ (-1)^{n+1} & 1 & x_n & \cdots & x_n^n \end{vmatrix} = A_0 + A_1 + \cdots + A_{n+1} > 0,$$

where we have expanded by cofactors along the first column and have used the fact that each minor A_k is a Vandermonde determinant (and hence each $A_k > 0$). If we apply Cramer's rule to find E we get

$$\begin{aligned} E &= \frac{f(x_0)A_0 - f(x_1)A_1 + \cdots + (-1)^{n+1} f(x_{n+1})A_{n+1}}{A_0 + A_1 + \cdots + A_{n+1}} \\ &= \lambda_0 f(x_0) - \lambda_1 f(x_1) + \cdots + (-1)^{n+1}\lambda_{n+1} f(x_{n+1}), \end{aligned}$$

where $\lambda_i > 0$ and $\sum_{i=0}^{n+1} \lambda_1 = 1$. Moreover, the λ_i satisfy $\sum_{i=0}^{n+1}(-1)^i \lambda_i q(x_i) = 0$ for every polynomial $q \in \mathcal{P}_n$ because $E = E_n(q; X_{n+2}) = 0$ for polynomials of degree at most n (and because Cramer's rule supplies the same coefficients for all f).

It may be instructive to see a more explicit solution to this problem. For this, recall that because we have $n+2$ points we may interpolate exactly out of $\mathcal{P}_{n+1}$. Given this, our original problem can be rephrased quite succinctly.

Let p be the (unique) polynomial in $\mathcal{P}_{n+1}$ satisfying $p(x_i) = f(x_i)$, $i = 0, 1, \ldots, n+1$, and let e be the (unique) polynomial in $\mathcal{P}_{n+1}$ satisfying $e(x_i) = (-1)^i$, $i = 0, 1, \ldots, n+1$. If it is possible to find a scalar λ so that $p - \lambda e \in \mathcal{P}_n$, then $p - \lambda e = p_n^*(X_{n+2})$ and $|\lambda| = E_n(f; X_{n+2})$. Why? Because $f - (p - \lambda e) = \lambda e = \pm\lambda$, alternately, on X_{n+2} and so $|\lambda| = \max_{x \in X_{n+2}} |f(x) - (p(x) - \lambda e(x))|$. Thus, we need to compare leading coefficients of p and e.

Now if p has degree less than $n+1$, then $p = p_n^*(X_{n+2})$ and $E_n(f; X_{n+2}) = 0$. Thus, $\lambda = 0$ would do nicely in this case. Otherwise, p has degree exactly $n+1$ and the question is whether e does too. Setting $W(x) = \prod_{i=0}^{n+1}(x - x_i)$, we have

$$e(x) = \sum_{i=0}^{n+1} \frac{(-1)^i}{W'(x_i)} \cdot \frac{W(x)}{(x - x_i)},$$

and so the leading coefficient of e is $\sum_{i=0}^{n+1}(-1)^i / W'(x_i)$. We'll be done if we can convince ourselves that this is nonzero. But

$$W'(x_i) = \prod_{j \neq i}(x_i - x_j) = (-1)^{n-i+1} \prod_{j=0}^{i-1}(x_i - x_j) \prod_{j=i+1}^{n+1}(x_j - x_i),$$

hence $(-1)^i / W'(x_i)$ is (nonzero and) of constant sign $(-1)^{n+1}$. Finally, writing

$$p(x) = \sum_{i=0}^{n+1} \frac{f(x_i)}{W'(x_i)} \cdot \frac{W(x)}{(x - x_i)},$$

we see that p has leading coefficient $\sum_{i=0}^{n+1} f(x_i)/W'(x_i)$, making it easy to find the value of λ.

Conclusion. $p_n^*(X_{n+2}) = p - \lambda e$, where

$$\lambda \;=\; \frac{\sum_{i=0}^{n+1} f(x_i)/W'(x_i)}{\sum_{i=0}^{n+1} (-1)^i/W'(x_i)} \;=\; \sum_{i=0}^{n+1} (-1)^i \lambda_i f(x_i)$$

and

$$\lambda_i = \frac{1/|W'(x_i)|}{\sum_{j=0}^{n+1} 1/|W'(x_j)|},$$

and $|\lambda| = E_n(f; X_{n+2})$. Moreover, $\sum_{i=0}^{n+1} (-1)^i \lambda_i q(x_i) = 0$ for every $q \in \mathcal{P}_n$.

Example H.4. Find the best linear approximation to $f(x) = x^2$ on $X_4 = \{0, 1/3, 2/3, 1\} \subset [\,0, 1\,]$.

Solution. We seek $p(x) = a_0 + a_1 x$ and we need only consider subsets of X_4 of size $1 + 2 = 3$. There are four:

$$X_{4,1} = \{0, 1/3, 2/3\},\;\; X_{4,2} = \{0, 1/3, 1\},\;\; X_{4,3} = \{0, 2/3, 1\},\;\; X_{4,4} = \{1/3, 2/3, 1\}.$$

In each case we find a p and a λ ($= E$ in our earlier notation). For instance, in the case of $X_{4,2}$ we would solve the system of equations $f(x) = \pm\lambda + p(x)$ for $x = 0, 1/3, 1$.

$$\begin{array}{rcl} 0 &=& \lambda^{(2)} + a_0 \\ \frac{1}{9} &=& -\lambda^{(2)} + a_0 + \frac{1}{3}a_1 \\ 1 &=& \lambda^{(2)} + a_0 + a_1 \end{array} \qquad \Longrightarrow \qquad \begin{array}{rcl} \lambda^{(2)} &=& \frac{1}{9} \\ a_0 &=& -\frac{1}{9} \\ a_1 &=& 1 \end{array}$$

In the other three cases you would find that $\lambda^{(1)} = 1/18$, $\lambda^{(3)} = 1/9$, and $\lambda^{(4)} = 1/18$. Because we need the largest λ, we're done: $X_{4,2}$ (or $X_{4,3}$) works, and $p_1^*(X_4)(x) = x - 1/9$. (Recall that the best approximation on all of $[\,0, 1\,]$ is $p_1^*(x) = x - 1/8$.) $\square$

Where does this leave us? We still need to know that there is some hope of finding an initial set X_m with $E_n(f) - \varepsilon \le E_n(f; X_m) \le E_n(f)$, and we need a more efficient means of searching through the $\binom{m}{n+2}$ subsets $X_{n+2} \subset X_m$. In order to attack the problem of finding an initial X_m, we'll need a few classical inequalities. We won't directly attack the second problem; instead, we'll outline an algorithm that begins with an initial set X_{n+2}^0, containing exactly $n + 2$ points, which is then "improved" to some X_{n+2}^1 by changing only a single point.

Convergence of Approximations over Finite Sets

In order to simplify things here, we will make several assumptions: For one, we will consider only approximation over the interval $I = [-1, 1\,]$. As before, we consider a fixed $f \in C[-1, 1\,]$ and a fixed integer $n = 0, 1, 2, \ldots$. For each integer $m \ge 1$ we choose a finite subset $X_m \subset I$, consisting of m points $-1 \le x_1 < \cdots < x_m \le 1$; in addition, we will assume that $x_1 = -1$ and $x_m = 1$. If we put

$$\delta_m \;=\; \max_{x \in I} \min_{1 \le i \le m} |x - x_i| > 0,$$

then each $x \in I$ is within δ_m of some x_i. If X_m consists of equally spaced points, for example, it's easy to see that $\delta_m = 1/(m-1)$.

Our goal is to prove

Theorem H.5. *If $\delta_m \to 0$, then $E_n(f; X_m) \to E_n(f)$.*

And we would hope to accomplish this in such a way that δ_m is a measurable quantity, depending on f, m, and a prescribed tolerance $\varepsilon = E_n(f; X_m) - E_n(f)$.

As a first step in this direction, let's bring Markov's inequality into the picture.

Lemma H.6. *Suppose that $\tau_m \equiv \delta_m^2 n^4/2 < 1$. Then, for any $p \in \mathcal{P}_n$, we have*

(1) $\max_{-1 \le x \le 1} |p(x)| \;\le\; (1-\tau_m)^{-1} \max_{1 \le i \le m} |p(x_i)|$, *and*

(2) $\omega_p([-1,1]; \delta_m) \;\le\; \delta_m n^2 (1-\tau_m)^{-1} \max_{1 \le i \le m} |p(x_i)|$.

Proof. (1): Take a in $[-1,1]$ with $|p(a)| = \|p\|$. If $a = \pm 1 \in X_m$, we're done (because $(1-\tau_m)^{-1} > 1$). Otherwise, we'll have $-1 < a < 1$ and $p'(a) = 0$. Next, choose $x_i \in X_m$ with $|a - x_i| \le \delta_m$ and apply Taylor's theorem:

$$p(x_i) = p(a) + (x_i - a)\, p'(a) + \frac{(x_i - a)^2}{2}\, p''(c),$$

for some c in $(-1,1)$. Re-writing, we have

$$|p(a)| \le |p(x_i)| + \frac{\delta_m^2}{2}\, |p''(c)|.$$

And now we bring in Markov:

$$\|p\| \le \max_{1 \le i \le m} |p(x_i)| \;+\; \frac{\delta_m^2 n^4}{2}\, \|p\|,$$

which is what we need.

(2): The real point here is that each $p \in \mathcal{P}_n$ is *Lipschitz* with constant $n^2\|p\|$. Indeed,

$$|p(s) - p(t)| = |(s-t)\, p'(c)| \le |s-t|\, \|p'\| \le n^2 \|p\|\, |s-t|$$

(from the mean value theorem and Markov's inequality). Thus, $\omega_p(\delta) \le \delta n^2 \|p\|$ and, combining this with (1), we get

$$\omega_p(\delta_m) \;\le\; \delta_m n^2 \|p\| \;\le\; \delta_m n^2 (1-\tau_m)^{-1} \max_{1 \le i \le m} |p(x_i)|. \qquad \square$$

Now we're ready to compare $E_n(f; X_m)$ to $E_n(f)$. Our result won't be as good as Rivlin's (he uses a fancier version of Markov's inequality), but it will be a bit easier to prove. As in Lemma H.6, we'll suppose that

$$\tau_m = \frac{\delta_m^2 n^4}{2} < 1,$$

and we'll set

$$\Delta_m = \frac{\delta_m n^2}{1 - \tau_m}.$$

[Note that as $\delta_m \to 0$ we also have $\tau_m \to 0$ and $\Delta_m \to 0$.]

Theorem H.7. *For $f \in C[-1,1]$,*

$$E_n(f; X_m) \;\le\; E_n(f) \;\le\; (1 + \Delta_m)\, E_n(f; X_m) \;+\; \omega_f([-1,1]; \delta_m) \;+\; \Delta_m \|f\|.$$

Consequently, if $\delta_m \to 0$, then $E_n(f; X_m) \to E_n(f)$ (as $m \to \infty$).

Proof. Let $p = p_n^*(X_m) \in \mathcal{P}_n$ be the best approximation to f on X_m. Recall that

$$\max_{1 \le i \le m} |f(x_i) - p(x_i)| \;=\; E_n(f; X_m) \;\le\; E_n(f) \;\le\; \|f - p\|.$$

Our plan is to estimate $\|f - p\|$.

Let $x \in [-1, 1]$ and choose $x_i \in X_m$ with $|x - x_i| \le \delta_m$. Then,

$$\begin{aligned} |f(x) - p(x)| &\le |f(x) - f(x_i)| + |f(x_i) - p(x_i)| + |p(x_i) - p(x)| \\ &\le \omega_f(\delta_m) \;+\; E_n(f; X_m) \;+\; \omega_p(\delta_m) \\ &\le \omega_f(\delta_m) \;+\; E_n(f; X_m) \;+\; \Delta_m \max_{1 \le i \le m} |p(x_i)|, \end{aligned}$$

where we've used (2) from Lemma H.6 to estimate $\omega_p(\delta_m)$. All that remains is to revise this last estimate, eliminating reference to p. For this we use the triangle inequality again:

$$\begin{aligned} \max_{1 \le i \le m} |p(x_i)| &\le \max_{1 \le i \le m} |f(x_i) - p(x_i)| \;+\; \max_{1 \le i \le m} |f(x_i)| \\ &\le E_n(f; X_m) \;+\; \|f\|. \end{aligned}$$

Putting all the pieces together gives us our result:

$$E_n(f) \;\le\; \omega_f(\delta_m) \;+\; E_n(f; X_m) \;+\; \Delta_m \big[\, E_n(f; X_m) \;+\; \|f\| \,\big]. \qquad \square$$

As Rivlin points out, it is quite possible to give a lower bound on m in the case of, say, equally spaced points, which will give $E_n(f; X_m) \le E_n(f) \le E_n(f; X_m) + \varepsilon$, but this is surely an inefficient approach to the problem. Instead, we'll discuss the *one point exchange algorithm*.

The One Point Exchange Algorithm

We're given $f \in C[-1, 1]$, n, and $\varepsilon > 0$.

1. Pick a starting "reference" X_{n+2}. A convenient choice is the set $x_i = \cos\left(\frac{n+1-i}{n+1}\,\pi\right)$, $i = 0, 1, \ldots, n+1$. These are the "peak points" of T_{n+1}; that is, $T_{n+1}(x_i) = (-1)^{n+1-i}$ (and so T_{n+1} is the polynomial e from our Conclusion on page 146).

2. Find $p = p_n^*(X_{n+2})$ and λ (by solving a system of linear equations). Recall that

$$|\lambda| \;=\; |f(x_i) - p(x_i)| \;\le\; \|f - p^*\| \;\le\; \|f - p\|,$$

where p^* is the best approximation to f on all of $[-1, 1]$.

3. Find (approximately, if necessary) the "error function" $e(x) = f(x) - p(x)$ and any point η where $|f(\eta) - p(\eta)| = \|f - p\|$. (According to Powell [43], this can be accomplished using "local quadratic fits.")

4. Replace an appropriate x_i by η so that the new reference set $X'_{n+2} = \{x'_1, x'_2, \ldots\}$ has the properties that $f(x'_i) - p(x'_i)$ alternates in sign and that $|f(x'_i) - p(x'_i)| \ge |\lambda|$ for all i. The new polynomial $p' = p_n^*(X'_{n+2})$ and new λ' must then satisfy

$$|\lambda| \;=\; \min_{0 \le i \le n+1} |f(x'_i) - p(x'_i)| \;\le\; \max_{0 \le i \le n+1} |f(x'_i) - p'(x'_i)| \;=\; |\lambda'|.$$

This is an observation due to de La Vallée Poussin: Because $f - p$ alternates in sign on an alternating set for $f - p'$, it follows that $f - p'$ increases the minimum error over this set. (See Theorem 4.9 for a precise statement.) Again according to Powell [43], the new p' and λ' can be found quickly through matrix "updating" techniques. (Because we've only changed one of the x_i, only one row of the matrix (H.1) needs to be changed.)

5. The new λ' satisfies $|\lambda'| \le \|f - p^*\| \le \|f - p'\|$, and the calculation stops when

$$\|f - p'\| - |\lambda'| \;=\; |f(\eta') - p'(\eta')| - |\lambda'| \;<\; \varepsilon.$$

References

[1] M. Abramowitz and I. A. Stegun *Handbook of Mathematical Functions with Formulas, Graphs, and Tables*, Dover, New York, 1965.

[2] S. Banach, *Theory of Linear Operations*, North-Holland Mathematical Library v. 38, North-Holland, New York, 1987.

[3] S. Banach and H. Steinhaus, Sur le principle de la condensation de singularités, *Fundamenta Mathematicae* **9** (1927), 50–61.

[4] B. Beauzamy, *Introduction to Banach Spaces and their Geometry*, North-Holland Mathematics Studies, v. 68, North-Holland, Amsterdam, 1982.

[5] S. N. Bernstein, Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités, *Communications of the Kharkov Mathematical Society* (2) **13** (1912), 1-2.

[6] S. N. Bernstein, Quelques remarques sur l'interpolation, *Mathematische Annalen* **79** (1918), 1–12.

[7] G. Birkhoff, ed., *A Source Book in Classical Analysis*, Harvard University Press, Cambridge, Mass., 1973.

[8] G. Birkhoff and S. Mac Lane, *A Survey of Modern Algebra*, 3rd. ed., Macmillan, New York, 1965.

[9] R. P. Boas, Jr., Inequalities for the derivatives of polynomials, *Mathematics Magazine* **42**, (1969), 165–174.

[10] N. L. Carothers, *Real Analysis*, Cambridge University Press, New York, 2000.

[11] N. L. Carothers, *A Short Course on Banach Space Theory*, London Mathematical Society Student Texts **64**, Cambridge University Press, New York, 2005.

[12] E. W. Cheney, *Introduction to Approximation Theory*, Chelsea, New York, 1982.

[13] P. J. Davis, *Interpolation and Approximation*, Dover, New York, 1975.

[14] M. M. Day, Strict convexity and smoothness of normed spaces, *Transactions of the American Mathematical Society* **78** (1955), 516–528.

[15] R. A. DeVore and G. G. Lorentz, *Constructive Approximation*, Grundlehren der mathematischen Wissenschaften, bd. 303, Springer-Verlag, Berlin-Heidelberg-New York, 1993.

[16] S. Fisher, Quantitative approximation theory, *The American Mathematical Monthly* **85** (1978), 318–332.

[17] G. Folland, *Real Analysis*, 2nd ed., Wiley-Interscience, New York, 1999.

[18] L. Fox and I. B. Parker, *Chebyshev Polynomials in Numerical Analysis*, Oxford University Press, New York, 1968.

[19] I. S. Gál, On sequences of operations in complete vector spaces, *The American Mathematical Monthly* **60** (1953), 527–538.

[20] C. Goffman and G. Pedrick, *First Course in Functional Analysis*, 2nd. ed., Chelsea, New York, 1983.

[21] A. Haar, Die Minkowskische Geometrie und die Annäherung an stetige Funktionen, *Mathematische Annalen* **78** (1918), 294–311.

[22] D. T. Haimo, ed., Orthogonal Expansions and their Continuous Analogues, Proceedings of a Conference held at Southern Illinois University, Edwardsville, April 27–29, 1967, Southern Illinois University Press, Carbondale, 1968.

[23] P. R. Halmos, *Finite-Dimensional Vector Spaces*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1974, and Litton Educational Publishing, 1958; reprint of the 2nd. ed., Van Nostrand, Princeton, New Jersey.

[24] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, 2nd. ed., Cambridge University Press, New York, 1983.

[25] E. R. Hedrick, The significance of Weierstrass's theorem, *The American Mathematical Monthly* **20** (1927), 211–213.

[26] J. Hennefeld, A nontopological proof of the uniform boundedness theorem, *The American Mathematical Monthly* **87** (1980), 217.

[27] C. Hermite, Sur la formule d'interpolation de Lagrange, *Journal für die Reine und Angewandte Mathematik* **84** (1878), 70–79.

[28] E. W. Hobson, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, 2 vols., Harren Press, Washington, 1950.

[29] K. Hoffman, *Analysis on Euclidean Spaces*, Prentice-Hall, Englewood Cliffs, New Jersey, 1975.

[30] R. B. Holmes, *Geometric Functional Analysis*, Graduate Texts in Mathematics 24, Springer-Verlag, New York, 1975.

[31] D. Jackson, The general theory of approximation by polynomials and trigonometric sums, *Bulletin of the American Mathematical Society* **27** (1920–1921), 415–431.

[32] D. Jackson, *The Theory of Approximation*, Colloquium Publications, vol. XI, American Mathematical Society, New York 1930 reprinted 1968.

[33] D. Jackson, *Fourier Series and Orthogonal Polynomials*, The Carus Mathematical Monographs, no. 6, Mathematical Association of America, Menasha, Wisconsin, 1941.

[34] P. Kirchberger, Über Tchebychefsche Annäherungsmethoden, *Mathematische Annalen* **57** (1903), 509–540.

[35] T. W. Körner, *Fourier Analysis*, Cambridge University Press, New York, 1988.

[36] P. P. Korovkin, *Linear Operators and Approximation Theory*, Hindustan Publishing, Delhi, 1960.

[37] C. J. de La Vallée Poussin, *Leçons sur l'Approximation des Fonctions d'une Variable Réele*, Gauthier-Villars, Paris, 1919; reprinted in *L'Approximation* by S. Bernstein and C. J. de La Vallée Poussin, Chelsea, New York, (reprinted) 1970.

[38] H. Lebesgue, Sur l'approximation des fonctions, *Bulletin des Sciences Mathématique* **22** (1898), 278–287.

[39] G. G. Lorentz, *Approximation of Functions*, Chelsea, New York, 1986.

[40] G. G. Lorentz, *Bernstein Polynomials*, Chelsea, New York, 1986.

[41] I. Natanson, *Constructive Function Theory*, 3 vols., Frederick Ungar, New York, 1964–1965.

[42] M. J. D. Powell, On the maximum errors of polynomial approximations defined by interpolation and by least squares criteria, *The Computer Journal* **9** (1967),404–407.

[43] M. J. D. Powell, *Approximation Theory and Methods*, Cambridge University Press, New York, 1981.

[44] T. J. Rivlin, *The Chebyshev Polynomials*, Wiley, New York, 1974.

[45] T. J. Rivlin, *An Introduction to the Approximation of Functions*, Dover, New York, 1981.

[46] W. Rudin, *Principles of Mathematical Analysis*, 3rd. ed., McGraw-Hill, New York, 1976.

[47] A. Shields, Years ago, *The Mathematical Intelligencer* **9**, No. 2 (1987), 61–63.

[48] A. Shields, Years ago, *The Mathematical Intelligencer* **9**, No. 3 (1987), 5–7.

[49] A. J. Shohat, On the development of functions in series of orthogonal polynomials, *Bulletin of the American Mathematical Society* **41** (1935), 49–82.

[50] J. A. Shohat and J. D. Tamarkin, *The Problem of Moments*, American Mathematical Society, New York, 1943.

[51] G. F. Simmons, *Introduction to Topology and Modern Analysis*, McGraw-Hill, New York, 1963; reprinted by Robert E. Krieger Publishing, Malabar, Florida, 1986.

[52] M. H. Stone, Applications of the theory of Boolean rings to general topology, *Transactions of the American Mathematical Society* **41** (1937), 375–481.

[53] M. H. Stone, A generalized Weierstrass theorem, in *Studies in Modern Analysis*, R. C. Buck, ed., Studies in Mathematics, vol. 1, Mathematical Association of America, Prentice-Hall, Englewood Cliffs, New Jersey, 1962.

[54] E. B. Van Vleck, The influence of Fourier's series upon the development of mathematics, *Science* **39** (1914), 113–124.

[55] B. L. van der Waerden *Modern Algebra*, 2 vols., translated from the 2nd rev. German ed. by Fred Blum, Frederick Ungar, New York, 1949–50.

[56] K. Weierstrass, Über die analytische Darstellbarkeit sogenannter willkürlicher Functionen einer reellen Veränderlichen, *Sitzungsberichte der K̈oniglich Preussischen Akademie der Wissenschaften zu Berlin*, (1885), 633–639, 789–805. See also, *Mathematische Werke*, Mayer and Müller, 1895, vol. 3, 1–37. A French translation appears as Sur la possibilité d'une représentation analytique des fonctions dites arbitraires d'une variable réele, *Journal de Mathématiques Pures et Appliquées* **2** (1886), 105–138.