1. Whether Polynomial and/or Trigonometric Approximation is Possible at All.

1.A. Introduction. Before we start, just a reminder of a definition:

Definition: Let $X$ be a set and $f : X \to \mathbb{R}$ or $f : X \to \mathbb{C}$ be a real- or complex-valued function on $X$. Let $\mathfrak{A}$ be a family of real- or complex-valued functions on $X$. We say that $f$ can be **uniformly approximated on $X$ by functions of the class $\mathfrak{A}$** if for every $\epsilon > 0$ there is a function $g \in \mathfrak{A}$ such that $|f(x) - g(x)| \leq \epsilon$ holds for every $x \in X$.

In words: “for every prescribed error tolerance $\epsilon > 0$, some (single) function $g \in \mathfrak{A}$ can be found that differs from $f$ at each point of $X$ by less than the prescribed error tolerance.”

The most universally useful non-effective theorem for establishing “uniform approximability” of all continuous functions by a relatively small class of continuous functions is the Stone-Weierstraß theorem. The proof of this theorem is not at all difficult but would take these notes rather far afield. I shall state the theorem in full generality, but wherever you see “compact Hausdorff space $X$” it would suffice for you to read “closed and bounded subset of a Euclidean space $\mathbb{R}^n$.” (Recall that closed means that $X$ contains the limit of every convergent sequence of elements of $X$.)

As usual, if $f : X \to \mathbb{R}$ or $\mathbb{C}$ is a bounded function on a set $X$, we put

$$
\| f \|_\infty = \sup \{|f(x)| : x \in X\}.
$$

In all the cases we shall consider, either $X$ will be a finite set or it will be a compact set in some Euclidean space and $f$ will be a continuous function: so if least upper bounds give you the vapors, you can read “max” for “sup” (since in either of those cases $|f(x)|$ actually takes a largest value on $X$). For a discussion of some of the properties of the “sup norm” that make the vector spaces of continuous functions $C_\mathbb{R}(X)$ and $C_\mathbb{C}(X)$ resemble Euclidean spaces (in their $\| \cdot \|_\infty$ norms), see Atkinson, pp. 200–201.

[Stone-Weierstraß] Theorem: Let $X$ be a compact Hausdorff space.

Real Scalars: Let $\mathcal{C}(X)$ denote the set of real-valued continuous functions on $X$, with the usual operations of addition and multiplication. Suppose $\mathfrak{A} \subseteq \mathcal{C}(X)$ is a family of real-valued continuous functions that has the following properties:

1. $1 \in \mathfrak{A}$ (where “1” denotes the identically-1 function);
2. $\mathfrak{A}$ is an algebra: if $\alpha$ and $\beta$ are real constants and $g$ and $h$ belong to $\mathfrak{A}$, then $\alpha g + \beta h$ and $g \cdot h$ also belong to $\mathfrak{A}$;
3. $\mathfrak{A}$ separates points of $X$: for any two distinct points $x_1$ and $x_2$ in $X$ there exists some $h \in \mathfrak{A}$ for which $h(x_1) \neq h(x_2)$.

Then any continuous real-valued function on $X$ can be uniformly approximated by elements of $\mathfrak{A}$: that is, given any $f \in \mathcal{C}(X)$ and any $\epsilon > 0$ there exists some $g \in \mathfrak{A}$ with $\| f - g \|_\infty \leq \epsilon$.

Complex Scalars: Let $\mathcal{C}(X)$ denote the set of complex-valued continuous functions on $X$, with the usual operations of addition and multiplication. Suppose $\mathfrak{A} \subseteq \mathcal{C}(X)$ is a family of complex-valued continuous functions that has the following properties:

1. $1 \in \mathfrak{A}$ (where “1” denotes the identically-1 function);
2. $\mathfrak{A}$ is an algebra: if $\alpha$ and $\beta$ are complex constants and $g$ and $h$ belong to $\mathfrak{A}$, then $\alpha g + \beta h$ and $g \cdot h$ also belong to $\mathfrak{A}$;
3. $\mathfrak{A}$ separates points of $X$: for any two distinct points $x_1$ and $x_2$ in $X$ there exists some $h \in \mathfrak{A}$ for which $h(x_1) \neq h(x_2)$.
4. If $g \in \mathfrak{A}$ then also its complex-conjugate function belongs to $\mathfrak{A}$: that is, if $g = \varphi + i\psi$ belongs to $\mathfrak{A}$ then so does $\overline{g} = \varphi - i\psi$. 

TRIGONOMETRIC AND POLYNOMIAL APPROXIMATION
Then the conclusion is the same: any continuous complex-valued function on $X$ can be uniformly approximated by elements of $\mathfrak{A}$: i.e., given any $f \in \mathcal{C}(X)$ and any $\epsilon > 0$ there exists some $g \in \mathfrak{A}$ with $\|f - g\|_\infty \leq \epsilon$.

A proof of this theorem can be found in almost every first-year functional analysis textbook. People who would like to see a proof and never have are probably students of mathematics and should look up a proof in a familiar text. It should be noted that many textbook treatments that reproduce Stone’s original “lattice” proof assume the classical Weierstraß approximation theorem as a lemma. However, the only lemma they really need invariably boils down to the following one, which is elementary.

**Lemma:** The function $|x|$ is uniformly approximable on the interval $-1 \leq x \leq 1$ by polynomials in $x$.

**Proof.** Using Stirling’s asymptotic approximation to $n!$ (or more elementary arguments) it is easy to show that the coefficients in the binomial expansion = Maclaurin series

\[(1 + z)^{1/2} = \sum_{k=0}^{\infty} \binom{1/2}{k} z^k \tag{1.A.1}\]

are asymptotic as $k \to \infty$ to a constant multiple of $\frac{1}{k^{3/2}}$. Since $\sum \frac{1}{k^{3/2}} < \infty$ by the integral test, the Maclaurin series (1.A.1) converges uniformly on the interval $[-1, 1]$ to a continuous function on the interval $[-1, 1]$. Replacing $z$ by $(x^2 - 1)$, we find that

\[|x| = (x^2)^{1/2} = \sum_{k=0}^{\infty} \binom{1/2}{k} (x^2 - 1)^k \tag{1.A.2}\]

converges uniformly for $-1 \leq x \leq 1$—in fact, the error after $n$ terms is $O(\sum_{k=n}^{\infty} 1/k^{3/2}) = O(n^{-1/2})$—and though the series is not a power series, it is a series of polynomials that converges uniformly for $-1 \leq x \leq 1$, and so given any $\epsilon > 0$ a suitable partial sum of this series will exhibit a polynomial that approximates the function $|x|$ uniformly for $-1 \leq x \leq 1$ with error uniformly smaller than $\epsilon$.

As an application (although we want to give a constructive proof of the classical Weierstraß approximation theorem below), here is the classical Weierstraß approximation theorem as a special case of Stone-Weierstraß.

**[Classical Weierstraß Approximation] Theorem:** Given any continuous real- or complex valued function $f$ on an interval $[a, b]$, there exists for any given $\epsilon > 0$ a polynomial $p(x)$ (with real or complex coefficients respectively) such that

\[\max\{|f(x) - p(x)| : x \in [a, b]\} \leq \epsilon.\]

**Proof.** In the Stone-Weierstraß theorem as given above, let $\mathfrak{A}$ be the polynomials with real or complex coefficients respectively, considered as functions on the interval $[a, b]$. It is trivial to verify that this $\mathfrak{A}$ satisfies conditions (1) and (2), and (3) is satisfied because $\mathfrak{A}$ contains the identity function $h(x) \equiv x$. In the case of complex scalars, we observe that given a polynomial

\[q(x) = a_0 + \cdots + a_n x^n\]

(1) “Almost everywhere” does not mean everywhere. John B. Conway’s *A Course in Functional Analysis*, 2nd ed., Springer GTM 96 (1990) is an exception. Leave it to Conway to not to give Stone’s insightful proof, but rather to give the 1959 proof of de Branges that uses the Krein-Mil’man theorem as its principal lemma! (Of course, de Branges’ argument can be adapted to prove other interesting things; but it’s not at all transparent in the way that Stone’s is.) You can do it that way too—but see Stone’s original argument in other places, for instance, N. Dunford and J. T. Schwartz, *Linear Operators*, v.1, and find out “why” it’s true. Stone’s proof is almost constructive, if you know the modulus of continuity of the function you’re trying to approximate.

(2) Henri Lebesgue (of Lebesgue measure) seems to have anticipated Stone by observing that if one could approximate functions of the form $|x - a|$ by polynomials, then one could approximate all continuous functions by polynomials.
we can form the polynomial
\[ q^*(x) = \overline{a}_0 + \cdots + \overline{a}_n x^n \]
and then, because \( x \in [a, b] \) is real and thus equals its own conjugate, we shall have
\[ q^*(x) = \overline{q(x)} , \]
so \( \mathfrak{A} \) contains with each of its functions the complex conjugate function, and thus satisfies (4). The original Weierstraß theorem thus appears as a special case of Stone’s generalization.

Exactly the same argument, *mutatis mutandis*, proves that if \( X \subseteq \mathbb{R}^k \) is a closed and bounded set in \( k \)-dimensional Euclidean space, then given a continuous real- or complex-valued function on \( X \) and given \( \epsilon > 0 \) there exists a polynomial \( p(x_1, \ldots, x_k) \) in the coördinates \( x_1, \ldots , x_k \), with real or complex coefficients respectively, such that
\[ \max\{ |f(x) - p(x_1, \ldots, x_k)| : (x_1, \ldots, x_k) = x \in X \} < \epsilon . \]

Without additional “smoothness” information about \( f \), there is no nice function \( d(f, n) \) (say) that will tell us what the degree of \( p(x) \) must be in order to approximate \( f \) uniformly with error \( < \frac{1}{n} \), say. There are such theorems, called *theorems of Jackson type* after their discoverer, and we shall introduce those below as we require them. Atkinson cites them as Theorem (4.11), pp. 224–225. {In fact, the approach to the “distance-measuring function” is a little different: we put
\[ \rho_n(f) = \min\{ ||f - p||_\infty : p \in \mathcal{P}_n \} \]
and Jackson-type theorems estimate \( \rho_n(f) \) with functions of \( n \) of the form \( \frac{\text{const.}}{n^\beta} \), where \( \beta > 0 \) is a measure of the smoothness of \( f \) [i.e., essentially, the number of continuous derivatives \( f \) has].}

One can specialize and ask for approximating polynomials of particular special forms: for example, the Bernstein polynomials that will be introduced below are linear combinations of particular fixed polynomials (the Bernstein basis), and Bernstein’s theorem gives a way to produce a sequence of linear combinations of Bernstein polynomials that converges uniformly on \([a, b]\) to the given \( f \in C([a, b]) \). For another example, if \([a, b]\) is contained in the half-line \( \{ x \in \mathbb{R} : x < -1 \} \), then every continuous function on \([a, b]\) can be uniformly approximated arbitrarily closely by polynomials \( c_n x^n + \cdots + c_1 x + c_0 \) with all the \( c_j \geq 0 \). But in the numerical-analysis context, what one needs most is something resembling an algorithm to generate approximating polynomials from a given continuous function.

1.B. Linear, Positive Approximation Methods. There is very little concrete information in the Stone-Weierstraß theorem: given an error tolerance \( \epsilon > 0 \) and a way to compute as (finitely) many values of the continuous function \( f \) on \([a, b]\) as you wished, you would have no clue about producing a polynomial that approximated \( f \) uniformly within \( \epsilon \) on \([a, b]\). There is an amazingly-easy-to-prove result that shows that if the computational scheme one uses to make approximations satisfies certain very natural requirements, then to show that it “works” in the sense that it yields approximations to arbitrary continuous functions, it is only necessary to check it on the functions \( 1, x \) and \( x^2 \).

It helps to make some preliminary definitions. Let \( X = [a, b] \subseteq \mathbb{R} \). A mapping (or “operator”) \( K : C(X) \to C(X) \) is said to be linear if it can be applied to linear combinations of functions term-by-term: \( K[\alpha f + \beta g] = \alpha K[f] + \beta K[g] \) should hold for all constants \( \alpha, \beta \) and functions \( f, g \in C(X) \). It is said to be positive (or nonnegative) if \( K[f] \) is a nonnegative function whenever \( f \) is: \( f \geq 0 \Rightarrow K[f] \geq 0 \). A positive linear operator preserves order relations: if \( f \) and \( g \) are two real-valued functions and \( g \leq f \) holds on \( X \), then \( f - g \geq 0 \) and so \( 0 \leq K[f-g] = K[f] - K[g] \), which says \( K[g] \leq K[f] \). A less obvious property of

---

(3) See R. D. Nussbaum and B. Walsh, *Approximation by polynomials with nonnegative coefficients* ... , Trans. Amer. Math. Soc. 350 (1998), pp. 2367–2391; this kind of approximation has applications in the spectral theory of linear operators. The proofs of approximability in this paper, however, are nonconstructive.
positive linear operators, however, is the following: if \( f, g \in \mathcal{C}(X) \) and we form the function \( f \wedge g \) defined by \( (f \wedge g)(x) = \min\{f(x), g(x)\} \), then \( f \wedge g \leq f \) and \( f \wedge g \leq g \), so \( K[f \wedge g] \leq K[f] \) and \( K[f \wedge g] \leq K[g] \) —but if those two relations hold, then \( K[f \wedge g] \leq K[f] \wedge K[g] \). This observation can be extended to apply to any finite collection of continuous function on \( X \): if \( \{f_1, \ldots, f_k\} \subseteq \mathcal{C}(X) \) is a finite collection of continuous functions on \( X \) and we form \( f_1 \wedge \cdots \wedge f_k \), whose value at any \( x \in X \) is the least of the values \( \{f_1(x), \ldots, f_k(x)\} \), then \( K[f_1 \wedge \cdots \wedge f_k] \leq K[f_1] \wedge \cdots \wedge K[f_k] \). Multiplication by \((-1)\) tells us that if \( f_1 \vee \cdots \vee f_k \) is defined by \( (f_1 \vee \cdots \vee f_k)(x) = \max\{f_1(x), \ldots, f_k(x)\} \), then \( K[f_1 \vee \cdots \vee f_k] \geq K[f_1] \vee \cdots \vee K[f_k] \).

It is now easy to state and prove the

**Bohman-Korovkin Theorem:** Let \([a, b] \subseteq \mathbb{R}\) be a closed interval of real numbers, and let \( \{K_n\}_{n=1}^{\infty} \) be a sequence of positive linear operators on \( \mathcal{C}_a([a, b]) \) with the property that the three sequences of continuous functions \( \{K_n[1]\}_{n=1}^{\infty}, \{K_n[x]\}_{n=1}^{\infty}, \) and \( \{K_n[x^2]\}_{n=1}^{\infty} \) converge uniformly on \([a, b]\) to the functions 1, \( x \) and \( x^2 \) respectively. Then for any continuous function \( f \in \mathcal{C}([a, b]) \), the sequence \( \{K_n[f]\}_{n=1}^{\infty} \) converges uniformly to \( f \) on \([a, b]\); that is, given any \( \epsilon > 0 \) we can find \( N \) such that \( n \geq N \Rightarrow \max\{|f(x) - K_n[f](x)| : a \leq x \leq b\} \leq \epsilon \).  

**Proof.** We need two calculus facts: the first is that a continuous function on a closed interval is bounded, \( i.e. \), there exists a constant \( M \geq 0 \) for which \( |f(x)| \leq M \) holds for all \( x \in [a, b] \). The second is the **Heine-Borel theorem:** if for each \( t \in [a, b] \) we are given an open interval \((a_t, b_t)\) containing \( t \), then it is possible to find a finite subcollection \( \{(a_j, b_j) : j = 1, \ldots, \ell\} \) of those subintervals that covers \([a, b]\) in the sense that each point \( x \in [a, b] \) is contained in at least one of the \((a_j, b_j)\)'s.

It also helps to make a preliminary observation: the assumption that the operators \( K_n \) applied to 1, \( x \) and \( x^2 \) gave sequences of functions that converged uniformly on \([a, b]\) to those functions respectively implies that for any second-degree polynomial \( p(x) = ax^2 + bx + c \) the same convergence relation holds, because for any \( t \in [a, b] \)

\[
K_n[ax^2 + bx + c](t) = (at^2 + bt + c) = a(K_n[x^2](t) - t^2) + b(K_n[x](t) - t) + c(K_n[1](t) - 1) \\
|K_n[ax^2 + bx + c](t) - (at^2 + bt + c)| \leq |a| |K_n[x^2](t) - t^2| + |b| |K_n[x](t) - t| + |c| |K_n[1](t) - 1|  
\]

and if \( N \) is so large that \( n \geq N \) implies that each of the 3 terms on the r. h. s. of (1.B.3) is \( < \eta/3 \) for all \( t \in [a, b] \)—which is possible by hypothesis—then the l. h. s. will be \( < \eta \) for all \( t \in [a, b] \).

Now for the proof of the theorem, let \( f \in \mathcal{C}_a([a, b]) \) and \( \epsilon > 0 \) be given. For each \( t \in [a, b] \) the fact that \( f \) is continuous at \( t \) lets us find a \( \eta_t > 0 \) such that if \( t - \eta_t < x < t + \eta_t \), then \( |f(x) - f(t)| < \epsilon/3 \). We can then find a quadratic polynomial \( p_t(x) = C_t(x-t)^2 + f(t) + 2\epsilon/3 \) with \( C_t > 0 \) so large that if \( x \leq t - \eta_t \) or \( t + \eta_t \leq x \), then \( p_t(x) \geq M \). [Draw a picture!] But then \( p_t(x) > f(x) \) at every point \( x \in [a, b] \); this is clearly true outside of \((t-\eta_t, t+\eta_t)\) by choice of \( C_t \), while for \( x \) inside that interval we have \( f(x) \leq f(t) + \epsilon/3 < f(t) + 2\epsilon/3 \leq p_t(x) \). Now by continuity of \( x \mapsto p_t(x) - f(x) \), the fact that \( 2\epsilon/3 > p_t(t) - f(t) = \epsilon/3 \) implies that there is a “small” interval \((a_t, b_t)\) containing \( t \) for which \( x \in (a_t, b_t) \Rightarrow 2\epsilon/3 > p_t(x) - f(x) \). Applying the Heine-Borel theorem, we find a finite collection \( \{(a_j, b_j) : j = 1, \ldots, \ell\} \) that covers \([a, b]\) in the sense that each point \( x \in [a, b] \) is contained in at least one of the \((a_j, b_j)\)'s. This says that for every \( x \in [a, b] \) there is some \( p_j \) for which the inequality \( p_j(x) < f(x) + 2\epsilon/3 \) holds. Now in all events \( f(x) < (p_t \wedge \cdots \wedge p_j)(x) \) holds for every \( x \in [a, b] \), and applying one of the \( K_n \)'s to that inequality, we see that

\[
K_n[f] \leq K_n[p_t \wedge \cdots \wedge p_j] \leq K_n[p_j] \wedge \cdots \wedge K_n[p_j]  
\]

holds throughout \([a, b]\). Because each sequence \( \{K_n[p_j]\}_{n=1}^{\infty} \) converges uniformly on \([a, b]\) to \( K_n[p_j] \), we can now find an \( N^+ \) so large that \( n \geq N^+ \Rightarrow K_n[p_j] \leq p_j + \epsilon/3 \) holds for all \( j = 1, \ldots, \ell \). It follows that for every \( x \in [a, b] \) we have

\[
K_n[f](x) \leq K_n[p_t \wedge \cdots \wedge p_j](x) \leq K_n[p_j](x) \wedge \cdots \wedge K_n[p_j](x) \\
\leq |p_j(x) + \epsilon/3| \wedge \cdots \wedge |p_j(x) + \epsilon/3| = \min\{p_j(x), \ldots, p_j(x)\} + \epsilon/3 \ .  
\]

For a particular \( x \in [a, b] \): by considering a particular \( p_j(\cdot) \) for which the inequality \( f(x) < p_j(x) < f(x) + 2\epsilon/3 \) holds, we see that (1.B.4) implies
whenever \( n \geq N^+ \); but that \( x \in [a, b] \) was arbitrary and the choice of \( t_j \) does not appear at either end of the inequality (1.B.5), so (1.B.5) holds for all \( x \in [a, b] \) whenever \( n \geq N^+ \). Applying exactly the same reasoning to the function \(-f\), we find that there exists an \( N^- \) such that

\[
-K_n[f](x) \leq -f(x) + \epsilon
\]

holds for all \( x \in [a, b] \) whenever \( n \geq N^- \). Putting (1.B.5) and (1.B.6) together, we see that if \( N = \max\{N^+, N^-\} \), then

\[
n \geq N \Rightarrow \text{For all } x \in [a, b], \ f(x) - \epsilon < K_n[f](x) < f(x) + \epsilon
\]

and that is precisely the definition (\( \epsilon > 0 \) having been arbitrary, and \( N \) having been determined from \( \epsilon \)) of “\( K_n[f] \) converges to \( f \) uniformly on \([a, b]\).”

This is such a good theorem that many proofs have been given of it—some in almost unrecognizable form, representing a generalization of the theorem to “what is really going on.” While there are versions that work on closed and bounded subsets of an \( \mathbb{R}^k \), these are in some sense the only versions that can exist.(4)

1.C. The Bernˇsteˇın basis. For any natural number \( n \), the Bernˇsteˇın basis(5) of \( P_n \) is the set of polynomials

\[
\left\{ \binom{n}{k} x^k (1-x)^{n-k} \right\}_{k=0}^{n} \subseteq P_n.
\]

While it is clear that each of these is a polynomial of degree exactly \( n \) and that there are the right number of them, it is not clear that they are linearly independent; in fact it will require a little work (which we shall do two \S\S below) to prove that.

The primary interest of the Bernˇsteˇın basis lies in its properties as a subset of the continuous real-valued functions on \([0, 1]\). These functions form a(n infinite-dimensional) vector space \( \mathcal{C}[0, 1] \) under the pointwise operations; because every polynomial is uniquely determined by its values on \([0, 1]\), we can (and do) think of each \( P_n \) as an \((n + 1)\)-dimensional subspace of \( \mathcal{C}[0, 1] \). It is obvious that each element of the Bernˇsteˇın basis is nonnegative on \([0, 1]\), taking the value zero at and only at the endpoints for \( 1 < k < n \); of course \( x^n \) takes the value 1 at \( x = 1 \) and \((1-x)^n = 1 \) at \( x = 0 \).

An element of \( P_n \) written as a linear combination of elements of the Bernˇsteˇín basis, i.e., in the form

\[
p(x) = \sum_{k=0}^{n} a_k \binom{n}{k} x^k (1-x)^{n-k},
\]

is said to be written in Bernˇsteˇín form (we will know that the coefficients \( \{a_k\} \) are uniquely determined by the function \( p(x) \) once we know that the Bernˇsteˇín basis is really a basis of \( P_n \)). Because each element of the Bernˇsteˇín basis is nonnegative on \([0, 1]\), a polynomial in Bernˇsteˇín form with nonnegative coefficients defines a nonnegative function on \([0, 1]\); similarly, if \( a_k \leq b_k \) for each \( k = 0, \ldots, n \), then

\[
\sum_{k=0}^{n} a_k \binom{n}{k} x^k (1-x)^{n-k} \leq \sum_{k=0}^{n} b_k \binom{n}{k} x^k (1-x)^{n-k} \quad \text{for all } 0 \leq x \leq 1.
\]

(5) Named after their discoverer—though in some sense they go back to Newton and the Bernoullis—the Russian mathematician Sergei N. Bernˇsteˇín. I thought of using the Cyrillic font to write his name out and decided that it was too much work.
The elements of the Bernštejn basis obviously have links with probability theory: \( \binom{n}{k} x^k (1-x)^{n-k} \) is the probability that on \( n \) flips of a coin with probability \( x \) of getting heads, one gets heads exactly \( k \) times. In what we do below we shall need three relations that explain the usefulness of the Bernštejn basis in approximating continuous functions: the derivation of these relations will seem most natural if we use the language of probability to discuss it. So we shall replace \( x \) by \( p \) and \( 1-x \) by \( q \) (with \( 0 \leq p \leq 1 \) and thus \( 0 \leq q = 1 - p \leq 1 \)) in the derivations. The fact that the “binomial distribution” \( \Pr(X = k) = \binom{n}{k} p^k q^{n-k} \) is a probability distribution—quite aside from the probabilists’ way to derive it (by adding \( n \) independent Bernoulli random variables with \( \Pr(X_i = 1) = p \) together)—follows from the binomial theorem:

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

If we introduce a parameter \( t \) and consider the function\(^6\)

\[
M_X(t) = (pe^t + q)^n = \sum_{k=0}^{n} e^{kt} \binom{n}{k} p^k q^{n-k}
\]

then it becomes easy to find the expected values of \( X \) and \( X^2 \), \textit{i.e.}, the sums

\[
\sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} \text{ and } \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k};
\]

just differentiate both sides with respect to \( t \) and set \( t = 0 \), then do it again.

\[
\frac{dM_X(t)}{dt} = n(pe^t + q)^{n-1} \cdot pe^t = \sum_{k=0}^{n} ke^{kt} \binom{n}{k} p^k q^{n-k}
\]

\[
n = n(pe^0 + q)^{n-1} \cdot pe^0 = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k}
\]

\[
\frac{d^2M_X(t)}{dt^2} = n(n-1)(pe^t + q)^{n-2} p^2 e^{2t} + n(pe^t + q)^{n-1} pe^t = \sum_{k=0}^{n} k^2 e^{kt} \binom{n}{k} p^k q^{n-k}
\]

\[
n^2 p^2 + npq = n(n-1)p^2 + np = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k}.
\]

Evidently this process could be carried on as many times as one might wish, but the information above will be sufficient\(^7\).

\textbf{1.D. The Bernštejn Polynomials of a Function.} Suppose \( f(x) \) is a bounded function defined for \( 0 \leq x \leq 1 \)—while it may be continuous, it need not be for the early part of the considerations of this §—and set of the the sum of \( n \) independent Bernoulli random variables, each with probability of success \( p \); its moment-generating function is thus the \( n \)-th power of a moment-generating function of a single Bernoulli random variable.

\( ^{7}\) A probabilist could already compute the mean and variance of \( X \) from this information.
We now connect with the Bohman-Korovkin ideas, because a striking fact about these polynomials is that for the functions \( f(x) = 1, x, \) and \( x^2, \) the corresponding Bernstein polynomials either equal the function identically for \( x \in [0, 1] \) or converge to it uniformly on \([0, 1]\) as \( n \to \infty. \) Indeed,

\[
1 = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^k (1-x)^{n-k}
\]

is just the binomial theorem, as we saw in the formula labeled (1) above. If we divide the “expectation” relation \((x)\) by \( n, \) we find for the function \( f(x) = x \) (with \( p \) and \( q \) replaced by \( x \) and \( 1-x \) respectively) that

\[
f(x) \equiv x \equiv \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} x^k (1-x)^{n-k} \equiv \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} = B_n[f](x).
\]

With \( f(x) = x^2 \) we are not quite so lucky, but from

\[
n^2 p^2 + npq = n(n-1)p^2 + np = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k}
\]

we get (by dividing by \( n^2 \))

\[
n^2 x^2 + nx(1-x) = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k}
\]

or

\[
B_n[f](x) = \sum_{k=0}^{n} \left(\frac{k}{n}\right)^2 \binom{n}{k} x^k (1-x)^{n-k} \equiv x^2 + \frac{x(1-x)}{n}.
\]

Thus, although \( B_n[f](x) \equiv f(x) = x^2 \) is false, it is the case that \( B_n[f](x) \to f(x) \) as \( n \to \infty, \) with a uniform error estimate on \([0, 1]\): since the maximum value of \( x(1-x) \) on the interval is \( 1/4, \) we have \( |B_n[f](x) - f(x)| \leq \frac{1}{4n} \) uniformly for all \( x \in [0, 1]. \)

We now need to verify that the mappings \( B_n : \mathcal{C}_\mathbb{R}(0, 1) \to \mathcal{C}(\mathbb{R}[0, 1]) \) satisfy the requirements of the Bohman-Korovkin theorem. We will then have \( \|f - B_n[f]\|_\infty \to 0 \) as \( n \to \infty \) for every \( f \in \mathcal{C}_\mathbb{R}([0, 1]), \) and so the Bernstein polynomials of \( f \) give a sequence of polynomials—with coefficients computable from the values of \( f(x) \)—that will approximate \( f(x) \) as closely as one wishes, provided that their degree \( n \) is sufficiently large.

So we start checking.

\[
f \mapsto B_n[f](x) \text{ is linear: } B_n[\alpha f + \beta g](x) \equiv \alpha B_n[f](x) + \beta B_n[g](x)
\]

holds for any functions \( f(x) \) and \( g(x) \) and any constants \( \alpha, \beta. \) This is obvious from the way that the coefficients of the Bernstein polynomials are defined:

\[
B_n[\alpha f + \beta g](x) \equiv \sum_{k=0}^{n} \left[ \alpha f\left(\frac{k}{n}\right) + \beta g\left(\frac{k}{n}\right) \right] \binom{n}{k} x^k (1-x)^{n-k}
\]

\[
= \alpha \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} + \beta \sum_{k=0}^{n} g\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}
\]

\[
= \alpha B_n[f](x) + \beta B_n[g](x).
\]

It is similarly obvious that \( f \mapsto B_n[f](x) \) is \textbf{positive,} or order-preserving:

\[
f(x) \leq g(x) \text{ for all } x \in [0, 1] \Rightarrow B_n[f](x) \leq B_n[g](x).
\]
Indeed, we have for each \(0 \leq k \leq n\)

\[
f\left(\frac{k}{n}\right) \leq g\left(\frac{k}{n}\right)
\]

\[
f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} \leq g\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}
\]

\[
B_n[f](x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} \leq \sum_{k=0}^{n} g\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} = B_n[g](x)
\]

The inequalities do not reverse on being multiplied by \(x^k (1-x)^{n-k}\), which is nonnegative provided that \(0 \leq x \leq 1\).

We have thus checked the hypotheses of the Bohman-Korovkin theorem, and so we can assert

**Theorem [Bernstein]:** For any continuous function \(f \in C([0,1])\), the sequence of polynomials \(\{B_n[f](x)\}_{n=0}^{\infty}\) converges to \(f(x)\) uniformly for \(x \in [0,1]\), as \(n \to \infty\).

This is, of course, the Weierstraß approximation theorem on the interval \([0,1]\), but in a computable form: if one can evaluate the function at \(0, 1/n, \ldots, (n-1)/n, 1\) then one can compute \(B_n[f]\) explicitly. By translation and a change of scale a function on any interval \([a, b]\) can be transformed into a function on \([0,1]\), approximated by a polynomial, and then transformed back; so this theorem is no less general than the full Weierstraß approximation theorem.

It may be interesting to know that Bernstein’s proof of this theorem, rather than using the Bohman-Korovkin construction, used probabilistic reasoning explicitly: the convergence of \(B_n[f](x)\) to \(f(x)\) at each point follows quite readily from the Chebyshev inequality (or weak law of large numbers) of probability theory, and the estimates can be made uniform for \(x \in [0,1]\). Indeed, the title of Bernstein’s original 1912 paper was *Démonstration du théorème de Weierstrass, fondé sur le calcul des probabilités.*

It is not difficult to check that even if the bounded function \(f\) is not everywhere continuous, the sequence \(\{B_n[f](x)\}_{n=0}^{\infty}\) of values of the polynomials still converges to \(f(x)\) at any point at which \(f\) is continuous. When \(f(x)\) is not continuous at \(x_0\) but the one-sided limits \(\lim_{x \to x_0^-} f(x) = L_-\) and \(\lim_{x \to x_0^+} f(x) = L_+\) exist, it is possible to show that \(\lim_{n \to \infty} B_n[f](x_0)\) exists and equals the average \((L_- + L_+)/2\); the easiest approach to that fact involves the central limit theorem of probability theory.

Note that in general very large \(n\) is necessary for good approximation: if \(f(x) = x^2\) the error in \(B_n[f](x)\) is still \(1/4n\) even though \(f(x)\) is itself a polynomial of small degree. Unlike interpolating polynomials, Bernstein approximators can be of large degree even when they (poorly!) approximate polynomials of small degree. There is specific quantitative information on this: a theorem of E. Woronovskaja, of which we shall not give a proof, shows that

\[
\lim_{n \to \infty} n \cdot \{f(x) - B_n[f](x)\} = -\frac{f''(x)}{2} x(1-x)
\]

at every point at which \(f\) has a second derivative, and thus the behavior of \(x^2\) is almost typical. In particular, one can expect \(\|f - B_n[f]\|_{\infty} = O(1/n)\) except in the trivial case where \(f(x) \equiv ax + b\) and \(B_n[f] \equiv f\) for all \(n \in \mathbb{N}\). Nonetheless, the Bernstein polynomials still have concrete uses because they inherit good properties from their function, as we shall see in the next §.

**1.E. Bernstein Polynomials Conform to the Shape of Their Functions.** The heading needs some explanation, but it will be most easily explained after we perform the following simple task: given a polynomial in Bernstein form, compute its derivative.
\[
\frac{d}{dt} \left[ \sum_{k=0}^{n} a_k \binom{n}{k} x^k (1-x)^{n-k} \right] = \\
\sum_{k=0}^{n} a_k \binom{n}{k} k x^{k-1} (1-x)^{n-k} - \sum_{k=0}^{n} a_k \binom{n}{k} x^k (n-k) (1-x)^{n-k-1} = \\
\sum_{k=0}^{n-1} a_{k+1} \binom{n}{k+1} (k+1) x^{k+1} (1-x)^{n-k-1} - \sum_{k=0}^{n-1} a_k \binom{n}{k} x^k (n-k) (1-x)^{n-k-1} = \\
\sum_{k=0}^{n-1} a_{k+1} \frac{n!}{(k+1)! (n-k-1)!} (k+1) x^k (1-x)^{n-k-1} \\
- \sum_{k=0}^{n-1} a_k \frac{n!}{k! (n-k)!} x^k (n-k) (1-x)^{n-k-1} = \\
\sum_{k=0}^{n-1} \left[ a_{k+1} \frac{n!}{k! (n-1-k)!} - a_k \frac{n!}{k! (n-1-k)!} \right] x^k (1-x)^{n-1-k} = \\
\sum_{k=0}^{n-1} [n \cdot (a_{k+1} - a_k)] \binom{n-1}{k} x^k (1-x)^{n-1-k} .
\]

In the case in which the Bernstein-form polynomial is the Bernstein polynomial of a function \( f(x) \), so that \( a_k = f(k/n) \), the form of the coefficient in the derivative should be familiar:

\[
\frac{n}{n+1} f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right) = \frac{f(k+1)/n - f(k/n)}{1/n} = f\left[ \frac{k}{n}, \frac{k+1}{n} \right]
\]

is a first divided difference of \( f(x) \) tabulated at the values \( x = 0, 1/n, 2/n, \ldots, 1 \). Thus if \( f(x) \) is an increasing function (it need not be differentiable) on \([0, 1]\), each of these coefficients is nonnegative: but then \( B_n[f]'(x) \geq 0 \) and so \( B_n[f](x) \) is also an increasing function on \([0, 1]\). The same is true for decreasing functions. This begins to show that the Bernstein polynomials of a function tend to preserve the “shape” of its graph. Note that if \( f(x) \) is differentiable, then each \( f[k/n, (k+1)/n] = f'(\xi_k) \) for a suitable \( k/n < \xi_k < (k+1)/n \) by application of the MVTDC, \( k = 0, \ldots, n-1 \). (Using this fact, it is not difficult to prove that if \( f \in C^1[0, 1] \) then \( B_n[f]'(x) \to f'(x) \) uniformly on \([0, 1]\), and similarly that if \( f \in C^k[0, 1] \) then \( B_n[f]^{(j)}(x) \to f^{(j)}(x) \) uniformly on \([0, 1]\) for \( 0 \leq j \leq k \). For details, see the book of G. G. Lorentz cited in \S 1.F below.)

Before we look at derivatives more closely, however, we need to clean up the following matter.

**Proposition:** The Bernstein basis \( \left\{ \binom{n}{k} x^k (1-x)^{n-k} \right\}_{k=0}^{n} \) is a basis of \( \mathcal{P}_n \).

**Proof.** Because they have exactly the right number of elements, it will suffice to show that each Bernstein basis is linearly independent. This is obvious for small \( n = 0 \) or \( 1 \). If linear independence fails for some \( n \), then there will be a smallest \( n > 1 \) for which it fails: let \( \sum_{k=0}^{n} a_k \binom{n}{k} x^k (1-x)^{n-k} = 0 \) be a nontrivial linear dependence relation for that \( n \). Differentiating this relation gives \( n \cdot \sum_{k=0}^{n-1} [a_{k+1} - a_k] \binom{n-1}{k} x^k (1-x)^{n-1-k} = 0 \), and by the minimality of \( n \) this linear dependence is possible only with zero coefficients; so \( a_{k+1} - a_k = 0 \) for all \( k = 0, \ldots, n-1 \). This says that all the coefficients in the linear dependence must be equal, say \( a_k = a \) for all \( k = 0, \ldots, n \). The linear dependence then takes the form

\[
0 \equiv \sum_{k=0}^{n} a_k \binom{n}{k} x^k (1-x)^{n-k} = a \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = a \cdot 1
\]

642:573:01 Fall 2000
by the binomial theorem; so all the \( a_k = a = 0 \) and the linear dependence was trivial contrary to assumption.\(^{(8)}\)

Now back to derivatives: a small inconvenience attached to the polynomial

\[
B_n[f]^n(x) = \sum_{k=0}^{n-1} n \cdot \left[ f\left( (k+1)/n \right) - f(k/n) \right] \binom{n-1}{k} x^k (1-x)^{n-1-k}
\]

is that it is not the Bernšteǐn polynomial of the function \( f'(x) \). Thus we cannot quite iterate the process that gave us values of the derivative of \( f(x) \) as coefficients in the Bernšteǐn form of \( B_n[f]'(x) \). There is no such problem if we just look at polynomials in Bernšteǐn form, however: for example

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

The coefficients are (up to factors involving \( n \)) “(higher) tabular differences” of the coefficients of the original polynomial. In the case of the Bernšteǐn polynomial \( B_n[f](x) \) of a function it is not difficult to work out the general form of the coefficients; we shall be satisfied to observe that for a second derivative we have

\[
B_n[f]''(x) = \frac{f((k+2)/n) - f((k+1)/n)}{1/n} - \frac{f((k+1)/n) - f(k/n)}{1/n}
\]

and since the second divided difference on the l. h. s. of that can be written as a value \( f''(\eta_k)/2! \), the coefficients in \( B_n[f]'' \) are values of \( f'' \) except for the multiplier \( n(n-1) \). Thus if \( f(x) \) is convex, so \( f''(x) \geq 0 \), then the same is true of \( B_n[f](x) \) (even without \( f(x) \) being twice-differentiable).

So if \( f(x) \) is an increasing function so is each \( B_n[f](x) \); if \( f''(x) \geq 0 \) the same is true of \( B_n[f](x) \), etc. Thus the graphs of Bernšteǐn approximators tend to have a “shape” that resembles that of the function they approximate.

1.1. Applications and Other Information. Unfortunately, the very slow convergence of \( B_n[f] \) to \( f \) — only \( O(1/n) \) even for quadratic polynomials—has made their use as an approximating scheme somewhat impractical in many situations. They are nonetheless to be preferred in many graphical applications because of their ability to follow the shape of the graph of a function. It is reasonable to hope that vector- and parallel-processing algorithms can be developed that will make real-time computation of their values less time-consuming than it is by ordinary serial techniques.

Bernšteǐn polynomials, having arisen in connection with probability theory, lend themselves to probability applications. For example, if \( X \) is an empirical random variable with values in \([0,1]\)—e.g., a table of percentage scores between 1 and 100—then its cumulative distribution function \( F_X(x) \) will be a nondecreasing function on \([0,1]\) with \( F_X(0) = 0 \) and \( F_X(1) = 1 \). Each \( B_n[F_X](x) \) will be nondecreasing and take the values 0 at 0 and 1 at 1; as \( n \) gets larger the Bernšteǐn polynomials will increasingly resemble the empirical distribution, but they will be continuously distributed approximations. Their derivatives (set identically equal to 0 outside \([0,1]\)) will then be civilized-looking probability density functions, each of which is a convex combination of beta densities. This fact should be suggestive.

Additional material on Bernšteǐn polynomials can be found in such places as S. Karlin, Total Positivity, v. 1, Stanford Univ. Press (1968); in G. G. Lorentz, Bernstein Polynomials, 2nd ed., Chelsea Pub. Co. (1986); and scattered around in various sources of material on approximation theory. (See, e.g., E. Ward Cheney, Introduction to Approximation Theory, 2nd ed., Chelsea Pub. Co. (1982), §3.3, especially the problems.) Unfortunately, one has to go to the journal literature for much beyond the now-standard Bernšteǐn-polynomial proof, given above, of the Weierstraß approximation theorem.

\(^{(8)}\) This strong induction could be rephrased as a straight induction with little work; any interested reader is welcome to make the effort.
2. Inner-Product and Hilbert Spaces.

2.A. Introduction. This is more introduction than the reader probably needs to this subject, but since I had it available in a more-or-less prepared form I saw no reason not to circulate it. The material that is most useful in numerical analysis is flagged with a bold-face star (*); the rest can be read in a rather cursory manner, at least at the first time. Material marked with (X), on the other hand, can pretty much be passed over. Do not read about only the real-scalar case, however: the Fourier transform on \([0, 2\pi]\), the unit circle in \(\mathbb{C}\) and the “finite circles” that underlie the fast-Fourier-transform algorithm make transparent sense over \(\mathbb{C}\) but are much less transparent over \(\mathbb{R}\).

(*) Let \(H\) be a vector space over \(K = \mathbb{R}\) or \(\mathbb{C}\). An inner product on \(H\) or positive definite sesquilinear function on \(H \times H\) is a function

\[
\langle \cdot, \cdot \rangle : H \times H \to K
\]

satisfying the following axioms:

1. For every \(x, y, z \in H\), \(\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle\);
2. For every \(x, y \in H\) and scalar \(\alpha \in K\), \(\langle \alpha x, y \rangle = \alpha \cdot \langle x, y \rangle\);
3. For every \(x, y \in H\), \(\langle x, y \rangle = \overline{\langle y, x \rangle}\), where the overbar denotes complex conjugation (and is thus vacuous if the scalar field \(K = \mathbb{R}\));
4. For every \(x \in H\), \(\langle x, x \rangle \geq 0\), and equality holds if and only if \(x = 0\) as an element of \(H\).

There is a fifth requirement, but one has to prove some preliminary results in order to state it. In the meantime, we can observe that it makes sense to define \(\|x\|\) to equal \(\sqrt{\langle x, x \rangle}\), even though we do not know that this function \(H \to \mathbb{R}^+\) actually has any of the properties of a norm.

(X) It costs nothing to make a small weakening of these requirements as we prove those preliminary results. A semi-inner product on \(H\) or positive semi-definite sesquilinear function on \(H \times H\) is a function \(\langle \cdot, \cdot \rangle : H \times H \to K\) satisfying all the requirements above except that (4) is replaced by

\(\langle \cdot, \cdot \rangle \) For every \(x \in H\), \(\langle x, x \rangle \geq 0\).

That is, there may be “null vectors” \(x \in H\) for which \(\|x\|^2 = \langle x, x \rangle = 0\) without \(x = 0\) holding in the vector-space structure of \(H\). Even with this weak requirement we have

(*) Proposition [Schwarz inequality]: For any \(x, y \in H\) the relation

\[
\|\langle x, y \rangle\| \leq \|x\|^2 \|y\|^2
\]

holds, with equality if and only if there is a scalar \(\alpha \in K\) for which one of \(x - \alpha y\) or \(y - \alpha x\) is a null vector (i.e., \(\|x - \alpha y\| = 0\) or \(\|\alpha x - y\| = 0\)).

Proof. If \(\langle x, y \rangle = 0\) there is nothing to prove: the inequality is correct, and the case of equality makes one of \(x\) or \(y\) a null vector. Otherwise, let \(e^{i\theta}\) be a complex number of modulus \(1\) for which \(\langle x, e^{i\theta} y \rangle > 0\) (if the scalars are the reals, then \(e^{i\theta} = \pm 1\)), and let \(z = e^{i\theta} y\). Consider the function

\[
\lambda \mapsto \langle x - \lambda z, \lambda z \rangle
\]

\[\mathbb{R} \to \mathbb{R}^+\]

It follows from the axioms—even without the preliminary preparation of \(z\)—that \(\langle x - \lambda z, x - \lambda z \rangle = \langle x, x \rangle - \lambda(x, z) - \lambda(z, x) + \lambda^2 \langle z, z \rangle\), or

\[0 \leq \langle x - \lambda z, x - \lambda z \rangle = \|x\|^2 - 2\lambda \text{Re} \langle x, z \rangle + \lambda^2 \|z\|^2. \tag{2.A.1}\]

Since \(\langle x, z \rangle > 0\), this expression further simplifies to \(0 \leq \|x\|^2 - 2\lambda \|z\|^2\). If \(\|z\| = 0\), then this expression defines a first-degree polynomial function of \(\lambda \in \mathbb{R}\) with nonzero coefficient \(-2\langle x, z \rangle\) of \(\lambda\), and
it cannot be nonnegative for all \( \lambda \in \mathbb{R} \), so \( z \) cannot be a null vector. This expression must therefore be a polynomial of second degree in \( \lambda \), and so it must take a minimum value at \( \lambda = \frac{\langle x, z \rangle}{\|z\|^2} \) (one can find this out using calculus, or by completing the square). Plugging in this minimizing value of \( \lambda \) results in the relation

\[
0 \leq \|x\|^2 - 2\frac{(\langle x, z \rangle)^2}{\|z\|^2} + \frac{(\langle x, z \rangle)^2}{\|z\|^2} \leq \|x\|^2
\]

\[
\langle x, z \rangle)^2 \leq \|x\|^2 \quad \text{(2.A.2)}
\]

Since our choice of \( z \) made \( \langle x, z \rangle = |\langle x, y \rangle| \), (2.A.2) is equivalent to the Schwarz inequality. If equality holds in (2.A.2), then tracing back through the derivation of (2.A.2) we see that equality must also have held in (2.A.1), so \( x - \lambda z = x - \lambda e^{i\theta} y \) is a null vector.

\[\text{(*) Proposition [Triangle Inequality]: For any } x, y \in H, \quad \|x + y\| \leq \|x\| + \|y\| \]  

\[\text{Proof.} \quad \|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle y, y \rangle + \langle y, x \rangle + \langle x, y \rangle = \|x\|^2 + 2\text{Re} \langle x, y \rangle + \|y\|^2 \leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \text{ Take square roots.} \]

\[\text{(X) Corollary: The null vectors form a subspace } N \subseteq H. \text{ The semi-inner product is well-defined on cosets of } N \text{ and thus defines a positive definite inner product on the elements of the quotient vector space } H/N. \]

\[\text{Proof. The fact that the null vectors form a subspace follows from the homogeneity of } x \mapsto \|x\| \text{ and the triangle inequality. If } m, n \in N \text{ are null vectors then for any } x, y \in H \]

\[\langle x + m, y + n \rangle = \langle x, y \rangle + \langle x, n \rangle + \langle m, y \rangle + \langle m, n \rangle \]

and all terms on the r. h. s. except perhaps the first are zero by the Schwarz inequality. In particular, all representatives \( x \) of the same coset of \( N \) give the same value for \( \|x\| \), so the norm on the cosets distinguishes cosets that are not \( N \) from \( N \) (the zero element of \( H/N \)), and the induced norm on \( H/N \) is thus definite (and it is a genuine norm rather than a semi-norm).

So we can now throw on the fifth requirement:

\[\text{(5) The normed space } (H, \|\cdot\|) \text{ is complete; that is, if } \{x_n\}_{n=0}^\infty \text{ is a sequence of vectors with the property that for every } \epsilon > 0 \text{ there is an } N \in \mathbb{N} \text{ such that } m, n \geq N \Rightarrow \|x_n - x_m\| < \epsilon, \text{ then there is a vector } x \in H \text{ for which } \lim_{n \to \infty} \|x - x_n\| \to 0. \]

It is fairly routine to use the Schwarz inequality to show that \( \langle x, y \rangle \mapsto \langle x, y \rangle \) is continuous from \( H \times H \) to \( \mathbb{R} \). One can also observe that if \( H \) is a vector space over \( \mathbb{C} \) then it is also a vector space over \( \mathbb{R} \), and that if \( \langle \cdot, \cdot \rangle \) is a complex inner product then \( \langle x, y \rangle \mapsto \text{Re} \langle \cdot, \cdot \rangle \) is a real-valued inner product on \( H \) that satisfies the real versions of the axioms and which defines the same norm. The resulting real Hilbert space is called the \textbf{underlying real Hilbert space} of the given complex Hilbert space. This fact makes it possible to “think real” when one is doing arguments involving only the norm (if one wishes to do that).

The primary examples of Hilbert spaces are the spaces \( L^2([a, b], w) \), where \([a, b] \subseteq \mathbb{R}\) is an interval and \( w(\cdot) \) is a weight function, as in Atkinson, p. 208; their inner product is \( \langle f, g \rangle = \int_a^b f(x)g(x)w(x) dx \) (where as usual the overbar is vacuous if the functions are real-valued). The axioms are theorems of measure theory; the Schwarz and triangle inequalities are the Hölder and Minkowski inequalities in the case \( p = 2 \). The Euclidean spaces \( \ell^2(n) \) for real scalars, as well as their “finite-dimensional unitary space” complex counterparts, are the examples that everyone knows. Some insight into what an underlying real Hilbert
space is can be gained by thinking of $\mathbb{C}^n$ as $n$-tuples $(\xi_1 + i\eta_1, \ldots, \xi_n + i\eta_n) = (\xi_1, \ldots, \xi_n) + i(\eta_1, \ldots, \eta_n)$; under the obvious correspondence with $\mathbb{R}^{2n}$ we have

$$\langle (\alpha_1 + i\beta_1, \ldots, \alpha_n + i\beta_n), (\xi_1 + i\eta_1, \ldots, \xi_n + i\eta_n) \rangle = \sum_{j=1}^{n} (\alpha_j + i\beta_j)(\xi_j - i\eta_j) = \sum_{j=1}^{n} (\alpha_j\xi_j + \beta_j\eta_j) + i\sum_{j=1}^{n} (\beta_j\xi_j - \alpha_j\eta_j)$$

from which one sees that the real part of the complex inner product “is” the real inner product on $\mathbb{R}^{2n}$. It is not absurdly simple—and it is frequently useful—to observe that if $y$ is a unique point that there are no compactness hypotheses: all that matters is (Cauchy-) completeness.

The relation $\|x + y\|^2 = \langle x, x \rangle + 2\Re \langle x, y \rangle + \|y\|^2$ is, of course, the law of cosines. One says that $x$ and $y$ are orthogonal or perpendicular if $\langle x, y \rangle = 0$, and writes $x \perp y$. If $x \perp y$ then the law of cosines becomes the Pythagorean theorem $\|x + y\|^2 = \|x\|^2 + \|y\|^2$. It is not absurdly simple—and it is frequently useful—to observe that $x = 0$ if and only if $\langle x, y \rangle = 0$ holds for all $y \in H$: one way is true by linearity, while the other way is true because one can take $y = x$ to get $0 = \langle x, x \rangle = \|x\|^2$.

The simple Hilbert-space relation with the most profound consequences is the parallelogram law, which asserts that the sum of the squares of the diagonals of a parallelogram is the sum of the squares of the sides:

**Proposition** [Parallelogram Law]: For any $x, y \in H$,

$$\|x + y\|^2 + \|x - y\|^2 = 2 \left[ \|x\|^2 + \|y\|^2 \right].$$

**Proof.**

$$\|x + y\|^2 = \|x\|^2 + 2\Re \langle x, y \rangle + \|y\|^2$$

$$\|x - y\|^2 = \|x\|^2 - 2\Re \langle x, y \rangle + \|y\|^2$$

$$\|x + y\|^2 + \|x - y\|^2 = 2 \left[ \|x\|^2 + \|y\|^2 \right].$$

The parallelogram law is the fundamental lemma in proving the nearest-point theorem. The consequences of that theorem are enormous, beginning with a sizeable fraction of the known existence proofs for solutions of the boundary-value problems of differential equations. What is amazing about this theorem is that there are no compactness hypotheses: all that matters is (Cauchy-) completeness.

**Nearest-Point Theorem** [F. Riesz]: If $K \subseteq H$ is a (norm-metric-)closed convex set and $x \in H$, then there is a unique point $p \in K$ nearest to $x$. This point is characterized by the property that for any other $y \in K$, $\Re \langle x - p, y - p \rangle \leq 0$.

**Proof.** Let $\{p_j\}_{j=1}^{\infty}$ be a minimizing sequence for dist$(x, K)$, i.e., a sequence for which $\|p_j - x\| \to \text{dist}(x, K)$. Applying the parallelogram law to the differences $p_j - x$ and $p_k - x$ gives

$$\|p_j - x + p_k - x\|^2 + \|p_j - p_k\|^2 = 2 \left[ \|p_j - x\|^2 + \|p_k - x\|^2 \right]$$

$$\|p_j - p_k\|^2 = 2 \left[ \|p_j - x\|^2 + \|p_k - x\|^2 \right] - \|p_j + p_k - 2x\|^2$$

$$= 2 \left[ \|p_j - x\|^2 + \|p_k - x\|^2 \right] - 4 \left[ \frac{p_j + p_k}{2} - x \right]^2$$

$$= 2 \left\{ \left[ \|p_j - x\|^2 - \left[ \frac{p_j + p_k}{2} - x \right]^2 \right] + \left[ \|p_k - x\|^2 - \left[ \frac{p_j + p_k}{2} - x \right]^2 \right] \right\}.$$
For \( j, k \geq N \to \infty \), the first term in each of the square brackets on the r. h. s. of the last set-off line approaches \( \text{dist}(x, K) \), while the second (subtracted) term must be \( \geq \text{dist}(x, K) \): the midpoint of \( p_j \) and \( p_k \) belongs to \( K \) since \( K \) was assumed to be convex. The square-bracketed quantities are thus \( \leq \|p_j - x\|^2 - \text{dist}(x, K)^2 \to 0 \) and \( \leq \|p_k - x\|^2 - \text{dist}(x, K)^2 \to 0 \) respectively; so the r. h. s. can be made as small as one pleases for \( j, k \geq N \) by taking \( N \in \mathbb{N} \) sufficiently large. Looking at the l. h. s., we see that the sequence \( \{p_j\}_{j=1}^\infty \) is therefore Cauchy, and thus convergent to a point \( p \in K \) at which the distance \( \|p - x\| = \text{dist}(x, K) \).

That proves the existence of a point \( p \in K \) at minimum distance from \( x \): the point is unique because if \( q \in K \) were another point at minimum distance from \( x \), the sequence \( p, q, p, q, \ldots \) would be a minimizing sequence and thus converge to a point at minimum distance—but of course it could converge only if \( p = q \). Finally, the point \( p \in K \) at minimum distance from \( x \) is characterized by the property that for any (other) \( y \in K \), \( \text{Re}[\langle x - p, y - p \rangle] \leq 0 \) because for any \( y \in K \) we may parametrize the line segment joining \( p \) to \( y \) by \( \lambda \mapsto (1 - \lambda)p + \lambda y \) for \( \lambda \in [0, 1] \), and then consider

\[
\lambda \mapsto \|x - [(1 - \lambda)p + \lambda y]\|^2 = \|(x - p) - \lambda(y - p)\|^2 = \|x - p\|^2 - 2\lambda\text{Re}[\langle x - p, y - p \rangle] + \lambda^2\|y - p\|^2.
\]

This function of \( \lambda \in [0, 1] \) will decrease at \( \lambda = 0 \), violating the minimizing property of \( p \in K \), unless the coefficient of \( \lambda \) is nonnegative, i.e., unless \( \text{Re}[\langle x - p, y - p \rangle] \leq 0 \). On the other hand, suppose \( p \in K \) is a point for which \( \text{Re}[\langle x - p, y - p \rangle] \leq 0 \) holds for every point \( y \in K \). Then the same computation shows that the function \( \lambda \mapsto \|x - [(1 - \lambda)p + \lambda y]\|^2 \) has a derivative everywhere nonnegative in \( [0, 1] \), so \( \|x - p\|^2 \leq \|x - y\|^2 \) for every \( y \in K \), and the inequality will be strict (at \( \lambda = 1 \)) unless \( y = p \). But that precisely says that \( p \in K \) is the (unique) point of \( K \) at minimum distance from \( x \).

The uniqueness assertions of the nearest-point theorem tells us that for any closed convex \( K \subseteq H \) we can define a function \( P : H \to K \) by mapping each \( x \in H \) to the nearest point \( Px \in H \). As things turn out, this function is not merely continuous, but distance-decreasing: it satisfies a Lipschitz condition with Lipschitz constant 1.

**Proposition** Let \( K \subseteq H \) be a closed convex set. The function \( x \mapsto Px \) which assigns to each \( x \in H \) the nearest point \( Px \in K \) satisfies \( \|Px - Px_0\| \leq \|x - x_0\| \) for each pair \( x, x_0 \) of points of \( H \).

**Proof.** Let \( x, x_0 \) be two points of \( H \) and \( p = Px, p_0 = Px_0 \) the corresponding nearest points of \( H \). Then the inner-product characterization of the nearest point, applied first to \( p \) and then to \( p_0 \), gives

\[
\text{Re}\{\langle x - p, p_0 - p \rangle\} \leq 0
\]
\[
\text{Re}\{\langle x_0 - p_0, p - p_0 \rangle\} \leq 0
\]
\[
\text{Re}\{\langle p - p_0 - (x - x_0), p - p_0 \rangle\} \leq 0
\]
\[
\|p - p_0\|^2 = \langle p - p_0, p - x_0 \rangle \leq \text{Re}\{\langle x - x_0, p - p_0 \rangle\} \leq \|x - x_0\| \|p - p_0\| \quad \text{(Schwarz inequality)}
\]
\[
\|p - p_0\| \leq \|x - x_0\|
\]

(denominator by zero is no problem, since if \( \|p - p_0\| = 0 \) there is nothing to prove).

Another direct consequence of the nearest-point theorem is the **dual geometric characterization of closed convex sets**. We make a standard

**Definition:** A closed (real) **half-space** in a Hilbert space \( H \) is a set of the form

\[
\{y \in H : \text{Re}\{\langle y_0, y \rangle\} \leq \alpha\},
\]

where \( \alpha \in \mathbb{R} \) and \( y_0 \neq 0 \). (Evidently such sets could be equivalently characterized as \( \{y \in H : \text{Re}\{\langle y_0, y \rangle\} \leq \alpha\} \), or by facing the inequality the other way as \( \{y \in H : \text{Re}\{\langle y_0, y \rangle\} \geq \alpha\} \), etc. Similarly, an open (real) **half-space** in a Hilbert space \( H \) is a set of the form \( \{y \in H : \text{Re}\{\langle y_0, y \rangle\} < \alpha\} \), where \( \alpha \in \mathbb{R} \) and \( y_0 \neq 0 \). Evidently closed half-spaces are closed and open half-spaces are open.

**Proposition:** Any closed convex set \( K \) in a Hilbert space \( H \) equals the intersection of the closed half-spaces that contain it. The smallest closed convex set that contains a given \( \emptyset \neq S \subseteq H \) is the intersection of all the half-spaces that contain \( S \).
Proof. If $K = H$ we invoke the convention that an empty intersection is the whole works. Assuming $K$ is proper, certainly $K$ is contained in that intersection. On the other hand, if $x_0 \notin K$ then $x_0 \neq P x_0$, and if we set $y_0 = x_0 - P x_0 \neq 0$, then for every $y \in K$ we have $\text{Re} \left[ \langle x_0 - P x_0, y - P x_0 \rangle \right] \leq 0$, or equivalently $\text{Re} \left[ \langle y_0, y \rangle \right] \leq \langle y_0, P x_0 \rangle$, and we may define $\alpha$ to be the (constant) r. h. s. of that relation. If $\emptyset \neq S \subseteq H$, then the intersection of the closed half-spaces that contain $S$ is convex (as an intersection of convex sets always is), closed, and contains $S$. If $K$ is a closed convex set that contains $S$, then—as we just saw—any point not in $K$ is excluded by some closed half-space that contains $K$ and therefore contains $S$, so the intersection of the closed half-spaces containing $S$ is contained in $K$.

This proposition, which is “geometrically evident” in $\mathbb{R}^2$ and $\mathbb{R}^3$, is not significantly easier to prove in the plane—the reader might try to give a proof.

Some of the most useful cases of these propositions occur when the set “$K$” is a (norm-)closed vector subspace of $H$.

(*) Proposition: Let $M \subseteq H$ be a norm-closed vector subspace of a Hilbert space $H$. Then

1. For any $x \in H$, the nearest point $P x \in M$ to $x$ is characterized by $x - P x \perp M$; it is the “foot of the perpendicular to $M$ dropped from $x$.”

2. The mapping $x \mapsto P x$ is linear and continuous: as an element of $\mathcal{L}(H)$, its norm $\|P\| = 1$ except when $M = \emptyset$ (in which case $P = 0$ and $\|P\| = 0$). For reasons that will appear shortly, it is called the orthogonal projection of $H$ onto $M$.

3. The mapping $x \mapsto x - P x = (I - P)x$ is the nearest-point mapping onto $M^\perp = \{ y \in H : \langle x, y \rangle = 0 \text{ for all } y \in M \}$, a subspace called the orthogonal complement of $M$. It thus has the same properties that were asserted of $P$ in (2) above.

4. The Hilbert space $H = M \oplus M^\perp$ in the following inner-product sense: $H$ is the algebraic direct sum of these two spaces, and the inner product on $H$ is given by $\langle x, y \rangle = \langle Px, Py \rangle + \langle (I - P)x, (I - P)y \rangle$ for any pair of vectors $x, y \in H$. In particular, $\|x\|^2 = \|Px\|^2 + \|(I - P)x\|^2$.

5. If $M \subseteq H$ is a subspace, then its norm closure $\overline{M}$ equals $[M^\perp]^\perp$.

Proof. Of (1): Let $p = P x \in M$ be the point nearest $x$, then $\text{Re} \left[ \langle x - p, y - p \rangle \right] \leq 0$ for every $y \in M$; but since $M$ is a vector subspace, one can choose $y = \alpha z + p$ for any $\alpha \in \mathbb{K}$ and any $z \in M$, so one gets $\text{Re} \left[ \langle x - Px, \alpha z \rangle \right] = \text{Re} \left[ \alpha \langle x - Px, z \rangle \right] \leq 0$ for every $z \in M$. One can choose $\alpha = e^{i\theta}$ in such a way that $\text{Re} \left[ \alpha \langle x - Px, z \rangle \right] = \langle x - Px, z \rangle$, so the only possibility is that $\langle x - Px, z \rangle = 0$. On the other hand, if $p \in M$ is such that $\langle x - p, z \rangle = 0$ for every $z \in M$, then $\text{Re} \left[ \langle x - p, y - p \rangle \right] \leq 0$ holds a fortiori for every $y \in M$; thus the orthogonality relation characterizes the nearest point of $M$. Of (2): If $x_1, x_2$ are two vectors in $H$ and one forms $(x_1 + x_2) - (Px_1 + Px_2)$, then for any $z \in M$ one has $\langle (x_1 + x_2) - (Px_1 + Px_2), z \rangle = \langle x_1, z \rangle + \langle x_2, z \rangle - \langle Px_1, z \rangle - \langle Px_2, z \rangle = 0 + 0 = 0$, so $P(x_1 + x_2) = Px_1 + Px_2$; passing a scalar through $P$ is even easier. We already know that $P$ is distance-decreasing, and so $\|x\| = \|x - 0\| \geq \|Px - P0\| = \|Px\|$. Thus $\|P\| \leq 1$; on the other hand, if $M \neq \emptyset$ then for every nonzero $x \in M$ we have $Px = x \neq 0$ so $\|P\| \geq 1$. Of (3) If $y \in M^\perp$ then for any $x \in H$ we have $0 = \langle Px, y \rangle = \langle x - (x - Px), y \rangle$, so $x - Px \in M^\perp$ has the property that characterizes the point of $M^\perp$ nearest to $x$. Thus $x \mapsto x - Px = (I - P)x$ is the nearest-point map of $H$ onto $M^\perp$. Of (4): We know that $x - Px \in M^\perp$, so for any $x \in H$ writing $x = Px + (x - Px)$ writes $x$ as a sum of a vector in $M$ and a vector in $M^\perp$. This “resolution into orthogonal components” is unique because if $x = x_1 + x_2$ and $x = z_1 + z_2$ are two competing resolutions of $x$, we must have $z = x_1 - z_1 = z_2 - x_2 \in M \cap M^\perp$ and so $\|z\|^2 = \langle z, z \rangle = 0$. For any two elements $x, y \in H$ we thus have

$$\langle x, y \rangle = \langle Px + (x - Px), Py + (y - Py) \rangle = \langle Px, Py \rangle + \langle (I - P)x, (I - P)y \rangle$$

since the other two inner products equal zero by orthogonality. Of (5): It is obvious (by continuity) that $\overline{M}$ is a subspace, that $[\overline{M}]^\perp = M^\perp$, and thus that $\overline{M} \subseteq [M^\perp]^\perp$. By (5) we can write $H = \overline{M} \oplus M^\perp$. Any $x \in [M^\perp]^\perp$ is a limit of convergent sequences in $M^\perp$. This condition means that $M$ contains the limit of any convergent (in the sense of the Hilbert-space norm) sequence of its elements. In almost every case we apply this proposition, the subspace $M$ will be finite-dimensional, and such subspaces are automatically closed.

\footnote{This condition means that $M$ contains the limit of any convergent (in the sense of the Hilbert-space norm) sequence of its elements. In almost every case we apply this proposition, the subspace $M$ will be finite-dimensional, and such subspaces are automatically closed.}
can be resolved into \( x = x_1 + x_2 \) with \( x_1 \in M^\perp \) and \( x_2 \in M^\perp \); but then \( x - x_1 = x_2 \in [M^\perp]^\perp \cap M^\perp = \{0\} \) so \( x = x_1 \in M^\perp \). This shows \([M^\perp]^\perp \subseteq M\), the inclusion reverse to the one we already knew. \( \) (Of course, (5) is just a restatement of the fact that the norm-closure of a convex set is the intersection of the closed half-spaces containing it.)

This geometry is easily turned into computation in view of the following considerations.

(* Definition: A(n indexed) set \( \{x_\alpha\}_{\alpha \in A} \) in a Hilbert space \( H \) is orthogonal if \( \alpha \neq \beta \Rightarrow \langle x_\alpha, x_\beta \rangle = 0 \). A(n indexed) set \( \{e_\alpha\}_{\alpha \in A} \) is orthonormal if \( \langle e_\alpha, e_\beta \rangle = \delta_{\alpha\beta} \) =

\[
\begin{cases}
1 & \text{if } \alpha = \beta \\
0 & \text{if } \alpha \neq \beta.
\end{cases}
\]

It is obvious that orthonormal sets—or, more generally, orthogonal sets not containing 0—are (finitistically) linearly independent: if one “dots” a linear dependence on \( x \) and this sum is orthogonal to the linear space spanned by \( \sum_{\alpha \in A} \lambda_\alpha e_\alpha \), one gets

\[
0 = \lambda_1 x_{\alpha_1} + \cdots + \lambda_n x_{\alpha_n}
0 = (0, x_\beta) = \lambda_1 \langle x_{\alpha_1}, x_\beta \rangle + \cdots + \langle \lambda_n x_{\alpha_n}, x_\beta \rangle
0 = \lambda_\beta \|x_\beta\|^2 \Rightarrow \lambda_\beta = 0.
\]

Moreover, the linear span of a finite orthogonal set is closed—and isomorphic as a Hilbert space to \( \mathbb{K}^n \) where \( n \) is its dimension. It suffices to consider orthonormal sets, for which

\[
(\lambda_1, \ldots, \lambda_n) \mapsto \lambda_1 x_{\alpha_1} + \cdots + \lambda_n x_{\alpha_n}
\]

has the property

\[
(\lambda_1, \ldots, \lambda_n) \cdot (\mu_1, \ldots, \mu_n) = \sum_{j,k=1}^n \lambda_j \mu_k \langle e_{\alpha_j}, e_{\alpha_k} \rangle = \left\langle \sum_{j=1}^n \lambda_j e_{\alpha_j}, \sum_{k=1}^n \mu_k e_{\alpha_k} \right\rangle.
\]

Since \( \mathbb{K}^n \) is (Cauchy-)complete, any such space is (norm-)closed.

Orthonormal sets abound in nature: \( e.g., \{e^{inx}\}_{n \in \mathbb{Z}} \) is an orthonormal set in \( L^2([-\pi, \pi]) \) with the usual inner product \( \langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) g(x) dx \). It is for this reason that one makes the

(*) Definition: If \( \{e_\alpha\}_{\alpha \in A} \) is an orthonormal set in \( H \), then for any \( x \in H \) the indexed family of numbers \( \{\langle x, e_\alpha \rangle\}_{\alpha \in A} \) are called the (generalized) Fourier coefficients of \( x \) (with respect to \( \{e_\alpha\}_{\alpha \in A} \)).

Evidently the Fourier coefficients of any vector are bounded, since the Schwarz inequality gives \( |\langle x, e_\alpha \rangle| \leq \|x\|\); but we shall soon be able to say much more. In any event

(*) Proposition: If \( \{e_\alpha\}_{\alpha \in A} \) is an orthonormal set in \( H \), then for any \( x \in H \) and finite subset \( F \subseteq A \), the vector \( \sum_{\alpha \in F} \langle x, e_\alpha \rangle e_\alpha \) is the point nearest to \( x \) in the (finite-dimensional) vector space spanned by \( \{e_\alpha\}_{\alpha \in F} \).

Easy Proof. For any \( \beta \in F \) one has \( \langle x - \sum_{\alpha \in F} \langle x, e_\alpha \rangle e_\alpha, e_\beta \rangle = \langle x, e_\beta \rangle - \langle x, e_\beta \rangle = 0 \). Thus the difference between \( x \) and this sum is orthogonal to the linear space spanned by \( \{e_\alpha\}_{\alpha \in F} \), and that property characterizes nearest points in vector subspaces.

Harder but perhaps more insightful Proof. Let \( \sum_{\alpha \in F} \lambda_\alpha e_\alpha \) be a linear combination of the \( \{e_\alpha\}_{\alpha \in F} \). Then if we write

\[
\|x - \sum_{\alpha \in F} \lambda_\alpha e_\alpha\|^2 = \|x - \sum_{\alpha \in F} \langle x, e_\alpha \rangle e_\alpha + \sum_{\beta \in F} ((\langle x, e_\beta \rangle - \lambda_\beta) e_\beta)\|^2
\]

(11) Do not be worried by the notion of an indexed set! The only indices you will see in numerical analysis will be the ordinary counting numbers \( \mathbb{N} \); our indexed sets will be either finite sequences or the usual indexed sequences. This will all seem very natural once we start using it.
(the change of the name of the index in the second sum is inconsequential) we see that each term of the second sum is orthogonal to each term of the first. It follows that the whole second sum is orthogonal to the first, and therefore this is the sum of two orthogonal terms. So the Pythagorean theorem is applicable, and we get

\[ \| x - \sum_{\alpha \in F} \lambda_{\alpha} e_{\alpha} \|^2 = \| x - \sum_{\alpha \in F} \langle x, e_{\alpha} \rangle e_{\alpha} \|^2 + \| \sum_{\beta \in F} (\langle x, e_{\beta} \rangle - \lambda_{\beta}) e_{\beta} \|^2. \]

The second term belongs to the finite-dimensional space spanned by \( \{ e_{\beta} \}_{\beta \in F} \), so by the isometry of that space with \( \mathbb{K}^n \) (or by direct computation), the norm-squared of that term is the sum of the squares of the components: we thus have

\[ \| x - \sum_{\alpha \in F} \lambda_{\alpha} e_{\alpha} \|^2 = \| x - \sum_{\alpha \in F} \langle x, e_{\alpha} \rangle e_{\alpha} \|^2 + \sum_{\beta \in F} |\langle x, e_{\beta} \rangle - \lambda_{\beta}|^2. \]

From the sum-of-squares form of the second term here, it is immediate that the minimum is attained if and only if \( \lambda_{\beta} = \langle x, e_{\beta} \rangle \) for each \( \beta \in F \).

A similar computation gives

**Lemma** [Bessel?]: If \( \{ e_{\alpha} \}_{\alpha \in A} \) is an orthonormal set in \( H \), then for any \( x \in H \) and finite subset \( F \subseteq A \),

\[ \| x - \sum_{\alpha \in F} \langle x, e_{\alpha} \rangle e_{\alpha} \|^2 = \| x \|^2 - \sum_{\alpha \in F} |\langle x, e_{\alpha} \rangle|^2. \]

**Proof.**

\[ \| x - \sum_{\alpha \in F} \langle x, e_{\alpha} \rangle e_{\alpha} \|^2 = \left( x - \sum_{\alpha \in F} \langle x, e_{\alpha} \rangle e_{\alpha} \right) \cdot \left( x - \sum_{\beta \in F} \langle x, e_{\beta} \rangle e_{\beta} \right) = \langle x, x \rangle - \sum_{\beta \in F} |\langle x, e_{\beta} \rangle|^2 - \sum_{\alpha \in F} \langle x, e_{\alpha} \rangle \cdot \sum_{\beta \in F} \langle x, e_{\beta} \rangle e_{\beta} \]

\[ = \langle x, x \rangle - \sum_{\beta \in F} |\langle x, e_{\beta} \rangle|^2 + \sum_{\alpha \in F} |\langle x, e_{\alpha} \rangle|^2 + \sum_{\alpha \in F} |\langle x, e_{\alpha} \rangle|^2 = \langle x, x \rangle - \sum_{\alpha \in F} |\langle x, e_{\alpha} \rangle|^2. \]

\[(*) \textbf{Corollary} [Bessel’s Inequality, First Version]: If \( \{ e_{\alpha} \}_{\alpha \in A} \) is an orthonormal set in \( H \), then for any \( x \in H \) and finite subset \( F \subseteq A \),

\[ \sum_{\alpha \in F} |\langle x, e_{\alpha} \rangle|^2 \leq \| x \|^2 \]

with equality if and only if \( x \) belongs to the linear space spanned by \( \{ e_{\alpha} \}_{\alpha \in F} \).

\[(*) \textbf{Corollary}: If \( \{ e_{\alpha} \}_{\alpha \in A} \) is an orthonormal set in \( H \), then(12) the filter \( \left\{ \sum_{\alpha \in F} |\langle x, e_{\alpha} \rangle|^2 \right\}_{F \subseteq A, F \text{ finite}} \) of finite sums of the absolute squares of the Fourier coefficients of \( x \)—an “unordered series of nonnegative terms”—converges (in \( \mathbb{R}^+ \)) to its supremum, which is \( \leq \| x \|^2 \).

For obvious reasons, one denotes the limit of such a filter of finite sums by \( \sum_{\alpha \in A} |\langle x, e_{\alpha} \rangle|^2 \). It is easy to see that such a convergent “unordered series” can have at most countably many nonzero terms; applying the usual argument is left to the reader.

---

(12) For numerical-analysis purposes, you can read these “filter of finite sums” constructions as if they were simply talking about an ordinary infinite series indexed by the counting numbers, with the series “unconditionally convergent”—that is, convergent to the same sum regardless of reordering of its terms.
then the linear space spanned by \( \{ e_\alpha \}_{\alpha \in A} \) is an orthonormal set in \( H \), then
\[
\sum_{\alpha \in A} |\langle x, e_\alpha \rangle|^2 \leq \|x\|^2
\]
with equality if and only if \( x \) belongs to the norm closure of the linear space spanned by \( \{ e_\alpha \}_{\alpha \in A} \).

\textit{Proof.} The statement makes sense because we now know the series converges, and the inequality follows from the first version of Bessel’s inequality. If equality holds, then given any \( \epsilon > 0 \) we can find finite \( F \subseteq A \) for which \( \| x - \sum_{\alpha \in F} \langle x, e_\alpha \rangle e_\alpha \|^2 < \epsilon^2 \); but that says \( \| x - \sum_{\alpha \in F} \langle x, e_\alpha \rangle e_\alpha \|^2 < \epsilon^2 \), so \( x \) can be approximated arbitrarily closely by finite linear combinations of the \( \{ e_\alpha \}_{\alpha \in A} \). On the other hand, if \( x \) can be approximated arbitrarily closely by finite linear combinations of the \( \{ e_\alpha \}_{\alpha \in A} \), then—the best approximation to \( x \) of the form \( \sum_{\alpha \in F} \lambda_\alpha e_\alpha \) is \( \sum_{\alpha \in F} \langle x, e_\alpha \rangle e_\alpha \)—given any \( \epsilon > 0 \) one can find \( F \subseteq A \) that makes \( \| x - \sum_{\alpha \in F} \langle x, e_\alpha \rangle e_\alpha \|^2 < \epsilon \), which means that one can make 0 \( \leq \langle x, x \rangle - \sum_{\alpha \in F} |\langle x, e_\alpha \rangle|^2 < \epsilon \), and that identifies the sum of the absolute squares of the Fourier coefficients.

This version of Bessel’s inequality shows that for any\(^{(13)} \) orthonormal set \( \{ e_\alpha \}_{\alpha \in A} \subseteq H \) the Fourier coefficients (relative to that set) of any \( x \in H \) belong to \( \ell^2(A) \). We already know that if \# \( A = n \in \mathbb{N}_0 \), then the linear space spanned by \( \{ e_\alpha \}_{\alpha \in A} \) is isometrically isomorphic—\( i.e. \), isomorphic with the inner product preserved—to \( \mathbb{K}^n = \ell^2(n) \). We can extend this to arbitrary \( A \), but we need to know a little about unordered series in normed spaces.

\( \textbf{(X)} \) An unordered series in \( (X, \| \cdot \|_X) \) is an indicated unordered sum \( \sum_{\alpha \in A} x_\alpha \), where the “terms” \( x_\alpha \) belong to \( X \). It converges if the (filter based on the) net \( \left\{ \sum_{\alpha \in F} x_\alpha \right\}_{F \subseteq A, \text{ } F \text{ finite}} \) converges. The limit (if it exists) is called the sum of the series, logically enough, and denoted by \( \sum_{\alpha \in A} x_\alpha \). It is easy to verify that convergent series can be multiplied term-by-term by scalars, and that convergent series indexed by the same set can be added term-by-term; the reader can easily supply the details of a verification, which requires nothing more than the continuity of the operations in \( X \). If \( "X" \) is a Hilbert space with the inner-product norm, then the continuity of the inner product also gives
\[
\left\langle \sum_{\alpha \in A} x_\alpha, \sum_{\beta \in B} y_\beta \right\rangle = \sum_{(\alpha, \beta) \in A \times B} \langle x_\alpha, y_\beta \rangle .
\]

In the most interesting cases the space \( (X, \| \cdot \|_X) \) is a Banach space, and the Cauchy condition for the convergence of unordered series takes an interesting and useful form:

\( \textbf{(X)} \) \textit{Proposition:} Let \( \sum_{\alpha \in A} x_\alpha \) be an unordered series in a Banach space. Then the series converges if and only if: for every \( \epsilon > 0 \) there exists a finite set \( F \subseteq A \) (depending, in general on \( \epsilon \)) such that if \( G \subseteq A \) is a finite set disjoint from \( F \), then \( \| \sum_{\alpha \in G} x_\alpha \| \leq \epsilon \).

\textit{Proof.} It is necessary and sufficient for convergence that the net of finite subsums of the series be a Cauchy net. If that net is Cauchy, then given \( \epsilon > 0 \) there exists finite \( F \subseteq A \) such that if \( F_1, F_2 \subseteq A \) are two finite subsets of \( A \) containing \( F \), then \( \| \sum_{\alpha \in F_1} x_\alpha - \sum_{\alpha \in F_2} x_\alpha \| \leq \epsilon \). If \( G \subseteq A \) is a finite subset disjoint from \( F \), then taking \( F_2 = G \cup F \) and \( F_2 = F \) in that inequality yields
\[
\left\| \sum_{\alpha \in G} x_\alpha \right\| = \left\| \sum_{\alpha \in G \cup F} x_\alpha - \sum_{\alpha \in F} x_\alpha \right\| \leq \epsilon
\]

\(^{(13)} \) Here again, the only index sets we have to work with are either finite: \( A = \{ 1, 2, \ldots, n \} \), or the counting-number indices \( A = \mathbb{N} \). The space \( \ell^2(n) \) is just our old friend \( \mathbb{R}^n \) or \( \mathbb{C}^n \), depending on what field of coefficients we’re using.
so the condition is necessary for convergence. On the other hand, if the condition holds then let \( \epsilon > 0 \) be given and take a finite set \( F \subseteq A \) with the property that if \( G \subseteq A \) is a finite set disjoint from \( F \), then \( \| \sum_{\alpha \in F} x_\alpha \| \leq \epsilon / 2 \). Given finite \( F_1, F_2 \subseteq A \) both of which contain \( F \), put \( G_1 = F_1 \setminus F \) and \( G_2 = F_2 \setminus F \). Because the terms with indices belonging to \( F \) cancel in the difference of the sums, we then have

\[
\left\| \sum_{\alpha \in F_1} x_\alpha - \sum_{\alpha \in F_2} x_\alpha \right\| = \left\| \sum_{\alpha \in G_1} x_\alpha - \sum_{\alpha \in G_2} x_\alpha \right\| \leq \left\| \sum_{\alpha \in G_1} x_\alpha \right\| + \left\| \sum_{\alpha \in G_2} x_\alpha \right\| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\]

and so the net of finite subsums is Cauchy.

**(X) Corollary:** A convergent unordered series in a normed space can have only countably many nonzero terms.

Indeed, for each \( n \in \mathbb{N} \) one can find finite \( F_n \subseteq A \) for which \( G \cap F_n = \emptyset \Rightarrow \| \sum_{\alpha \in G} x_\alpha \| \leq 1/n \), and thus in particular \( \alpha \in A \setminus F_n \Rightarrow \| x_\alpha \| \leq 1/n \). It follows that any term whose index does not belong to \( \bigcup_{n=1}^{\infty} F_n \) must be zero.

**(X)** The reader can easily verify that a series of scalars \( (\mathbb{R} = \mathbb{R} \) or \( \mathbb{C} ) \) is unordered convergent if and only if it is absolutely convergent; it follows that if \( \dim X < \infty \) then the unordered convergent series are exactly those for which \( \sum_{\alpha \in A} \| x_\alpha \| < \infty \). In \( (\mathbb{C}, [0,1], \| \cdot \|_\infty ) \), on the other hand, a series \( \sum_{\alpha \in A} f_\alpha \) is unordered convergent if and only if the series \( \sum_{\alpha \in A} | f_\alpha | \) is unordered convergent,\(^{14}\) but it is possible for that to happen even though \( \sum_{\alpha \in A} | f_\alpha | \|_\infty = \infty \). Conversely, however, it is easy to see that for a series \( \sum_{\alpha \in A} x_\alpha \) in any Banach space \( X \), the convergence of the series \( \sum_{\alpha \in A} \| x_\alpha \| \) of positive scalars—necessarily converging to the supremum of its finite subsums, so its convergence can be characterized as \( \sum_{\alpha \in A} \| x_\alpha \| < \infty \)—implies the condition given in the proposition above, and thus implies the convergence of the given vector series in \( X \).

In a Hilbert space, the condition for convergence of (unordered) series with orthogonal terms “scalarizes” in a very neat way:

**(\ast) Proposition:** Let \( \sum_{\alpha \in A} x_\alpha \) be an unordered series in a Hilbert space \( (H, \langle \cdot, \cdot \rangle) \), and suppose that its terms are pairwise orthogonal, i.e., that \( \alpha \neq \beta \Rightarrow \langle x_\alpha, x_\beta \rangle = 0 \). Then the series converges if and only if the (unordered nonnegative) scalar series \( \sum_{\alpha \in A} \| x_\alpha \|^2 < \infty \).

**Proof.** For any finite \( G \subseteq A \) we have the Pythagorean

\[
\left\| \sum_{\alpha \in G} x_\alpha \right\|^2 = \left\langle \sum_{\alpha \in G} x_\alpha, \sum_{\beta \in G} x_\beta \right\rangle = \sum_{\alpha \in G} \langle x_\alpha, x_\alpha \rangle = \sum_{\alpha \in G} \| x_\alpha \|^2.
\]

It follows that the scalar series in \( \mathbb{R}^+ \) satisfies the necessary and sufficient condition for Cauchy-ness established in the proposition above if and only if the vector series in \( H \) satisfies it.

Putting this together with the Bessel inequality, we have

**(\ast) Theorem:** If \( \{ e_\alpha \}_{\alpha \in A} \) is an orthonormal set in a Hilbert space \( H \), then for each \( x \in H \) the unordered series

\[ \sum_{\alpha \in A} \langle x, e_\alpha \rangle e_\alpha \]

converges (in the norm metric of \( H \)) to the point nearest to \( x \) in the norm closure of the linear space spanned by \( \{ e_\alpha \}_{\alpha \in A} \). In particular, the following three conditions are equivalent:

1. The vector series \( \sum_{\alpha \in A} \langle x, e_\alpha \rangle e_\alpha \) converges to \( x \);
2. \( x \) belongs to the norm closure of the linear space spanned by \( \{ e_\alpha \}_{\alpha \in A} \);
3. \( \sum_{\alpha \in A} | \langle x, e_\alpha \rangle |^2 = \| x \|^2 \).

\(^{14}\) Checking this assertion is elementary but not entirely trivial; it makes a nice exercise for the mathematically inclined reader.
Proof. The series obviously has pairwise orthogonal terms, so its convergence is guaranteed by the proposition just proved—joined with Bessel’s inequality, which gives a bound on the (scalar) series of squares of the lengths of the terms:

\[ \sum_{\alpha \in A} \| (x, e_\alpha) e_\alpha \|^2 = \sum_{\alpha \in A} | \langle x, e_\alpha \rangle |^2 \leq \| x \|^2 . \]

The sum of the series evidently belongs to the norm closure of the linear space spanned by \( \{ e_\alpha \}_{\alpha \in A} \), and since

\[ \left( \sum_{\alpha \in A} \langle x, e_\alpha \rangle e_\alpha, e_\beta \right) = \langle x, e_\beta \rangle - \langle x, e_\beta \rangle = 0 \]

holds for each \( \beta \in A \), the difference between \( x \) and the sum of the series is orthogonal to all the \( \{ e_\beta \}_{\beta \in A} \) and thus to the closure of the subspace that they span. It follows that the sum of the series is the projection of \( x \) on that subspace. As to the equivalence of the three conditions listed: (1) \( \Rightarrow \) (2) is clear. (2) \( \Rightarrow \) (3) is the sufficient condition for equality in Bessel’s inequality (second version). Finally, (3) and the relation

\[ \| x - \sum_{\alpha \in F} \langle x, e_\alpha \rangle e_\alpha \|^2 = \| x \|^2 - \sum_{\alpha \in F} | \langle x, e_\alpha \rangle |^2 \]

of the Lemma on p. 17 above show that when (3) holds the limit of the finite sums \( \sum_{\alpha \in F} \langle x, e_\alpha \rangle e_\alpha \) is precisely the vector \( x \in H \) with which one started.

The equivalence among (1), (2) and (3) above is sometimes called the Parseval relation, and (3) is called the Parseval identity or Parseval equation for the norm-closed subspace generated by the \( \{ e_\alpha \}_{\alpha \in A} \)—although that usage is frequently restricted to the case in which that subspace is all of the Hilbert space \( H \).

It is now appropriate to look at the relation between the norm-closed subspace generated by an orthonormal \( \{ e_\alpha \}_{\alpha \in A} \) and the orthonormal set itself—but from the viewpoint of the subspace rather than from that of the orthonormal set.

(*) Definition: Let \((H, \langle \cdot, \cdot \rangle)\) be a Hilbert space and \( M \subseteq H \) a norm-closed subspace. An orthonormal set \( \{ e_\alpha \}_{\alpha \in A} \subseteq M \) is called an orthonormal basis (or base) of \( M \) if \( M \) is the norm closure of the linear subspace generated by \( \{ e_\alpha \}_{\alpha \in A} \).

One skips the phrase “of \( M \)” in the case \( M = H \) and simply refers to an orthonormal basis.

(*) Proposition: Let \((H, \langle \cdot, \cdot \rangle)\) be a Hilbert space and \( M \subseteq H \) a norm-closed subspace. The following conditions on an orthonormal set \( \{ e_\alpha \}_{\alpha \in A} \subseteq M \) are logically equivalent:

1. \( \{ e_\alpha \}_{\alpha \in A} \) is an orthonormal basis of \( M \);
2. If \( y \in M \) is orthogonal to all the \( \{ e_\alpha \}_{\alpha \in A} \), then \( y = 0 \);
3. \( \{ e_\alpha \}_{\alpha \in A} \) is a maximal orthonormal subset of \( M \), i.e., one not properly contained in a larger orthonormal set;
4. For every \( x \in M \), the vector series \( \sum_{\alpha \in A} \langle x, e_\alpha \rangle e_\alpha \) converges to \( x \), i.e.,

\[ x = \sum_{\alpha \in A} \langle x, e_\alpha \rangle e_\alpha ; \]

5. For every \( x \in M \), \( \sum_{\alpha \in A} | \langle x, e_\alpha \rangle |^2 = \| x \|^2 \).

Proof. The equivalence of (1), (4) and (5) follows from the theorem just proved. (2) \( \Leftrightarrow \) (3) is clear: (2) prevents one from finding a new \( e_\beta \) that is orthogonal to all the \( e_\alpha \)’s one already has, while if (3) holds then any vector \( y \in M \) orthogonal to all the \( \{ e_\alpha \}_{\alpha \in A} \) must be zero, because otherwise one could set \( e = y/\| y \| \) and then \( \{ e_\alpha \}_{\alpha \in A} \cup \{ e \} \subseteq M \) would be orthonormal, contrary to maximality. (4) \( \Rightarrow \) (2): if all the coefficients in the series are zero, as they will be if \( x \perp \{ e_\alpha \}_{\alpha \in A} \), then \( x = 0 \). (2) \( \Rightarrow \) (4): Let \( x \in M \) be given; the
point nearest to \( x \) in the norm closure of the linear span of \( \{ e_\alpha \}_{\alpha \in A} \) is known to be \( \sum_{\alpha \in A} \langle x, e_\alpha \rangle e_\alpha \). Put \( y = x - \sum_{\alpha \in A} \langle x, e_\alpha \rangle e_\alpha \). For each \( \beta \in A \) we have \( \langle y, e_\beta \rangle = \langle x, e_\beta \rangle - \langle x, e_\beta \rangle = 0 \), so by (2) we have \( y = 0 \) and therefore \( x = \sum_{\alpha \in A} \langle x, e_\alpha \rangle e_\alpha \), which is (4).

(*) **Corollary:** Every norm-closed subspace \( M \neq \{0\} \) of a Hilbert space possesses orthonormal bases; indeed, every orthonormal set in such an \( M \) can be enlarged to an orthonormal basis. In particular, every orthonormal basis of a subspace \( M \subseteq H \) can be enlarged to an orthonormal basis of \( H \).

**Proof.** Let \( \{ e_\alpha \}_{\alpha \in B} \subseteq M \) be an orthonormal set. The family \( \mathcal{E} \) of orthonormal subsets of \( M \) that contain \( \{ e_\alpha \}_{\alpha \in B} \) is partially ordered under set-inclusion (in \( 2^M \)) and it is obvious that the union of a chain of orthonormal subsets of \( M \) is again an orthonormal subset of \( M \). Zorn’s lemma produces maximal orthonormal subsets \( \{ e_\alpha \}_{\alpha \in A} \subseteq M \), and the proposition just proved implies that these are orthonormal bases.

(X) It is surprisingly complicated to give a proof of the following proposition for uncountable cardinals, so we content ourselves with the countable cases. For the uncountable cases see, e.g., N. Dunford and J. T. Schwartz, Linear Operators, Vol. I, IV.4.14, pp. 253–255.

(X) **Lemma:** A Hilbert space \( H \) is separable (i.e., possesses a countable subset dense in its norm metric topology) if and only if every orthonormal basis of \( H \) is either finite or countably infinite,\(^{(16)} \) and for this it suffices that \( H \) possess one countable orthonormal basis.

**Proof.** Let \( \{ e_\alpha \}_{\alpha \in A} \subseteq M \) be an orthonormal basis of \( H \). We have already observed that all but countably many of the terms in a series \( x = \sum_{\alpha \in A} \langle x, e_\alpha \rangle e_\alpha \) must have zero coefficients. Thus if \( \{ x_n \}_{n=1}^\infty \) is dense in \( H \) and if for each \( n \in \mathbb{N} \) we put \( A_n = \{ \alpha \in A : \langle x_n, e_\alpha \rangle \neq 0 \} \), then \( \bigcup_{n=1}^\infty A_n \subseteq A \) is countable. If there were some \( \beta \in A \setminus \bigcup_{n=1}^\infty A_n \) it would have the property that \( \langle x_n, e_\beta \rangle = 0 \) for all \( x_n \) belonging to a dense subset of \( H \); approximating \( e_\beta \) arbitrarily closely by \( x_n \)'s, we see that \( \| e_\beta \|^2 = 0 \) which is impossible for an element of an orthonormal set. Thus \( A = \bigcup_{n=1}^\infty A_n \) and it is countable. Conversely, if \( \{ e_n \}_{n=1}^N \) or \( \{ e_n \}_{n=1}^\infty \subseteq H \) is a countable orthonormal basis of \( H \), then the set of all finite linear combinations \( \sum_{j=1}^k \lambda_n_j e_n_j \) (coefficients in \( \mathbb{K} \)) of its elements is dense in \( H \), and since each coefficient \( \lambda_n_j \) can be approximated arbitrarily closely in \( \mathbb{K} \) by elements of \( \mathbb{Q} \) (if \( \mathbb{K} = \mathbb{Q} \)) or \( \mathbb{Q}[i] \) (if \( \mathbb{K} = \mathbb{C} \)), the existence of the countable orthonormal basis forces \( H \) to be separable.

(X) **Proposition:** Let \( H \) be a separable Hilbert space. Then if \( M \subseteq H \) is a closed linear subspace the cardinality of all orthonormal bases of \( M \) is the same, and is either a finite cardinal or \( \aleph_0 \) (the cardinality of \( \mathbb{N} \)); hereinafter this cardinal is called the **dimension** of \( M \) and written \( \dim M \). If \( M \subseteq N \) are two closed linear subspaces, then \( \dim M \leq \dim N \). In general, all orthonormal bases of a given Hilbert space \( H \), separable or not, have the same cardinality.

**Proof.** If \( M \) is finite-dimensional (and thus automatically closed, as we observed on p. 44 above) then the invariance of the basis cardinality is a well-known theorem of linear algebra. If \( M \) is not finite-dimensional, then it is separable (as a subspace of a separable metric space) and the lemma shows that the cardinality of any orthonormal base must be \( \aleph_0 \). The fact that dimension grows with subspaces follows from the fact that if \( M \subseteq N \) are subspaces, then any orthonormal basis of \( M \) can be enlarged to an orthonormal basis of \( N \). For the general case of the invariance of cardinality of orthonormal bases, see Dunford & Schwartz, loc. cit.

It is natural to define an **isomorphism** of Hilbert spaces, also called an **isometric isomorphism**, an **orthogonal transformation** (if \( \mathbb{K} = \mathbb{R} \)) or a **unitary transformation** (if \( \mathbb{K} = \mathbb{C} \)) to be an algebraic

\(^{(15)} \) In these as in purely algebraic considerations about dimension and bases, one can avoid dealing with very-low-dimensional exceptions by agreeing to say that the zero subspace has the empty set as its basis, and that the empty set is (vacuously) orthonormal.

\(^{(16)} \) The original axiomatization of Hilbert spaces insisted that they be separable and thus finessed cardinality complications.
isomorphism $U : H_1 \to H_2$ of one Hilbert space onto another, such that $\langle Ux, Uy \rangle_2 = \langle x, y \rangle_1$ holds for all $x, y \in H_1$. The considerations for finite linear spans on p. 16 above can be viewed as having shown us that all Hilbert spaces of dimension $n < \aleph_0$ are isometrically isomorphic to $K^n$ with its usual inner product; we want to extend this to give us “standard models” of all Hilbert spaces.

(**) Proposition: Let $H$ be a Hilbert space$^{(17)}$ with orthonormal basis $\{e_\alpha\}_{\alpha \in A} \subseteq M$. Then the mapping

$$\ell^2(A) \to H \quad \text{(coefficients for $\ell^2$ from the scalar field of $H$)}$$

$$(\ldots, \lambda_\alpha, \ldots) \mapsto \sum_{\alpha \in A} \lambda_\alpha e_\alpha$$

is an isometric isomorphism (where $\ell^2(A)$ is given the inner product that goes with $L^2(A, 2^A, \#)$, namely $\langle (\ldots, \lambda_\alpha, \ldots), (\ldots, \mu_\alpha, \ldots) \rangle = \sum_\alpha \lambda_\alpha \overline{\mu_\alpha}$).

Proof. Not much is left to prove. If $(\ldots, \lambda_\alpha, \ldots) \in \ell^2(A)$, then the formal unordered series $\sum_{\alpha \in A} \lambda_\alpha e_\alpha$ has orthogonal terms, and $\sum_{\alpha \in A} |\lambda_\alpha|^2 < \infty$ (the definition of belonging to $\ell^2(A)$) is exactly the necessary and sufficient condition for this series to converge. Thus the mapping is well-defined, and it is trivial to verify—extending what one knows in the finite-sum case by continuity—that the mapping is linear and preserves the inner product. The mapping is onto in view of the theorem on p. 19 above.

Of course the first application one wants to consider in this setting is

(**) Proposition: For the Hilbert space $L^2([−\pi, \pi])$ equipped with the inner product

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) \overline{g(\theta)} \, d\theta,$$

the orthonormal set $\{e^{i n \theta}\}_{n \in \mathbb{Z}}$ is an orthonormal basis. Thus the unitary mappings

$$L^2 \to \ell^2 \quad \text{(Fourier transform)}$$

$$f \mapsto \hat{f} = \left( n \mapsto \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-i n \theta} \, d\theta \right)$$

$$\ell^2 \to L^2 \quad \text{(inverse Fourier transform)}$$

$$(\ldots, \lambda_k, \ldots) \mapsto \sum_{k \in \mathbb{Z}} \lambda_k e^{i k \theta}$$

(these mappings are each other’s inverses) are isomorphisms of Hilbert spaces between the space of functions $L^2([−\pi, \pi])$ and the “space of square-summable Fourier coefficients” $\ell^2(\mathbb{Z})$.$^{(18)}$

Proof. One needs only to show that the orthonormal set $\{e^{i n \theta}\}_{n \in \mathbb{Z}}$ is an orthonormal basis, since the correspondence between inner products is easily verified. It suffices to show that the linear span of $\{e^{i n \theta}\}_{n \in \mathbb{Z}}$ is dense in $L^2$. Because the “identity injection” of $(\mathbb{C}[−\pi, \pi], \| \cdot \|_\infty)$ into $(L^2, \| \cdot \|_2)$ is linear and norm-continuous, to prove density it suffices to observe two things: (1) The Stone-Weierstraß theorem implies that the linear span of $\{e^{i n \theta}\}_{n \in \mathbb{Z}}$, which consists exactly of the trigonometric polynomials, is uniformly dense in the space of continuous functions on $[−\pi, \pi]$ which take the same value at both endpoints, since these can be identified with the continuous functions on the unit circle $T$. (2) The continuous functions on $[−\pi, \pi]$ which take the same value at both endpoints are dense in $L^2$, because the characteristic functions of any measurable set $E \subseteq [−\pi, \pi]$ can be $L^2$-norm approximated by the characteristic function of a measurable set $F$ whose closure $\overline{F}$ is contained in $[−\pi, \pi]$, and any such function can be approximated in the $L^2$-norm metric by a continuous function $0 \leq f \leq 1$ whose support is also contained in $[−\pi, \pi]$. (The reader should

---

$^{(17)}$ In the numerical-analysis context, the only ones we shall consider are the finite-dimensional ones (for which the index set “$A$” is $\{1, 2, \ldots, n\}$ for some $n$), the spaces $L^2([a, b], w)$, and the space $\ell^2$ of “square-summable sequences of coefficients” indexed by the counting numbers. That was a lie, because in the case of the trigonometric = complex-exponential basis we used the signed integers $\mathbb{Z}$, but those are in 1-1 correspondence with the counting numbers.

$^{(18)}$ Of course, this is the example that started the whole thing.
check these assertions; they follow easily from the countable additivity of Lebesgue measure.) Since any such function can be uniformly approximated—and thus $L^2$-norm approximated—by a trigonometric polynomial, it follows that any simple function can be thus $L^2$-norm approximated, and that suffices to prove the density in $L^2$ of the trigonometric polynomials.

What you just read, of course, was a completely “soft” proof that the Fourier series of an $L^2$-function converges to the function in the $L^2$ norm. The only bit of “hard” analysis in the proof amounted to checking that the Taylor series of $\sqrt{1+x}$ has coefficients that are $o(1/n^a)$ for $1 < a < 3/2$ and that its series thus converges absolutely and uniformly for $|x| \leq 1$.

(*) This is the place at which we join up with Atkinson again, at his §4.4, p. 207 ff. There are ways to get orthonormal bases of separable Hilbert spaces that don’t involve infinite choice processes. One favorite is to take a sequence of vectors $\{x_n\}_{n=1}^\infty \subseteq H$ whose (finistic) linear span is dense in $H$ and orthogonalize them by the Gram-Schmidt orthogonalization process (for the details, see any good first-course-in-linear-algebra textbook, or see Kenneth M. Hoffman & Ray A. Kunze, Linear Algebra, 2nd ed., Prentice-Hall (1971), p. 280 and p. 287(19)) For the case with which we will be most concerned, see Atkinson’s Theorem 4.2, p. 209 ff. For spaces $L^2([a,b],w)$ on intervals of $\mathbb{R}$, a favorite choice of the sequence of vectors is the monomials $\{x^n\}_{n=0}^\infty \subseteq H$; this choice has certain peculiar properties that lead to closed-form formulas in a number of classical cases, and will be most important (along with the trigonometric or complex-exponential basis) for numerical work. The two examples that will most concern us are: for $L^2([-1,1],w)$ with $w(x) = \frac{1}{\sqrt{1-x^2}}$, Gram-Schmidt orthogonalization leads to the Chebyshev polynomials $T_n(x)$ which are characterized by the fact that $T_n(\cos \theta) = \cos n \theta$ for all $n \in \mathbb{N}$. (Easy verification: make the substitution $x = \cos \theta$ and use the orthogonality of the cosines on the interval $[0,\pi]$.) Taking the weight function $w(x) \equiv 1$, we shall find that Gram-Schmidt orthogonalization leads to the Legendre polynomials, beloved of classical potential theorists. For more details, see, e.g., H. S. Wilf, Mathematics for the Physical Sciences,(20) Dover (1978), Ch. 2, Orthogonal Polynomials, p. 48 ff.

(X) Measure-theoretic arguments involving the Radon-Nikodym theorem show that the dual of any $L^p$-space can be identified (for $1 < p < \infty$) with $L^q$, and so we know that any Hilbert space is its own dual (it looks like $l^2(A)$). It’s just as easy to give a geometric proof of the following theorem, avoiding measure theory altogether (except for having shown that $L^2$-spaces are complete in their norm metrics—and that required the full power of the Lebesgue measure theory).

(X) Theorem: Let $(H,\langle \cdot,\cdot \rangle)$ be a Hilbert space. The real-linear (but conjugate-linear if $\mathbb{K} = \mathbb{C}$) mapping

$$H \to H^*, \quad y \mapsto (x \mapsto \langle x,y \rangle)$$

is an isometry of $H$ onto $H^*$.

Proof. It is evident that the mapping is real-linear, complex-conjugate-linear, and sends $H$ into $H^*$; indeed, the Schwarz inequality shows that the $H^*$-norm of $(x \mapsto \langle x,y \rangle)$ is $\leq ||y||$, and since equality is attained in the Schwarz inequality for $x = y$, the mapping is an isometry (and therefore 1-1). To see that it is onto, let $h^* \in H^*$ be given. If $h^* = 0$ there is nothing to prove, so assume $h^* \neq 0$; then its null space $M = \{x : h^*(x) = 0\}$ is a proper closed (because $h^*$ is continuous) subspace of $H$ of codimension 1. We can write $H = M \oplus M^\perp$ where $M^\perp \neq \{0\}$; indeed, by elementary linear algebra, $\dim M^\perp = 1$. Let $e \in M^\perp$ be a unit vector (evidently $\{e\}$ is an orthogonal basis of $M^\perp$) and let $y = \overline{h^*(e)} e$ (the complex conjugation is vacuous if $\mathbb{K} = \mathbb{R}$). Then for any $x \in H$ orthogonal decomposition gives us

$$x = (x - \langle x,e \rangle e) + \langle x,e \rangle e \quad \text{(evidently } x - \langle x,e \rangle e \in [M^\perp] = M)$$

$$h^*(x) = h^*(x - \langle x,e \rangle e) + \langle x,e \rangle h^*(e)$$

$$= 0 + \langle x,h^*(e) e \rangle e = \langle x,y \rangle$$

(19) In making the alternative between this and a good (etc.), I am not necessarily implying that this is not a good book, just that it's not a first-course-(etc.) book. *Honi soit qui mal y pense.*

(20) A modest and disarming title for a book that might well have been called “an exceedingly lovely introduction to some basic classical mathematics with which every analyst who doesn’t want to look like a complete idiot should be acquainted.”
so \( h^* = (x \mapsto \langle x, y \rangle) \).

**X** Corollary [Hahn-Banach Theorem for Hilbert spaces]: Let \( M \) be a subspace of a Hilbert space \( H \) and let \( \Psi \in M^* \). Then \( \Psi \) can be extended to an element of \( H^* \) without change of norm.

Proof. Since \( \Psi \in M^* \) is continuous, it can be extended to \( M \) in a unique way by simply taking limits.\(^{(21)}\) Thus there is no loss of generality in assuming that \( M \) is closed; so \( M \) is a Hilbert space in the relativized inner product of \( H \), and thus—by the theorem—there is a (unique) element \( y \in M \) for which \( \Psi(x) = \langle x, y \rangle \) for \( x \in M \). But the r. h. s. of that identity defines a linear functional \( x \mapsto \langle x, y \rangle \) which obviously extends \( \Psi \), and since the norms of \( \Psi \) and \( y \) were the same and the norm of \( x \mapsto \langle x, y \rangle \) on \( H \) is the same as \( ||y|| \), we have extended \( \Psi \) from \( M \) to \( H \) without increase of norm.

{Remark: In the Hilbert-space situation, the norm-preserving extension of \( \Psi \) to an element of \( H^* \) is unique: if \( x \mapsto \langle x, z \rangle \) is an(other) extension of \( \Psi \), then \( z - y \in M^\perp \) and so the resolution of \( z \) into orthogonal components along \( M \) and \( M^\perp \) is

\[
\begin{align*}
z &= y + (z - y) \\
\|z\|^2 &= \|y\|^2 + \|z - y\|^2
\end{align*}
\]

where the norm relation follows from the Pythagorean theorem. But clearly if \( z \neq y \) then \( \|z\| > \|y\| = \|\Psi\| \), so any choice of extension other than \( x \mapsto \langle x, y \rangle \) will increase the norm of the extension. Uniqueness of norm-preserving extension does not hold for Banach spaces in general: any element of \( \ell^\infty(2) \) of the form \((1, \mu) \) with \(|\mu| \leq 1 \) will have \( \|(1, \mu)\|_\infty = 1 \) and will extend the linear functional \( \Psi \) defined on the subspace of \( \ell^1(2) \) spanned by the first standard basis vector \((1, 0)^T \) by \( \Psi : (\lambda, 0) \mapsto \lambda \). The fact that the \( \ell^\infty \) unit ball has “flats” is responsible for this situation. Similar and worse things happen in the spaces \( L^1([a, b]) \).}


We now have the equipment to look at least-squares approximation problems from a geometric viewpoint, and so we can do a leisurely read through Atkinson’s §4.4, absorbing the things we need about orthogonal polynomials. This § of these notes will thus consist of commentary on Atkinson’s §4.4, and so it will not have the “n.K” paragaphic scheme of its predecessors.

Most of A.’s pp. 208–209 is devoted to defining the inner product \( \langle f, g \rangle_w = \int_a^b f(x) g(x) w(x) \, dx \). While he only considers \( f, g \in \mathcal{C}([a, b]) \) in (4.4.1), the formula makes sense for Lebesgue-measurable functions with \( \int_a^b |f(x)|^2 w(x) \, dx < \infty \), and those form a Hilbert space. A. uses only real scalars; the formulas with the complex-conjugation bar make sense for reals and give back (4.4.1), so the generalization may be handy but isn’t an essential one. We already have the Cauchy-Schwarz inequality (4.4.3) back on p. 11 above, and we know it’s also valid in the complex-scalar case.

These notes will use \( \langle \cdot, \cdot \rangle_w \) instead of A.’s \( \langle \cdot, \cdot \rangle \). Many writers now prefer the angle brackets, the subscript \( w \) (which may be omitted when it’s clear from the context) reminds the reader that a weight function is involved, and the ordered-pair notation (\( \langle \cdot, \cdot \rangle \)) is overworked anyway.

A.’s Theorem 4.2 is the Gram-Schmidt process mentioned (but not discussed) on p. 23 above. Everybody does this in the same way; there is nothing much to add other than the fact that sometimes one doesn’t “normalize” the \( \varphi_n(x) \)’s to have \( \| \cdot \|_w \)-norm 1, but rather omits the step (4.4.9) and is willing to tolerate \( \{\varphi_n\}_{n=0}^\infty \) obeying the requirement \( \langle \varphi_m, \varphi_n \rangle_w = 0 \) for \( m \neq n \), but not necessarily obeying \( \|\varphi_n\|_w = 1 \). The “standard” normalization choices for the classical (Chebyshev, Legendre, Hermite, etc.) orthogonal polynomials do not normalize to \( \|\cdot\|_2 = 1 \), and while there are some generally-agreed-upon choices (the Chebyshev polynomials are firm, the Legendre polynomials less so), one always has to check that the two books or papers one happens to be reading are using the same normalizations.

\(^{(21)}\) The argument involved here is the one showing that a uniformly continuous function from a dense subspace of a metric space \((X,d)\) to a complete metric space \((Y,\rho)\) has a unique continuous extension to all of \( X \), obtained by taking limits along convergent sequences.
The **Particular Cases** 1, 2, 3 of A.’s pp. 210–212 are extremely important examples; so is the example of the Hermite polynomials, which he does not mention. For the sake of uniformity, I’ll hand out charts listing, for each of these orthogonal families, one of the standard choices for weight function, normalization (that is, the values of $\|\varphi_n\|_w^2$), recurrence formulas (A.’s **Theorem 4.5**), Rodrigues’ Formula (all of these families—not merely the Legendre polynomials—can be produced by formulas analogous to (4.4.10)), Gaussian quadrature formulas (important in A.’s §5.3), differential equations (important in potential theory) and differential recurrence (important in some differential-equations settings).

A.’s **Theorem 4.4** turns out to be extremely valuable in approximate integration, and his proof is the industry-standard one. Another thing that one can extract from the argument he uses, along with the three-term recursion relations (4.4.21), is that the $(n - 1)$ (real) roots of $\varphi_{n-1}(x) = 0$ always separate the (real) roots of $\varphi_n(x) = 0$; that is, for $n \geq 2$ each of the roots of “equation $n - 1$” lies in one of the (finite) open intervals whose endpoints are roots of “equation $n$”. See his problem 21, p. 243.

On the other hand, §4.5 is just a recap of the geometric **Proposition** at the bottom of p. 16 above. The orthonormal set \{\varphi_j\}_{j=0}^n is an orthonormal basis of $P_n$ (considered as a subspace of the Hilbert space $L^2([a, b], w)$, so the geometry tells us that $r^*_n = \sum_{j=0}^n (f, \varphi_j)_w \varphi_j$ is the nearest point of $P_n$, and that the “remaining distance to $f$ from $P_n$” is exactly $\|f - r^*_n\|_w$. **Theorem 4.7** is similarly clear from the standpoint of Lebesgue integration, because the continuous functions (of compact support, if we’re dealing with an infinite interval $[a, b]$) are always a dense subspace of the Hilbert space $L^2([a, b], w)$. However, Atkinson passes by a consequence of his proof that has considerable conceptual interest, and we need to investigate it before proceeding.

This consequence is the **Erdős-Turán Theorem**. To prove it we need to take a little bit of Gaussian integration theory for granted, but this is a borrowing that we shall pay back later in the course. We borrow A.’s **Theorem 5.3**, p. 272, with a small change in wording:

**Theorem**: For each $n \geq 1$, there is a unique numerical integration formula of the form (5.3.1) which is exact for polynomials of degree $\leq 2n - 1$. Assuming $f(x)$ is $2n$-times differentiable, the numerical integration formula and its error term are given by

$$
\int_a^b f(x) \, w(x) \, dx = \sum_{j=1}^n w_{n,j} f(x_{n,j}) + \frac{\gamma_n}{A_n^2(2n)!} f^{(2n)}(\eta)
$$

for some “sample point” $a < \eta < b$. The evaluation points $\{x_{n,j}\}_{j=1}^n$ are the zeros of $\varphi_n(x)$, and the weights $w_{n,j}$ are given by

$$
w_{n,j} = -\frac{a_n \gamma_n}{\varphi'_n(x_{n,j}) \varphi_{n+1}(x_{n,j})} \quad \text{for } j = 1, \ldots, n \, .
$$

These weights are positive and their sum $\sum_{j=1}^n w_{n,j} = \int_a^b 1 \cdot w(x) \, dx$ (two facts which we need in the proof of Erdős-Turán).

**[Erdős-Turán] Theorem**: Let $[a, b] \subseteq \mathbb{R}$ be a finite interval on the real line and $w(\cdot) \geq 0$ a nonnegative, integrable ($\int_a^b w(x) \, dx < \infty$) weight function on it. Let $\{\varphi_n\}_{n=0}^\infty$ be the orthogonal polynomials on $[a, b]$ with respect to the weight $w$, and let $\{x_{n,j}\}_{j=1}^n$ be the roots of $\varphi_n(x) = 0$. Let $f \in C((a, b))$ be given, and for each $n \in \mathbb{N}$ let $P_n(x) \in P_n$ be the polynomial of degree $\leq n$ that interpolates $f$ at the points $\{x_{n+1,j}\}_{j=1}^{n+1}$. Then $\|f - P_n\|_w \to 0$ as $n \to \infty$.

**Proof**. Let $W = \int_a^b w(x) \, dx$ denote the “total weight” of $w$ on $[a, b]$. Let $\epsilon > 0$ be given, and let $p(x) \in P_N$ be a polynomial of some degree $N$ such that $|f(x) - p(x)| \leq \epsilon$ holds uniformly for $x \in [a, b]$; such a $p(x)$ exists by the Weierstraß (or Bernshtein) approximation theorem. Then

(22) Recall from Atkinson, Theorem 4.4, p. 213 that these roots all lie in $[a,b]$. 

25
\[ \|f - p\|_w^2 = \int_a^b |f(x) - p(x)|^2 w(x) \, dx \leq \epsilon^2 \cdot W \]

or, equivalently, \( \|f - p\|_w \leq \epsilon \sqrt{W} \).

If \( n \geq N \), then \( \deg(p(x) - P_n(x)) \leq n \) and so \( p(x) - P_n(x) \) is its own unique interpolant of degree \( \leq n \) at the points \( \{x_{n+1,j}\}_{j=1}^{n+1} \). Its square therefore has degree at most \( 2n \), and so the Gaussian integration formula (5.3.1)—with \( n \) replaced by \( n + 1 \)—is exact for \( |p(x) - P_n(x)|^2 \): we have

\[
\int_a^b |p(x) - P_n(x)|^2 w(x) \, dx = \sum_{j=1}^{n+1} [p(x_{n+1,j}) - P_n(x_{n+1,j})]^2 \cdot w_{n+1,j}.
\]

But \( P_n(x_{n+1,j}) = f(x_{n+1,j}) \) because \( P_n(x) \) interpolates \( f(x) \) at these points, and so also each difference \( |p(x_{n+1,j}) - P_n(x_{n+1,j})| \leq \epsilon \). Since the \( w_{n+1,j} \)'s are all positive, we obtain

\[
\int_a^b |p(x) - P_n(x)|^2 w(x) \, dx = \sum_{j=1}^{n+1} [p(x_{n+1,j}) - P_n(x_{n+1,j})]^2 \cdot w_{n+1,j} \leq \epsilon^2 \sum_{j=1}^{n+1} w_{n+1,j} = \epsilon^2 \cdot W.
\]

Putting these together, we find that for \( n \geq N \)

\[ \|f - P_n\|_w \leq \|f - p\|_w + \|p - P_n\|_w \leq 2\epsilon \cdot \sqrt{W} \]

which shows that \( \|f - P_n\|_w \) can be made arbitrarily small by taking \( n \) sufficiently large, as desired.

What this theorem does (if anything) is to convince one that interpolation as a means of approximation is fundamentally an honest technique: even though the interpolants of a continuous function on a finite interval may not converge to it uniformly on that interval, it is possible—by picking the interpolation points correctly, in accordance with the weight—to guarantee weighted-root-mean-square convergence of the interpolants to the function for any weight function assigning finite total weight to the interval. To demonstrate the validity of some applications, this kind of convergence may be good enough—e.g., if the interpolating polynomial is to be integrated, or more generally used as an approximate r. h. s. in a differential equation. The problem with the theorem, of course, is that to determine "\( N \)" effectively for a given \( \epsilon \) requires determination of the degree (at least) of a polynomial \( p(x) \) that approximates \( f(x) \) uniformly within a constant multiple of \( \epsilon \)—and one could use that polynomial instead of the interpolant. We shall be looking at trigonometric approximation later, and will prove the corresponding theorem (which is easier, and will probably seem more natural): if \( f \) is a continuous function of period \( 2\pi \), or (equivalently) a function on the unit circle, then its trigonometric interpolants at equally spaced points \( 0, 2\pi/n, 4\pi/n, \ldots \) (if \( n = 2k + 1 \) is odd, the interpolant should use only \( 1, \cos t, \ldots, \cos kt; \sin t, \ldots, \sin kt \); the condition for even \( n \) is slightly stranger-looking without some previous preparation) converge in root-mean-square to \( f \) as \( n \to \infty \). Again, this theorem justifies interpolation as a means of approximation, even though the rate of approach as \( n \to \infty \) is not very well estimated.

On the other hand, since the best one could find out about the root-mean-square error of \( \text{any} \) function \( g(x) \) from knowing that \( |f(x) - g(x)| \leq \epsilon \) for all \( x \in [a, b] \) would be to integrate the square of that estimate and get

\[
\|f - g\|_w^2 = \int_a^b |f(x) - g(x)|^2 w(x) \, dx \leq \int_a^b \epsilon^2 w(x) \, dx = \epsilon^2 W^2
\]

\[ \|f - g\|_w \leq W \cdot \epsilon, \]

a more optimistic way to interpret this result is: the \( \text{best} \) uniform approximation you could make to \( f(x) \) with a polynomial of degree \( \leq n \) would not necessarily give you a much better root-mean-square error than you could get by simply interpolating \( f(x) \) at the zeros of \( \varphi_{n+1}(x) \): you could decrease the error estimate, which is sharp (it cannot be improved for general \( f(x) \)) only by a factor of \( 1/2 \). But you can compute the interpolating polynomial, while the uniform approximating polynomial may be hard to find. Similarly,
the best approximation in the sense of weighted root-mean-square error would be $\sum_{k=0}^{n} \frac{\langle f, \varphi_k \rangle_w}{\| \varphi_k \|_w^2} \varphi_k(x)$; but computing it would involve $n + 1$ integrations rather than $n + 1$ evaluations (assuming that the norms of the $\varphi_k$’s are known).

A.’s Theorem 4.5—the three-term recursion relations that polynomials orthogonal with respect to a weight function always satisfy—looks more complicated than it is, because he has to be careful to set up notation for the constants in the recursions. The basic idea is the following rather simple one. Suppose you already know the polynomials $\{\varphi_0(x), \ldots, \varphi_n(x)\}$. Then $x \cdot \varphi_n(x)$ is a polynomial of degree $n + 1$ and so is $\varphi_{n+1}$, so you can find some constant $a$ for which the terms of degree $n + 1$ in the difference $\varphi_{n+1}(x) - ax\varphi_n(x)$ cancel—which means that this is a polynomial of degree at most $n$, so you can write

\[
\varphi_{n+1}(x) - ax \cdot \varphi_n(x) = b\varphi_n(x) - c_{n-1}\varphi_{n-1}(x) - c_{n-2}\varphi_{n-2}(x) - \cdots - c_0\varphi_0(x)
\]

\[
c_{n-2}\varphi_{n-2}(x) + \cdots + c_0\varphi_0(x) = -\varphi_{n+1}(x) + ax \cdot \varphi_n(x) + b\varphi_n(x) - c_{n-1}\varphi_{n-1}(x) .
\]

If you “dot” both sides of (*) with any $\varphi_j(x)$ whose degree $j < n - 1$, you get

\[
c_j \langle \varphi_j, \varphi_j \rangle_w = -\langle \varphi_{n+1}, \varphi_j \rangle_w + a\langle x \cdot \varphi_n, \varphi_j \rangle_w + b\langle \varphi_n, \varphi_j \rangle_w - c_{n-1}\langle \varphi_{n-1}, \varphi_j \rangle_w
\]

\[
c_j \| \varphi_j \|_w^2 = -\langle \varphi_{n+1}, \varphi_j \rangle_w + a\langle x \cdot \varphi_n, \varphi_j \rangle_w + b\langle \varphi_n, \varphi_j \rangle_w - c_{n-1}\langle \varphi_{n-1}, \varphi_j \rangle_w
\]

\[
= 0 + a \cdot 0 + b \cdot 0 - c_{n-1} \cdot 0 = 0
\]

where the zeros on the r. h. s. of the last set-off line occur because the polynomial on the r. h. s. of the $\langle \cdot, \cdot \rangle_w$ is of lower degree than the $\varphi_{n+1}$, $\varphi_n$ or $\varphi_{n-1}$ on the l. h. s.—and any $\varphi_k$ is orthogonal to all polynomials of degree $< k$. The only non-self-explanatory thing is how the “$x$” migrated from one side of the dot product to the other. (23) and the reason is the definition:

\[
\langle x \cdot f, g \rangle_w = \int_a^b [x \cdot f(x)]g(x) w(x) dx = \int_a^b f(x)[x \cdot g(x)] w(x) dx = \langle f, x \cdot g \rangle_w .
\]

So all the $c_j$’s are zero for $j \leq n - 2$, which means that the l. h. side of (*) is actually the zero function! But then

\[
\varphi_{n+1}(x) = (ax + b) \cdot \varphi_n(x) - c_{n-1}\varphi_{n-1}(x) ,
\]

as advertised. Of course the values of $a$, $b$, and $c$ will vary for various values of $n$; the formulas giving them in terms of the other constants $A_n$, $B_n$ and $\gamma_n$ are Atkinson’s (4.4.22), p. 214.

In practice, working out the values of these constants isn’t an enormously enlightening exercise, and of course their values depend on the normalizations one has chosen in the course of constructing the $\{\varphi_n(x)\}_{n=0}^\infty$. Most people look them up. There is a “general theory of good normalizations” for the so-called “classical orthogonal polynomials.” The reference in H. S. Wilf’s book of note (20) above is a good place to start; the standard treatise, old as it is, is G. Szegö, *Orthogonal Polynomials*, Amer. Math. Soc. Colloquium Publications 23, 3rd ed. 1967. The literature on this subject begins early, and the amount of it is formidable.

The best-known and probably most useful orthogonal polynomials are the Chebyshev polynomials. While we can’t get the most mileage from them until we look at trigonometric approximation, let us take a small break from Atkinson’s treatment and look at some of their most important properties.

**A Very Brief Introduction to the Chebyshev Polynomials.**

### 1. Definition and Basic Properties.

The Chebyshev polynomials (“of the first kind”: we shall say very briefly what the polynomials of the second kind are, below) are a sequence $\{T_n(x)\}_{n=0}^\infty$ of polynomials,

(23) What is happening here, to the functional analyst’s eye, is that the multiplication-by-$x$ operator is self-adjoint; it behaves toward the inner product of functions in the same way that a (conjugate-) symmetric matrix does toward the usual inner product in Euclidean space (or its complex counterpart).
indeed the trig identity \( \cos 2t \) behaves for \(-1 \leq x \leq 1\) the way \( \cos nt \) does for \(0 \leq t \leq \pi\)."

One way to define them—although it is not immediately clear that the statement is a definition—is to say that \( T_n(x) \) is the (unique) polynomial of degree \( n \) for which the identity \( T_n(\cos t) \equiv \cos nt \) holds. To see that this is meaningful we have to see why such polynomials should exist. The basic reason is that the addition formula for the cosine function can be used inductively. For any arguments \( A \) and \( B \) it is true that

\[
\cos(A \pm B) \equiv \cos A \cos B \mp \sin A \sin B .
\]

(1.1)

If one adds the two versions of this with the ambiguous signs, then the sine terms cancel and one is left with

\[
\cos(A + B) + \cos(A - B) \equiv 2 \cos A \cos B .
\]

(1.2)

Now for any real \( t \) and any natural number \( n \) we may take \( A = nt \) and \( B = t \) to get

\[
\cos(n + 1)t + \cos(n - 1)t \equiv 2 \cos nt \cos t
\]

\[
\cos(n + 1)t \equiv 2 \cos t \cos nt - \cos(n - 1)t .
\]

(1.3)

It follows that if we define the Chebyshev polynomials to be the polynomials determined by the recursion

\[
T_0(x) \equiv 1, \quad T_1(x) \equiv x, \quad T_{n+1} = 2x \cdot T_n(x) - T_{n-1}(x) \quad \text{for } n \geq 2
\]

(1.4)

then these will have the property that \( T_n(\cos t) \equiv \cos nt \). For example, \( T_2(x) = 2x \cdot x - 1 = 2x^2 - 1 \), and indeed the trig identity \( \cos 2t = 2 \cos^2 t - 1 \) is well known (it is the one that makes it possible to integrate \( \int \cos^2 t \, dt \)); \( T_3(x) = 2x \cdot (2x^2 - 1) - x = 4x^3 - 3x \) and thus \( \cos 3t = 4 \cos^3 t - 3 \cos t \) is less well-known, but also occasionally useful.\(^{(24)}\)

Because the \( k \)-th Chebyshev polynomial is of degree exactly \( k \), the Chebyshev polynomials \( \{T_k(x)\}_{k=0}^n \) form a basis of \( \mathcal{P}_n \). Moreover, because the Chebyshev polynomials are defined by a two-term recursion, namely (1.4), the values of a polynomial written out in its basis expansion in terms of the Chebyshev polynomials can be computed by a recursive algorithm only slightly more complicated than the nested-multiplication process for evaluating a polynomial written out in terms of a power basis or a Newton basis. One incarnation of the recursion is A.'s Chebeval on pp. 221–222. Another incarnation of the algorithm—where \( p_n(x) \in \mathcal{P}_n \) has the basis expansion

\[
p_n(x) = a(0) + a(1)T_1(x) + a(2)T_2(x) + \cdots + a(n)T_n(x)
\]

(1.5)

—is realized in the C routine below (the coefficients \( \{a(k)\}_{k=0}^n \) have been written as an array, rather than with subscripts)\(^{(25)}\)

The Chebeval Algorithm in C

```c
float chebeval(a,n,x)
float a[],x;
int n;
{
float b=0.0, bb=0.0, s, y, y2;
int j;
y2 = 2.0*x;
for(j=n;j>=1;j--)
{
s = b;
b = y2*b - bb + a[j];
bb = s;
}
return y*b - bb + a[0];
}
```

\(^{(24)}\) For example, the proof that it is impossible to trisect the 30° angle by a ruler-and-compass construction (so the angle-trisection problem of classical Euclidean geometry has no solution) uses that identity.

\(^{(25)}\) For a full discussion of this algorithm, see Numerical Recipes in C, Cambridge (1988), pp. 152–154 and pp. 158–162, or one of the other Numerical Recipes versions in other high-level languages. (Java programmers should be able to follow the trend of the C code.) Another discussion, with slightly different notation, can be found on p. 449 of Cheney & Kincaid. The reader might have fun writing out exactly what the algorithm embodied in this code does for a few moderately small values of \( n \).
and indeed one can establish the trig identity
\[
\sin(A \pm B) \equiv \sin A \cos B \pm \cos A \sin B
\] (1.6)
and adds the two versions of this with the ambiguous signs, then the second terms cancel and one is left with
\[
\sin(A + B) + \sin(A - B) \equiv 2 \sin A \cos B.
\] (1.7)
For any real \( t \) and any natural number \( n \) we may take \( A = nt \) and \( B = t \) to get
\[
\sin(n + 1)t + \sin(n - 1)t \equiv 2 \sin nt \cos t
\]
\[
\sin(n + 1)t \equiv 2 \sin nt \cos nt - \sin(n - 1)t.
\] (1.8)

Define the Chebyshev polynomials of the second kind to be the polynomials determined by the recursion
\[
U_0(x) \equiv 1, \quad U_1(x) \equiv 2x, \quad U_{n+1} = 2xU_n(x) - U_{n-1}(x) \quad \text{for } n \geq 1;
\] (1.9)
then these have the property that \( U_n(\cos t) \sin t \equiv \sin(n + 1)t \). For example, \( U_2(x) = 2x \cdot 2x - 1 = 4x^2 - 1 \), and indeed one can establish the trig identity
\[
\sin 3x = \sin 2x \cos x + \cos 2x \sin x = 2 \sin x \cos^2 x + (2 \cos^2 x - 1) \sin x = U_2(\cos x) \sin x
\]
directly. The relations between the Chebyshev polynomials of the second kind and the sines imply that these polynomials satisfy the integral orthogonality relations
\[
\int_{-1}^{1} U_j(x) U_k(x) \sqrt{1 - x^2} \, dx = \left\{ \begin{array}{ll} \frac{\pi}{2} & \text{if } j = k > 0; \\ 0 & \text{otherwise}. \end{array} \right.
\] (1.10)
These follow simply by taking the trigonometric orthogonality relations
\[
\int_{0}^{\pi} \sin j t \sin k t \, dt = \left\{ \begin{array}{ll} \frac{\pi}{2} & \text{if } j = k > 0; \\ 0 & \text{otherwise}. \end{array} \right.
\] (1.11)
and making the change of variable \( x = \cos t \). From this fact it follows that these are in fact the orthogonal polynomials produced by the weight function \( w(x) = \sqrt{1 - x^2} \), so of course everything we know about orthogonal polynomials applies to them. They also satisfy certain discrete orthogonality relations that they inherit from the corresponding relations for the sine functions. The Chebyshev polynomials of the second kind play a rôle in approximation in the \( L^1[-1,1] \) norm that is parallel to the one that the polynomials of the first kind play in approximation in the \( \| \cdot \|_\infty \) norm.

2. Chebyshev Nodes and Interpolation. The reason that “each \( T_n(x) \) is a polynomial that behaves like a cosine on \([-1,1]\)” is that the map \( t \mapsto x = \cos t \) sends \([0, \pi]\) onto \([-1,1]\): \( x \) is a number between \(-1\) and \(1\) if and only if \( x = \cos t \) for some uniquely determined \( t \) with \( 0 \leq t \leq \pi \). Thus one can frequently think trigonometrically rather than algebraically when one is working with the Chebyshev polynomials. A particularly useful consequence of this fact is that because \( \cos(n+1)t \) has \( n + 1 \) zeros in the interval \( 0 \leq t \leq \pi \), the corresponding Chebyshev polynomial \( T_{n+1}(x) \) has all its \((n + 1)\) distinct zeros in \(-1 \leq x \leq 1\); similarly, it satisfies \( \max_{-1 \leq x \leq 1} |T_{n+1}(x)| = 1 \). The zeros of \( T_{n+1}(x) \), called the Chebyshev zeros or Chebyshev points (sometimes Chebyshev nodes), can be concretely represented (for each \( n \)) as \( x_j = \cos t_j \) where \( t_j = \frac{\pi}{2(n + 1)}, \frac{3\pi}{2(n + 1)}, \ldots, \frac{(2j+1)\pi}{2(n + 1)}, \ldots, \) and thus all the roots of a Chebyshev polynomial can be determined without using root-finding or approximation (other than that inherent in the built-in trigonometric functions of a high-level language). It follows easily from the recursion (1.4) above that the leading coefficient of \( T_{n+1}(x) \) is \( 2^n \):
where $T_n(x) = 2^n x^n + (\text{lower order terms})$ \hfill (2.1)
and thus that if \{\(x_0, x_1, \ldots, x_n\)\} is a list of the roots of $T_{n+1}(x)$ in increasing order, then (by factoring them all out of $T_{n+1}(x)$)
\[
T_{n+1}(x) = 2^n (x-x_0)(x-x_1) \cdots (x-x_n) \quad \text{or} \quad \Psi(x) = \frac{1}{2^n} T_{n+1}(x) \hfill (2.2)
\]
\[
\max_{-1 \leq x \leq 1} |\Psi(x)| = \frac{1}{2^n} \quad \text{for these nodes.} \quad \hfill (2.4)
\]
This can be shown to be the smallest $|\Psi(x)|$ among all possible choices of nodes in $[-1, 1]$.

As a consequence of this estimate of $\Psi(x)$, we see that if $p_n(x)$ interpolates a (sufficiently differentiable) function $f(x)$ at the Chebyshev nodes $\{x_0, \ldots, x_n\}$, then for $-1 \leq x \leq 1$ we have the error estimate
\[
|f(x) - p_n(x)| \leq \frac{|f^{(n+1)}(\xi)|}{(n+1)!} \frac{1}{2^n} \leq \frac{\max_{-1 \leq x \leq 1} |f^{(n+1)}(\xi)|}{(n+1)!} \frac{1}{2^n} \hfill (2.5)
\]
and from the standpoint of making the r. h. s. of this estimate as small as possible without knowing anything about $f^{(n)}(x)$ but an estimate of its size, the Chebyshev nodes are the best possible choice for $-1 \leq x \leq 1$.

While we will not use the fact in this course, Hermite interpolation of an arbitrary continuous function on $[-1, 1]$ at the Chebyshev nodes, with the derivative values set equal to zero (it doesn’t matter whether the function is differentiable or not) yields a sequence of polynomials that converges to the function uniformly on $[-1, 1]$. This discovery was made in 1930 by the Hungarian mathematician Leopold Fejér:

**Theorem** (Fejér): Let $f(x)$ be a function continuous on the interval $[-1, 1]$. For each natural number $n$ let $p_{2n+1}(x) \in P_{2n+1}$ be the polynomial for which $p_{2n+1}(x_i) = f(x_i)$ and $p'_{2n+1}(x_i) = 0$, where $\{x_0, \ldots, x_n\}_{i=0}^n$ are the zeros of the $(n+1)$-st Chebyshev polynomial $T_{n+1}$. Then the sequence of polynomials $\{p_{2n+1}(x)\}_{n=1}^\infty$ converges to the function $f(x)$ uniformly on $[-1, 1]$; i.e., for any preassigned error tolerance $\epsilon > 0$ there is a natural number $N$ such that if $n \geq N$ then $|p_{2n+1}(x) - f(x)| < \epsilon$ holds for all $-1 \leq x \leq 1$.

If $\{\ell_{i,n}\}_{i=0}^n$ are the Lagrange polynomials constructed for the the Chebyshev nodes that are the zeros $\{x_0, \ldots, x_n\}$ of $T_{n+1}(x)$, and
\[
A_{i,n}(x) = \left[1 - 2(x-x_i)\ell_{i,n}^2(x)\right] \ell_{i,n}(x) \quad \text{for } 0 \leq i \leq n
\]
\[
B_{i,n}(x) = (x-x_i) \ell_{i,n}(x) \quad \text{for } 0 \leq i \leq n
\] \hfill (2.6)
are the $2n + 2$ basis polynomials used in Hermite interpolation at those nodes, then an explicit formula for the polynomial of Fejér’s theorem is furnished by
\[
p_{2n+1}(x) = \sum_{i=0}^n f(x_i) A_{i,n}(x) . \hfill (2.7)
\]

A similar expression can be given using Newton polynomials and divided differences. However, the computation of these values is too expensive, and the convergence of the approximations too slow, to justify using this theorem for machine approximation as of this writing.


---

(26) This will follow from the characterization of best approximating polynomials in the $\| \cdot \|_\infty$-norm below; or read Atkinson’s treatment on pp. 229–231. Kincaid & Cheney also contains other information about the Chebyshev polynomials and nodes; see their index.
3. Orthogonality of Chebyshev Polynomials. The relations between the Chebyshev polynomials and the cosines imply that these polynomials satisfy the integral orthogonality relations

\[
\int_{-1}^{1} T_j(x) T_k(x) \frac{dx}{\sqrt{1-x^2}} = \begin{cases} 
\pi & \text{if } j = k = 0; \\
\frac{\pi}{2} & \text{if } j = k \neq 0; \\
0 & \text{otherwise}.
\end{cases}
\] (3.1)

These follow simply by taking the trigonometric orthogonality relations

\[
\int_{0}^{\pi} \cos jt \cos kt \, dt = \begin{cases} 
\pi & \text{if } j = k = 0; \\
\frac{\pi}{2} & \text{if } j = k \neq 0; \\
0 & \text{otherwise}.
\end{cases}
\] (3.2)

and making the change of variable \(x = \cos t\). From this fact it follows that the Chebyshev polynomials are in fact the orthogonal polynomials produced by the weight function \(w(x) = \frac{1}{\sqrt{1-x^2}}\), so of course everything we know about orthogonal polynomials applies to them. They also satisfy certain discrete orthogonality relations that they inherit from the corresponding relations for the cosine functions.

4. Chebyshev Economization. This is a technique for decreasing the degrees of approximating polynomials while keeping error of approximation acceptably small. It uses the fact that the relation (2.4) above can be rewritten—in a rather “unofficial” form—as

\[
\max_{-1 \leq x \leq 1} |x^{n+1} - \frac{1}{2^n} \cdot (the \ rest \ of \ T_{n+1}(x))| \leq \frac{1}{2^n}.
\] (4.1)

For example, since \(T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1\), we have

\[
\max_{-1 \leq x \leq 1} |x^6 - \left[ \frac{3}{2} x^4 - \frac{9}{16} x^2 + \frac{1}{32} \right]| \leq \frac{1}{32}.
\] (4.2)

To exhibit a rather simple-minded application that illustrates the way that this fact is exploited: it follows from the estimate just given that in the Taylor-series approximation

\[
\cosh x = 1 + x^2 + \frac{x^4}{24} + \frac{x^6}{720} + R_7(x)
\] (4.3)

for \(-1 \leq x \leq 1\), in which the error term \(R_7(x)\) is absolutely smaller than \(\frac{\cosh 1}{40320} \approx 3.827 \times 10^{-5}\), we could replace the \(x^6\) by \(\left[ \frac{3}{2} x^4 - \frac{9}{16} x^2 + \frac{1}{32} \right]\) and yet introduce an additional absolute error of only at most \(\frac{1}{720} \cdot \frac{1}{32} \approx 4.340 \times 10^{-5}\). This might be perfectly acceptable in some situations: the total absolute error would be at most \(8.167 \times 10^{-5}\). By contrast, the usual error estimate for the plain 4-th degree Taylor polynomial would give \(\frac{\cosh 1}{720} \approx 2.143 \times 10^{-3}\), over 25 times as large. We can almost have the benefit of the 6-th degree Taylor polynomial while only having to evaluate a polynomial of degree 4.

Of course, given an error tolerance \(\epsilon > 0\) to start with, one could iterate the process until the error estimate would be larger than \(\epsilon\). For more on this subject, as well as examples, see, e.g., R. W. Hamming, Numerical Methods for Scientists and Engineers, Ch. 29, p. 483 ff.
is minimal among all polynomials of degree \( \leq n \) that could approximate \( x^{n+1} \) “in the mean”—i.e., in the \( L^1[-1,1] \) norm. Thus if one wanted to do Chebyshev economization in the mean, one would use the Chebyshev polynomials of the second kind. Minimizing in the \( L^2[-1,1] \) norm, of course, would be accomplished with “the rest of the Legendre polynomial \( P_{n+1}(x) \).”

5. A Note on Change of Scale. The entire development above has taken place on the interval \([-1,1]\). However, one can transfer the results to an arbitrary interval \([a,b]\) (with \( a < b \)) with little work. If \( f(x) \) is defined on \([a,b]\), set \( z = \frac{1}{2}[(1-x)a + (1+x)b] \) and observe that as \( x \) goes from \(-1\) to \(1\), the corresponding \( z \) goes from \( a \) to \( b \) and \( F(x) = f\left(\frac{1}{2}[(1-x)a + (1+x)b]\right) \) takes the values that \( f(z) \) did in that interval. If Chebyshev methods are used to approximate \( F(x) \) on \([-1,1]\)—say \( p_n(x) \) is a polynomial approximation to \( F(x) \)—then the function \( p_n\left(\frac{1}{b-a}(2z-b-a)\right) \), which is a polynomial of the same degree \( n \) in the variable \( z \), approximates \( f(z) \) just as well on \([a,b]\) as \( p_n(x) \) approximated \( F(x) \) on \([-1,1]\).

4. Polynomial Approximation in the Uniform Norm.

Here I really must deviate from A.’s §§4.6-4.7, and you should delay reading him until you read the treatment of this matter below. In particular, §4.6 will become easy, and you will be able to see that the proof of A.’s Theorem 4.9, p. 222, is really a specialization of an argument that you already know and that is easier to understand if it is given in full generality.

4.A. Understanding the Problem. When one does polynomial approximation in the uniform norm, everything takes place in the context of the following general problem: given a continuous real-valued function \( f \) defined on a finite interval \([a,b] \subseteq \mathbb{R} \), and given a tolerable error \( \epsilon > 0 \), find a polynomial \( p(x) \) of degree at most \( n \) for which

\[
\|f - p\|_{\infty} = \max\{|f(x) - p(x)| : a \leq x \leq b\} < \epsilon.
\]

There is no loss of generality in taking \([a,b] = [-1,1]\), and we shall do that from now on; and there are obvious aesthetic and practical reasons for trying to have the degree \( n \) as small as possible consistent with the desired small error tolerance.

This problem is obviously not always solvable if one insists on specifying the degree of \( p(x) \) in advance! For example, no constant \((n = 0)\) function will approximate \( f(x) \equiv x \) on \([-1,1]\) within \( \epsilon < 1/2 \). It therefore makes sense to start one’s consideration of the problem by making “pessimistic estimates.” So, as before, we let \( \mathcal{P}_n \) denote the \((n+1)\)-dimensional real vector space of all polynomial functions of degree \( \leq n \) (with coefficients in \( \mathbb{R} \)), and given a function \( f \in \mathcal{C}([-1,1]) \) we put

\[
\rho_n(f) = \min\{\|f - p_n(x)\|_{\infty} : p_n \in \mathcal{P}_n\}.
\]

Then our original problem can only be solved at all if \( \rho_n(f) < \epsilon \). It thus makes sense to try to estimate \( \rho_n(f) \) from below, since every such estimate will give a necessary condition that \( n \) and \( \epsilon > 0 \) must satisfy if there is to be any chance of solving the original problem with a polynomial \( p(x) \) of degree \( \leq n \).

It may help to make things more concrete if we realize that there is always at least one \( p_n(x) \) for which the greatest lower bound that should have been used to define \( \rho_n(f) \) is attained. {Non-mathematicians are excused from filling in the details of the following argument and may read over it very lightly. We really should have written \( \rho_n(f) = \inf\{\|f - p_n(x)\|_{\infty} : p_n \in \mathcal{P}_n\} \), because without further investigation it is not clear that there is a minimizing polynomial. But in fact, all one has to do is to observe that the function

\[
\mathbb{R}^{n+1} \times [-1,1] \to \mathbb{R}
\]

\[
((a_0, \ldots, a_n), x) \mapsto f(x) - a_0 - a_1 x - \cdots - a_n x^n
\]

is (jointly) continuous (in all its arguments), and therefore that

\[
(a_0, \ldots, a_n) \mapsto \max\{|f(x) - a_0 - a_1 x - \cdots - a_n x^n| : -1 \leq x \leq 1\}
\]
is continuous on \( \mathbb{R}^{n+1} \). It is easy to see that the r. h. s. of the last set-off line can be made arbitrarily large provided only that the \( \ell^\infty \)-norm \( \|(a_0, \ldots, a_n)\|_{\ell^\infty} \) of the vector of coefficients is made large (try an induction on \( n \)). It follows that the function \( (a_0, \ldots, a_n) \mapsto \max\{|f(x) - a_0 - a_1 x - \cdots - a_n x^n| : -1 \leq x \leq 1\} \) attains its minimum in a suitably large bounded set in \( \mathbb{R}^{n+1} \), and that establishes the existence of a polynomial of degree \( n \) that is nearest to \( f \) in the \( \| \cdot \|_\infty \) norm. Functional analysts will see that the fact that any \((n+1)\)-dimensional subspace of a real Banach space is norm-isomorphic (though not in general isometric) to \( \mathbb{R}^{n+1} \) can be used to make things even easier.

Let us agree to call a polynomial \( p \) of degree \( \leq n \) that minimizes \( \| f - p \|_\infty \) over all \( p \in \mathcal{P}_n \) an **optimal polynomial** (for the function \( f \)). This may not be the standard name, but it gets the idea across.

One way of getting estimates for \( \rho_n(f) \) from below is to use the fact that for any \((n + 2)\) points \( \{x_0, \ldots, x_{n+1}\} \subseteq [-1, 1] \), the divided-difference operation

\[
\mathcal{E}([-1, 1]) \to \mathbb{R} \\
f \mapsto f[x_0, \ldots, x_{n+1}]
\]

"kills polynomials of degree \( \leq n \)." Two ways of seeing this are: first, if we interpolate a polynomial of degree \( \leq n \) at \((n + 2)\) points, then the polynomial must come back as its own interpolant. But the coefficient of \( x^n \) to \( 0 \). So we should now be convinced.

For any \((n + 2)\) points \( \{x_0, \ldots, x_{n+1}\} \subseteq [-1, 1] \) and any polynomial \( p_n(x) \) of degree \( \leq n \) we can now write (by "subtracting zero")

\[
|f[x_0, \ldots, x_{n+1}]| = |f[x_0, \ldots, x_{n+1}] - p_n[x_0, \ldots, x_{n+1}]| = \| f - p_n \|_{[x_0, \ldots, x_{n+1}]} \tag{4.A.1.1}
\]

\[
= \sum_{j=0}^{n+1} \left| \frac{f(x_j) - p_n(x_j)}{\Psi(x_j)} \right| \cdot \frac{1}{\Psi(x_j)} \quad \text{where } \Psi(x) = (x-x_0) \cdots (x-x_{n+1}) \tag{4.A.1.2}
\]

\[
\leq \sum_{j=0}^{n+1} \left| \frac{1}{\Psi(x_j)} \right| \cdot \left| f(x_j) - p_n(x_j) \right| \leq \sum_{j=0}^{n+1} \left| \frac{1}{\Psi(x_j)} \right| \cdot \| f - p_n \|_{\infty} \tag{4.A.1.3}
\]

so if we put \( W(x_0, \ldots, x_{n+1}) = \sum_{j=0}^{n+1} \left| \frac{1}{\Psi(x_j)} \right| \), this gives us (writing \( W \) instead of \( W(x_0, \ldots, x_{n+1}) \) when the choice of points is understood)

\[
\frac{|f[x_0, \ldots, x_{n+1}]|}{W} \leq \| f - p_n \|_{\infty}
\]

and by the definition of infimum (or because \( p_n(x) \) could have been taken to be an optimal polynomial) we have the fundamental estimate-from-below

\[
\frac{|f[x_0, \ldots, x_{n+1}]|}{W} \leq \rho_n(f). \tag{4.A.2}
\]

People who know some functional analysis will realize that this construction is a concrete version of the following simple theorem about the geometry of normed spaces: if \( \mathcal{M} \) is a closed subspace of the normed space \( (X, \| \cdot \|) \) and if \( f \in X \), then the distance from \( f \) to \( \mathcal{M} \) is the norm of its coset in the quotient space \( X/\mathcal{M} \). The dual of \( X/\mathcal{M} \) is the annihilator \( \mathcal{M}^\circ \) of \( \mathcal{M} \), and so for any \( F \in \mathcal{M}^\circ \) of norm 1 we have

\[
|F(f)| = |F(f + \mathcal{M})| \leq \| f + \mathcal{M} \| = \inf \{ \| f - m \| : m \in \mathcal{M} \}. 
\]
In this case $X = \mathcal{C}([-1,1])$, $\mathcal{M} = \mathcal{P}_n$, and the effect of dividing by $W$ is to make the functional $f \mapsto f[x_0, \ldots, x_{n+1}]/W$ have norm $1$.

Even without the functional-analytic conceptualization, readers will see that the following argument is similar to the one just given (from the functional-analytic viewpoint it is identical), and gives us another way of making pessimistic estimates of $\rho_n(f)$: let $\varphi(x)$ be a continuous (Lebesgue-integrable would do) function on $[-1,1]$ for which $\int_{-1}^{1} p_n(x) \varphi(x) w(x) \, dx = 0$ holds for every $p_n \in \mathcal{P}_n$. (For example, $\varphi(x)$ could be the $\varphi_{n+1}(x)$ from the sequence of orthogonal polynomials given by the weight function $w(x)$.) Then

$$\left| \int_{-1}^{1} f(x) \varphi(x) w(x) \, dx \right| = \left| \int_{-1}^{1} [f(x) - p_n(x)] \varphi(x) w(x) \, dx \right| \leq \left\{ \int_{-1}^{1} |\varphi(x)| w(x) \, dx \right\} \cdot \|f - p_n\|_\infty$$

so if we set $W$ equal to the integral inside curly braces in the last set-off line above, we again get the estimate

$$\frac{1}{W} \left| \int_{-1}^{1} f(x) \varphi(x) w(x) \, dx \right| \leq \|f - p_n\|_\infty,$$

and since $p_n(x)$ could have been the optimal polynomial, the l. h. s. of this expression again offers an estimate from below for $\rho_n(f)$.

But back to the construction with the divided differences. The beginning of wisdom about optimizing the choice of $p_n(x)$ begins with trying to see under what circumstances the lower estimate

$$\frac{|f[x_0, \ldots, x_{n+1}]|}{W} \leq \rho_n(f) \quad (4.4.2)$$

turns into an equality because the absolute-value calculations don’t allow any cancellation to take place.

**Theorem [de la Vallée-Poussin](27):** Suppose that $f \in \mathcal{C}([-1,1])$, that $p_n \in \mathcal{P}_n$, and that $x_0 < x_1 < \cdots < x_{n+1}$ are $(n + 2)$ points in $[-1,1]$ for which

$$f(x_j) - p_n(x_j) = (-1)^j \cdot e_j \quad j = 0, \ldots, n + 1$$

with all the $e_j$’s of the same sign. Then

$$\rho_n(f) \geq \min\{|e_j|: 0 \leq j \leq n + 1\}.$$

**Proof.** Just as when we proved the estimate (4.4.1), we have

$$|f[x_0, \ldots, x_{n+1}]| = \sum_{j=0}^{n+1} \{f(x_j) - p_n(x_j)\} \cdot \frac{1}{\Psi'(x_j)} \quad \text{where } \Psi(x) = (x - x_0) \cdots (x - x_{n+1}).$$

But now we observe that the signs of both the factors in $\{f(x_j) - p_n(x_j)\} \cdot \frac{1}{\Psi'(x_j)}$ alternate with each increase in the index $j$: the first alternates by explicit hypothesis, while the second alternates because at each successive node in the list $x_0 < x_1 < \cdots < x_{n+1}$ the sign of the error factor $\Psi(x) = (x - x_0) \cdots (x - x_{n+1})$ changes, and therefore the values $\Psi'(x_j)$ of the derivative must be alternately positive and negative. It follows

---

(27) This is Atkinson’s Theorem 4.9, p. 222 ff.
that all the terms of the sum inside the absolute value bars in 
\[ \sum_{j=0}^{n+1} \{ f(x_j) - p_n(x_j) \} \cdot \frac{1}{\Psi'(x_j)} \] have the same 
sign, and thus the absolute value can be computed term-by-term: we have 
\[ \left| \sum_{j=0}^{n+1} \{ f(x_j) - p_n(x_j) \} \cdot \frac{1}{\Psi'(x_j)} \right| = \sum_{j=0}^{n+1} \left| \{ f(x_j) - p_n(x_j) \} \right| \cdot \frac{1}{\Psi'(x_j)} = \sum_{j=0}^{n+1} \frac{1}{\Psi'(x_j)} \cdot |e_j| \]
where before we had only the inequality \( \leq \). We can now estimate the r. h. s. from below by 
\[ \sum_{j=0}^{n+1} \frac{1}{\Psi'(x_j)} \cdot |e_j| \geq W(x_0, \ldots, x_{n+1}) \cdot \min\{|e_j| : 0 \leq j \leq n + 1\}. \]
Dividing both sides by \( W \) and using (4.A.2) we have 
\[ \rho_n(f) \geq \frac{|f[x_0, \ldots, x_{n+1}]|}{W} \geq \min\{|e_j| : 0 \leq j \leq n + 1\}, \quad (4.A.3) \]
as advertised.

The differences between this approach and that of Atkinson’s Theorem 4.9 are twofold. First, this 
approach specializes a sharper inequality to give the de la Vallée-Poussin theorem. Second, it gives the 
sufficient condition for optimality that goes with this necessary one.

**Corollary [Chebyshev sufficient condition for optimality]**: If all the absolute errors \( |e_j| \) of the 
de la Vallée-Poussin theorem are equal to \( \| f - p_n \|_\infty \), then \( p_n(x) \) is an optimal polynomial, i.e., it is at 
minimum distance \( \rho_n(f) \) from \( f \) among all polynomials of degree \( \leq n \).

**Proof.** We can then finish the inequality (4.A.3) by writing 
\[ \rho_n(f) \geq \frac{|f[x_0, \ldots, x_{n+1}]|}{W} \geq \min\{|e_j| : 0 \leq j \leq n + 1\} = \| f - p_n \|_\infty \geq \rho_n(f) \quad (4.A.4) \]
(where the last inequality is trivial!) and we’re done.

Note that this is the sufficient condition for optimality promised (but not delivered) by Atkinson’s 
**Theorem 4.10**, p. 224. It’s true that the proof that the optimal \( n \)-th-degree polynomial always possesses 
interpolation points \( x_0 < \cdots < x_{n+1} \) at which this “equioscillation condition” holds is a little tricky, and for 
our purposes it is honest to omit it—generally speaking, in applications we simply want a best approximation, 
and the facts that it is unique and that it must interpolate at certain points in a certain way are interesting 
but relatively less important that the fact that we can show that we have actually got the minimal polynomial.

As an example, let us establish the result promised\(^{28}\) back in our discussion of Chebyshev polynomials above: (1) the zeros of the Chebyshev polynomial \( T_{n+1}(x) \) furnish the smallest \( \| \Psi(x) \|_\infty \) among all choices 
of nodes in \([-1, 1]\), and (2) the best approximant to \( x^{n+1} \) uniformly on \([-1, 1]\) by a polynomial of degree 
\( \leq n \) is furnished by \( \frac{1}{2^n} \cdot (\text{the rest of } T_{n+1}(x)) \) as in (4.1) of \( \S \) 4 above. These two statements are actually 
equivalent under the side condition that the difference polynomial occurring in (2) has enough zeros, because 
if \( \Psi(x) = (x - x_0) \cdot (x - x_{n+1}) = x^{n+1} - [a_n x^n + \cdots + a_0] \) has uniform norm \( \| \Psi \|_\infty = \epsilon \), then the polynomial 
\( a_n x^n + \cdots + a_0 \) approximates \( x^{n+1} \) uniformly within \( \epsilon \); conversely, if the polynomial \( a_n x^n + \cdots + a_0 \) is 
such that \( x^{n+1} - [a_n x^n + \cdots + a_0] \) has \( n + 1 \) zeros in \([-1, 1]\) and approximates \( x^{n+1} \) at distance \( \epsilon \), then 
we can take \( \Psi(x) = x^{n+1} - [a_n x^n + \cdots + a_0] \) and have \( \| \Psi \|_\infty = \epsilon \). So all we have to do is check that 
\( x^{n+1} - \frac{1}{2^n} \cdot (\text{the rest of } T_{n+1}(x)) \) satisfies the Chebyshev sufficient condition: but of course it does, because it behaves on \([-1, 1]\) in exactly the same way that \( \frac{1}{2^n} \cos(n + 1)t \) behaves on \([0, \pi]\), hitting 
\( \pm 1/2^n \) alternately at the \( (n + 2) \) successive extrema of \( \cos(n + 1)t \) on that interval.

\(^{28}\) In footnote 26, p. 30.
4.B. **Approximation by Interpolation.** In most application situations, one would try to produce polynomial approximants to \( f(x) \) by polynomial interpolation of \( f \) at suitably chosen points \( x_0 < \cdots < x_n \); the amount of “information” required is limited to a few evaluations of \( f \), the computation of the interpolating polynomial in Newton form is short and efficient, and no integrals have to be computed (or approximated). So you may find it nice to have the results of this §, even though they’re not in Atkinson’s book. Rather than making all the definitions required to state the principal theorem, let me work up to the theorem so that you can see how the definitions grow organically.

The basic idea is to take an \( n \)-th degree polynomial interpolant \( p_n(x) \) of \( f(x) \) on \([-1,1]\) and estimate the error of interpolation in terms of \( \rho_n(f) \). Thus let \( p_n \) interpolate \( f \) at the points \( \{x_0, \ldots, x_n\} \), and recall the error expression

\[
f(x) - p_n(x) = f[x_0, \ldots, x_n, x](x-x_0)\cdots(x-x_n) .
\]

Take any point \( x \in [-1,1] \) but call it \( x_{n+1} \), and this expression becomes

\[
|f(x_{n+1}) - p_n(x_{n+1})| = |f[x_0, \ldots, x_n, x_{n+1}]| \cdot |x_{n+1} - x_0| \cdots |x_{n+1} - x_n| . \tag{4.B.1}
\]

The first thing we did in the preceding § was to observe that

\[
f[x_0, \ldots, x_{n+1}] = \{f - p_n^x\}[x_0, x_{n+1}]
\]

for any choice of \( p_n^x \in \mathcal{P}_n \), because taking \((n + 1)\)-st divided differences kills polynomials of degree \( \leq n \); we thus got the estimate (4.A.1)

\[
|f[x_0, \ldots, x_{n+1}]| \leq \left[ \sum_{j=0}^{n+1} \frac{1}{\Psi'(x_j)} \right] \cdot \rho_n(f)
\]

because \( p_n^x \) was arbitrary (or could have been optimal). Plugging this estimate into (4.B.1) above and bringing the factor \(|x_{n+1} - x_0| \cdots |x_{n+1} - x_n| \) into the sum gives

\[
|f(x_{n+1}) - p_n(x_{n+1})| \leq \left[ \sum_{j=0}^{n+1} \frac{|x_{n+1} - x_0| \cdots |x_{n+1} - x_n|}{\Psi'(x_j)} \right] \cdot \rho_n(f) . \tag{4.B.2}
\]

We now have to analyze the complicated-looking sum with the \( \Psi'(x_j) \)'s in its denominator. Recall that \( \Psi(t) = (t-x_0)\cdots(t-x_{n+1}) \), and so \( \Psi'(t) \) is the sum of \((n + 2)\) terms, each of the factors having been differentiated once:

\[
\Psi'(t) = \sum_{i=0}^{n+1} \prod_{k \neq i} (t-x_k) .
\]

However, if you plug in \( t = x_j \), then the only term without a zero in the indicated product is the \( j \)-th, so

\[
\Psi'(x_j) = \prod_{0 \leq k \leq n+1, k \neq j} (x_j - x_k) \quad \text{and}
\]

\[
|\Psi'(x_j)| = \prod_{0 \leq k \leq n+1, k \neq j} |x_j - x_k| .
\]

In the \( j = n + 1 \) term of (4.B.2) this means that everything cancels multiplicatively:

\[
\frac{|x_{n+1} - x_0| \cdots |x_{n+1} - x_n|}{\Psi'(x_{n+1})} = \frac{|x_{n+1} - x_0| \cdots |x_{n+1} - x_n|}{|x_{n+1} - x_0| \cdots |x_{n+1} - x_n|} = 1 .
\]

In the \( 0 \leq j \leq n \) terms of (4.B.2), however, the quotient instead becomes

\[
\frac{|x_{n+1} - x_0| \cdots |x_{n+1} - x_n|}{|\Psi'(x_j)|} = \frac{|x_{n+1} - x_0| \cdots |x_{n+1} - x_n|}{|x_j - x_0| \cdots |x_j - x_{n+1}|} (|x_j - x_j| \text{ missing})
\]

36
and because we don’t have to be concerned with the ± signs, we can cancel the \(|x_{n+1} - x_j|\) with the \(|x_j - x_{n+1}|\) in the denominator to find that this quotient in fact is

\[
\prod_{0 \leq k \leq n, k \neq j} \left| \frac{x_{n+1} - x_j}{x_j - x_k} \right|.
\]

Now this should look familiar. The \(j\)-th Lagrange polynomial for the nodes at which \(p_n(x)\) was originally interpolated, namely \(\{x_0, \ldots, x_n\}\), is given by

\[
\ell_j(x) = \prod_{0 \leq i \leq n, i \neq j} \frac{x - x_j}{x_i - x_j}
\]

with the \(j\)-th term of the sum is just \(|\ell_j(x)|\) with \(x = x_{n+1}\) plugged in. We can now rewrite (4.B.2) as

\[
|f(x_{n+1}) - p_n(x_{n+1})| \leq \left[ 1 + \sum_{j=0}^{n} |\ell_j(x_{n+1})| \right] \cdot \rho_n(f).
\]  

(4.B.3)

But \(x_{n+1}\) was arbitrary, so we might as well call it \(x\) again. Define the **Lebesgue function for the nodes** \(\{x_0, \ldots, x_n\}\) by

\[
\Lambda_n(x) = 1 + \sum_{j=0}^{n} |\ell_j(x)|
\]

and put \(\Lambda_n = \max\{\Lambda_n(x) : x \in [-1, 1]\}\). Then (4.B.3) implies

\[
|f(x) - p_n(x)| \leq \left[ 1 + \sum_{j=0}^{n} |\ell_j(x)| \right] \cdot \rho_n(f) \leq [1 + \Lambda_n] \cdot \rho_n(f).
\]  

(4.B.4)

Take the maximum for \(x \in [-1, 1]\) on the l. h. s. of (4.B.4), and you’ll see that we have proved

**Theorem:** If \(p_n(x)\) is the \(n\)-th degree polynomial that interpolates \(f(x)\) at \(\{x_0, \ldots, x_n\}\), then the distance between \(p_n(f)\) and \(f\) in the uniform norm of \(C([-1, 1])\) is estimated by

\[
\|f - p_n\|_{\infty} \leq [1 + \Lambda_n] \cdot \rho_n(f)
\]  

(4.B.5)

where \(\Lambda_n\) is the maximum of the Lebesgue function for the nodes at which the polynomial \(p_n(x)\) interpolates \(f(x)\).

To make this theorem useful, we need some results from some heavy hitters. I am not going to give proofs of these, as I would have to do in a course or seminar devoted purely to approximation theory. The results are

(1) **The theorems of D. Jackson**, which should be familiar from Atkinson pp. 224–225: if \(f \in C^{(k)}([a, b])\), and if moreover there are constants \(K > 0\) and \(0 < \alpha \leq 1\) for which

\[
|f^{(k)}(x) - f^{(k)}(t)| \leq K \cdot |x - t|^\alpha \text{ for all } x, t \in [a, b],
\]

then

\[
\rho_n(f) \leq \frac{d_{k, \alpha} \cdot K}{n^{k+\alpha}}
\]

where the constant \(d_{k, \alpha}\) is independent of \(f\) and \(n\). (See Atkinson *loc. cit.*, where he doesn’t prove it either. If you are interested in seeing a proof, try §1.1 of T. J. Rivlin, *An Introduction to the Approximation of Functions*, Dover, 1981. The vector space of all functions \(f(x)\) satisfying conditions of the four above for fixed \(k\) and \(\alpha\) is sometimes called \(C^{k, \alpha}([a, b])\) or \(C^{(k, \alpha)}([a, b])\). As you recall, the condition

\[
|g(x) - g(t)| \leq K \cdot |x - t|^\alpha
\]

is called a **Hölder condition** (with exponent \(\alpha\)).}
(2) **Bad Lebesgue constants:** no matter how the nodes are chosen, one always has

\[ \Lambda_n > \frac{2}{\pi} \ln n - c \quad (c \text{ some constant}) \]

so the factor \((1 + \Lambda_n)\) can’t stay bounded above as \(n \to \infty\) [Erdős]. Even worse, for \(n = 2m\) uniformly spaced nodes there are constants \(K_1\) and \(K_2\) for which

\[ K_1 \cdot \left( \frac{3}{2} \right)^m < \Lambda_{2m} < K_2 \cdot 2^{m} e^{2m} . \]

(3) **Good Lebesgue Constants:** if \(\{x_0, \ldots, x_n\}\) are the zeros of the Chebyshev polynomial \(T_{n+1}(x)\), then \(\Lambda_n \leq \frac{2}{\pi} \ln n + 4\). And there are even better choices: if the Chebyshev nodes are “expanded” or “scaled out” so that the end nodes become \(\pm 1\)—i.e., if

\[ x_j = \frac{\cos \left( \frac{(2j + 1)\pi}{2n + 2} \right)}{\cos \left( \frac{\pi}{2n + 2} \right)} \]

then the Lebesgue constants come within 0.201 \cdots of their smallest possible values (though one still has \(\Lambda_n = O(\ln n)\)). (See Rivlin, op. cit., §4.2. The result\(^{(29)}\) about the expanded Chebyshev nodes is less that a quarter-century old, and may still be subject to improvement. For another discussion, see Conti and de Boor, *Elementary Numerical Analysis: ...*, §6.1.)

Putting (1) and (3) together tells us that if \(f \in C^{k,\alpha}([-1,1])\) then we can expect

\[ \|f - p_n\|_\infty \leq \text{const.} \times \left\{ \left[ \frac{2}{\pi} \ln n \right] + 4 \right\} \rightarrow 0 \quad \text{as } n \to \infty \]

when \(p_n(x)\) interpolates \(f(x)\) at the zeros of \(T_{n+1}(x)\) (or their “expanded” versions). So for functions with even the little bit of smoothness represented by a Hölder condition on \(f(x)\) itself, we now have a viable strategy for uniform polynomial approximation via polynomial interpolation, provided that we use the Chebyshev or expanded-Chebyshev nodes. Remember that the Erdős-Turán theorem already told us to expect \(\|\cdot\|_w\text{-convergence}, \text{ where } w(x) = (1 - x^2)^{-1/2}, \text{ on the mere assumption of continuity of } f.\]

*It is important to realize that one should not try to write algorithms to interpolate \(f(x)\) at the Chebyshev nodes by divided-difference methods.* Via the replacement of \(f(x)\) by \(f(\cos \theta)\), the Chebyshev interpolation computation turns into a very “regular” trigonometric interpolation that can be handled with the (fast, finite) Fourier transform. We shall beat all this to death below.

Before taking leave both of the Chebyshev polynomials and of the topic of approximation by interpolation, we should make two more small observations.

First, in the chain of inequalities

\[ \rho_n(f) \geq \left| \frac{f[x_0, \ldots, x_{n+1}]}{W} \right| \geq \min\{|e_j| : -1 \leq x \leq 1\} = \|f - p_n\|_\infty \geq \rho_n(f) \]

that establishes the Chebyshev sufficient condition for \(p_n(x) = x^{n+1} - \frac{1}{2n} T_{n+1}(x)\) to be the best approximant of \(x^{n+1}\) in \(\mathcal{P}_n\), we know all the numbers except \(f[x_0, \ldots, x_{n+1}]\). But in fact, its value is 1: we may observe

that \( f^{(n+1)}(x) \equiv (n+1)! \) and \( f[x_0, \ldots, x_{n+1}] = f^{(n+1)}(x) \) if we must, though it would be better to observe that \( f[x_0, \ldots, x_{n+1}] = 1 \) by definition of divided difference for the function \( f(x) = x^{n+1} \). In any event, since we now know this number we may deduce that

\[
W(x_0, \ldots, x_{n+1}) = 2^n \tag{4.8.6}
\]

holds for the “Chebyshev extrema”—the \( x_j \) which at which \( T_{n+1}(x) \) takes its \((n+2)\) extreme values on \([-1, 1]\). Finding these is no problem, since \( T_{n+1}(\cos \theta) \equiv \cos(n+1)\theta \) takes extreme values at 0 and \( \pi \), while the zeros of the derivative \( -(n+1) \sin(n+1)\theta \) of \( \cos(n+1)\theta \) interior to \((0, \pi)\) occur at \( \theta_j = \frac{j - (n+1)\pi}{n+1}, \ j = 1, \ldots, n \).

Thus

\[
x_j = \cos \left( \frac{j - n - 1}{n+1} \pi \right), \ j = 0, \ldots, (n+1)
\]

do so if \( f \) is sufficiently-smooth. This relation is frequently useful, because these extrema make good test points to use in the lower estimate

\[
\left| \frac{f[x_0, \ldots, x_n]}{W(x_0, \ldots, x_{n+1})} \right| \leq \rho_n(f) \tag{4.8.7}
\]

In his discussion of interpolation at the Chebyshev zeros on p. 228 ff. (but particularly at the bottom of p. 230), Atkinson discusses something I brought up in class a while ago but have not followed up in these notes, namely: if \( \{x_0, \ldots, x_n\} \) are the zeros of \( T_{n+1}(x) \), then (since we know the leading coefficient of \( T_{n+1}(x) \)) the relation

\[
\frac{1}{2^n} T_{n+1}(x) = (x - x_0) \cdots (x - x_n) = \Psi(x)
\]

(where \( \Psi(x) \) is the error factor for interpolation at these nodes) holds, and we know that \( |T_{n+1}(x)| \leq 1 \) for \( x \in [-1, 1] \). It follows that for sufficiently-smooth \( f(x) \) we have

\[
|f(x) - p_n(x)| = \left| \frac{f^{(n)}(\xi)}{(n+1)!} \Psi(x) \right| \leq \frac{\|f^{(n)}\|_{\infty}}{2^n \cdot (n+1)!} \tag{4.8.8}
\]

The r. h. s. of that inequality looks like it stands a good chance of going to zero for large \( n \) (indeed, it must do so if \( f(x) \) is an analytic function whose Taylor-Maclaurin series at zero converges in an interval containing \([-1, 1]\); moreover, no choice of nodes in \([-1, 1]\] can improve on the factor \( 1/2^n \).

Putting (4.8.6), (4.8.7) and (4.8.8) together gives us the amusing

\[
\left| \frac{f^{(n+1)}(\eta)}{2^n \cdot (n+1)!} \right| = \left| \frac{f[x_0, \ldots, x_{n+1}]}{W(x_0, \ldots, x_{n+1})} \right| \leq \rho_n(f) \leq \max_{-1 \leq \xi \leq 1} |f(x) - p_n(x)| \leq \frac{f^{(n+1)}(\xi)}{2^n \cdot (n+1)!} \tag{4.8.9}
\]

in which, of course, the sample points \( \xi, \eta \in [-1, 1] \) are generally different. Perhaps the surprising thing about the l. h. end of the inequality is that a specific choice of interpolation points (the Chebyshev extrema) produces it.

---

(30) Actually it suffices that \( f \) be analytic at each point of \([-1, 1]\), which is not the same thing; but a discussion of this would take us too far into complex analysis.