Interpolating polynomials and divided differences

Distinct points: the Lagrange form

We shall take it as known that a polynomial of degree n has at most n distinct zeros (a proof is given in Lemma 1 below). Given $n+1$ distinct real numbers x_j and any numbers α_j $(0 \le j \le n)$, there is a unique polynomial p of degree at most n satisfying $p(x_j) = \alpha_j$ $(0 \le j \le n)$. The polynomial is unique, since if p_1 and p_2 were two such polynomials, then $p_1 - p_2$ would be zero at each x_j: since it has degree at most n, it can only be zero. Existence can be deduced from the fact that the matrix with entries x_j^k $(0 \le j \le n,\ 0 \le k \le n)$ is non-singular, but it is easy to describe an explicit construction, as follows.

First, let
$$q(x) = (x - x_0)(x - x_1)\dots(x - x_n).$$
Note that $q(x_j) = 0$ for all j and $q(x)$ is of the form $x^{n+1} + c_n x^n + \dots + c_0$, so $q^{(n+1)}(x) = (n+1)!$ for all x.

For each j, write $N_j = \{0, 1, \dots, n\} \setminus \{j\}$, and let
$$q_j(x) = \frac{q(x)}{x - x_j} = \prod_{k \in N_j} (x - x_k),$$
$$r_j(x) = \frac{q_j(x)}{q_j(x_j)} = \prod_{k \in N_j} \frac{x - x_k}{x_j - x_k}.$$
Then r_j is a polynomial of degree n and we have
$$r_j(x_k) = \begin{cases} 1 & \text{if } k = j, \\ 0 & \text{if } k \ne j. \end{cases}$$
So the required polynomial satisfying $p(x_j) = \alpha_j$ for all j is:
$$p(x) = \sum_{j=0}^{n} \alpha_j r_j(x). \tag{1}$$
So, given a function f, there is a unique polynomial p of degree at most n such that $p(x_j) = f(x_j)$ for each j. It is called the "polynomial interpolating f at $x_0, x_1, \dots, x_n$". Expression (1), with $\alpha_j = f(x_j)$, is called the *Lagrange form* of the interpolating polynomial.

Note that since $q(x) = (x - x_j)q_j(x)$, we have $q'(x_j) = q_j(x_j)$.

It is clear, both from uniqueness and from expression (1), that $p(x)$ is independent of the order in which the points x_j are listed.

If $n = 2$, then of course $p(x)$ is the linear function $ax + b$ agreeing with f at x_0 and x_1. Note that $a = [f(x_1) - f(x_0)]/(x_1 - x_0)$.

To find the polynomial in a particular case, it is usually simpler to solve for the coefficients, as in the next example. (A very effective alternative method will be described below.)

Example 1. To interpolate $f(x) = 2^x$ at 0, 1, 2: let the required polynomial be $a + bx + cx^2$. Equating values at $x = 0$, 1 and 2, we get the equations

$$a = 1, \qquad a + b + c = 2, \qquad a + 2b + 4c = 4,$$

hence $b = c = \frac{1}{2}$, so the polynomial is $1 + \frac{1}{2}x + \frac{1}{2}x^2$. Note that $p(\frac{1}{2}) = 1\frac{3}{8}$, while $f(\frac{1}{2}) = \sqrt{2}$.

If $f(x) = x^k$, where $0 \le k \le n$, then of course $p(x)$ is also x^k. In terms of the functions $r_j(x)$, this says:

PROPOSITION 1. *For $0 \le k \le n$, we have $\sum_{j=0}^{n} x_j^k r_j(x) = x^k$. In particular, $\sum_{j=0}^{n} r_j(x) = 1$ for all x.*

COROLLARY. $\displaystyle\sum_{j=0}^{n} \frac{1}{q_j(x_j)} = 0.$

Proof. This is the coefficient of x^{n-1} in $\sum_{j=0}^{n} r_j(x)$, which equals 1. □

Example 2. Let $f(x) = x^{n+1}$. Then, with notation as above, the required $p(x)$ is $f(x) - q(x)$, since this has degree at most n (the x^{n+1} term cancels) and agrees with $f(x)$ at each x_j.

It is clear from (1) that the leading term of $p(x)$ is $a_n x^n$, where

$$a_n = \sum_{j=0}^{n} \frac{f(x_j)}{q_j(x_j)}. \tag{2}$$

We shall see below that the leading coefficient a_n has particular significance.

Repeated points

In this section, we show that it is not hard (apart from the complexities of notation) to extend the notion of interpolating polynomials to the situation where there are repetitions among the points x_j. This version will be used in the application to Simpson's rule. However, the reader who is so inclined could omit, or defer, this section after glancing at Lemma 1 and its corollaries (suitably simplified by ignoring the orders of the zeros).

Suppose, then, that the *distinct* elements among $x_0, x_1, \ldots, x_n$ are listed as $y_0, y_1, \ldots, y_r$, and that k_j of the x_i's are equal to y_j, so that $\sum_{j=0}^{r} k_j = n+1$. By the "polynomial interpolating f at $x_0, x_1, \ldots, x_n$" we mean the polynomial p of degree at most n such that for each j, $p^{(k)}(y_j) = f^{(k)}(y_j)$ for $0 \le k \le k_j - 1$. We require agreement of the function and the first $k_j - 1$ derivatives at y_j. (The term *osculating polynomial* is sometimes used for p.) Of course, f must have at least $K-1$ derivatives, where $K = \max k_j$. We now establish existence and uniqueness of this polynomial.

Let f be a function having at least $k-1$ derivatives. We say that f has a zero of of *order* (or *multiplicity*) k at the point a if $f^{(j)}(a) = 0$ for $0 \le j \le k-1$ and either $f^{(k)}(a) \neq 0$ or $f^{(k)}(a)$ does not exist (this second case is a technicality of no real importance). Note that f' then has a zero of order $k-1$ at a (this even works for $k=1$ if a "zero of order 0" is taken to mean a point that is not a zero!)

Given an interval I, we denote by $Z(f, I)$ the number of zeros of f in I, *counted with their orders*. We will just write $Z(f)$ for $Z(f, \mathbb{R})$. By Rolle's theorem, we have:

LEMMA 1. *For any function f (having enough derivatives) and any interval I, we have $Z(f', I) \ge Z(f, I) - 1$.*

Proof. Let f have a zero of order k_r at a_r $(1 \le r \le n)$, so that $\sum_{r=1}^{n} k_r = Z(f, I)$. Then f' has a zero of order $k_r - 1$ at a_r (with the above comment about order 0): these add up to

$$\sum_{r=1}^{n} (k_r - 1) = Z(f, I) - n.$$

By Rolle's theorem, f' also has at least $n-1$ zeros in the gaps between the points x_r. Together, these two facts give $Z(f', I) \ge Z(f, I) - 1$. □

COROLLARY 1. *If $Z(f, I) \ge n$, then there exists $\xi \in I$ such that $f^{(n-1)}(\xi) = 0$.*

COROLLARY 2. *Let p be a polynomial of degree n. Then $Z(p) \le n$.*

Proof. Let the leading term be $a_n x^n$. If $Z(p) \ge n+1$, then there exists ξ such that $p^{(n)}(\xi) = 0$. But this is not true, since $p^{(n)}(x) = n! a_n$ for all x. □

This establishes uniqueness of the interpolating polynomial. Existence could now be deduced from non-singularity of the (rather unpleasant) matrix corresponding to the implied set of equations for the coefficients. However, as in the case of distinct points, it is much more satisfying to prove it directly, as follows.

LEMMA 2. *Let $f(x) = (x-a)^k g(x)$, where g has at least k derivatives. Then f has a* *zero of order at least k at a, and $f^{(k)}(a) = k!g(a)$.*

Proof. This follows at once from Leibniz's rule for the higher derivatives of a product. □

PROPOSITION 2. *Suppose that x_i $(0 \le i \le n)$ are points of an interval I (possibly with repetitions), and that f has at least $K-1$ derivatives on I, with K as above. Then there is a unique polynomial interpolating f at the points x_i.*

Proof. We prove the statement by induction on n. It is trivial for $n = 0$ (and indeed for $n = 1$). Assume it is correct for a certain n, and let points x_i $(0 \le i \le n+1)$ be given. As above, suppose that points y_j $(0 \le j \le r)$ are such that k_j of the points $x_0, x_1, \dots, x_n$ equal y_j. Let $p(x)$ be the polynomial interpolating f at $x_0, x_1, \dots, x_n$, and let $q(x) = \prod_{i=0}^{n}(x-x_i) = \prod_{j=0}^{r}(x-y_j)^{k_j}$. By Lemma 2, $q(x)$ has a zero of order k_j at y_j, and $q^{(k_j)}(y_j) \neq 0$. Let

$$p_1(x) = p(x) + a_{n+1}q(x),$$

where a_{n+1} is to be chosen (we use this notation because a_{n+1} is the coefficient of x^{n+1}). Clearly, $p_1^{(k)}(y_j) = f^{(k)}(y_j)$ for $0 \le k \le k_j - 1$. We distinguish two cases.

Case 1: x_{n+1} different from all y_j. Then $q(x_{n+1}) \neq 0$, so we can choose a_{n+1} to ensure that $p_1(x_{n+1}) = f(x_{n+1})$.

Case 2: $x_{n+1} = y_j$, say, so k_j+1 of the extended list of x_i's equal y_j. Since $q^{(k_j)}(y_j) \neq 0$, we can choose a_{n+1} to ensure that $p_1^{(k_j)}(y_j) = f^{(k_j)}(y_j)$, which is what is required. □

However, we cannot offer an explicit expression for $p(x)$ corresponding to the Lagrange form for distinct points.

The most extreme case, of course, is when *all* the points x_j coincide. It is then elementary that the interpolating polynomial is the Taylor expansion

$$\sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k,$$

and it is clear that this is indeed the expression generated by the proof of Proposition 2.

Example 3. To interpolate $f(x) = 2^x$ at 0, 1, 1, 2 (equally, at 0, 1, 2, 1). We saw in Example 1 that the polynomial interpolating f at 0, 1, 2 is $p(x) = 1 + \frac{1}{2}x + \frac{1}{2}x^2$. Let $q(x) = x(x-1)(x-2)$. Then the required polynomial is $p_1(x) = p(x) + a_3q(x)$, with a_3 chosen so that $p_1'(1) = f'(1) = 2\log 2$. Now $p'(1) = \frac{3}{2}$ and $q'(1) = -1$, so $a_3 = \frac{3}{2} - 2\log 2$. Note that $q(\frac{1}{2}) = \frac{1}{2}(-\frac{1}{2})(-\frac{3}{2}) = \frac{3}{8}$, hence $p_1(\frac{1}{2}) \approx 1.4176$.

Estimation of the leading coefficient and the error

We now apply Rolle's theorem (more exactly, Corollary 1 of Lemma 1) to give estimations for the leading coefficient a_n and the "error" $f(x) - p(x)$.

PROPOSITION 3. *Suppose that the points $x_0, x_1, \ldots, x_n$ (some possibly repeated) lie in an interval I and that f is n times differentiable on I. Let $p(x)$ be the polynomial interpolating f at the points x_j, and let the leading term of $p(x)$ be $a_n x^n$. Then there exists a point ξ in I such that*

$$a_n = \frac{f^{(n)}(\xi)}{n!}.$$

Proof. We have $Z(f - p, I) \geq n + 1$: this is obvious when the points x_j are distinct, and follows from our definition of the interpolating polynomial when there are repetitions. By Corollary 1 of Lemma 1, it follows that there exists $\xi \in I$ such that $f^{(n)}(\xi) - p^{(n)}(\xi) = 0$. But $p^{(n)}(x) = n! a_n$ for all x, so $a_n = f^{(n)}(\xi)/n!$. $\square$

The case $n = 1$ (with $x_1 \neq x_0$) equates to the mean-value theorem, since $a_1 = [f(x_1) - f(x_0)]/(x_1 - x_0)$.

The estimation of $f(x) - p(x)$ is derived by a slight elaboration of the same reasoning:

THEOREM 4. *Suppose that the points $x_0, x_1, \ldots, x_n$ (some possibly repeated) lie in an interval I and that f is $n + 1$ times differentiable on I. Let $p(x)$ be the polynomial interpolating f at the points x_j, and let $q(x) = \prod_{j=0}^{n}(x - x_j)$. Then, given a point x in I, there exists ξ in I such that*

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

Proof. Choose a point x^* in I, different from all the x_j. We will show that the given statement applies with $x = x^*$. Define

$$G(x) = f(x) - p(x) - kq(x),$$

with k chosen so that $G(x^*) = 0$, in other words, $f(x^*) - p(x^*) = kq(x^*)$. We need to show that $k = f^{(n+1)}(\xi)/(n+1)!$ for some ξ. Now $Z(G, I) \geq n+2$. In the case where the points x_j are distinct, this is simply because G is zero at $x_0, \ldots, x_n$ and x^*. When there are repetitions (for readers interested in this case), it follows from the fact in the notation of Proposition 2, that $q(x)$, and hence also $G(x)$, has a zero of order at least k_j at y_j for each j. By Corollary 1, it follows that there is a point ξ in I such that $G^{(n+1)}(\xi) = 0$. But $p^{(n+1)}(x) = 0$ (since p has degree at most n) and $q^{(n+1)}(x) = (n+1)!$ for all x. So $0 = G^{(n+1)}(\xi) = f^{(n+1)}(\xi) - (n+1)!k$, hence $k = f^{(n+1)}(\xi)/(n+1)!$, as required. $\square$

So if we know that $f^{(n+1)}(x)$ is between m and M for $x \in I$, then $f(x)-p(x)$ is between $(m/(n+1)!)\,q(x)$ and $(M/(n+1)!)\,q(x)$. It makes sense for $q(x)$ to appear in the estimation, since of course the error is 0 at each x_j.

Example 4. Revisit Examples 1 and 3. In Example 1, $f^{(3)}(x) = 2^x(\log 2)^3$, which is between $(\log 2)^3$ and $4(\log 2)^3$ for x in $[0, 2]$. Also, $q(\frac{1}{2}) = \frac{3}{8}$. So the bounds for $f(\frac{1}{2}) - p(\frac{1}{2})$ given by Theorem 4 are $\frac{1}{16}(\log 2)^3 \approx 0.021$ and $\frac{1}{4}(\log 2)^3 \approx 0.083$. As we saw, the actual value is ≈ 0.039.

In Example 3, we apply $f^{(4)}(x) = 2^x(\log 2)^4$, and $q(x)$ is $x(x-1)^2(x-2)$, so $q(\frac{1}{2}) = \frac{3}{16}$. The bounds are $\frac{1}{128}(\log 2)^4 \approx 0.0018$ and $\frac{1}{32}(\log 2)^4 \approx 0.0072$. The actual value is ≈ 0.0034.

Note. Example 2 is actually a special case of Theorem 4, since if $f(x) = x^{n+1}$, then $f^{(n+1)}(x) = (n+1)!$. In this case, the error estimation is exact.

Application: error estimates for the trapezium rule and Simpson's rule

(This section could be deferred.) The trapezium rule estimates $\int_a^b f$ by $T(f) = \frac{1}{2}(b-a)[f(a) + f(b)]$, the integral of the linear function $p(x)$ interpolating f at a and b. From Theorem 4, we can derive the following bounds for its error:

PROPOSITION 5. *Suppose that $m \le f''(x) \le M$ for $x \in [a, b]$, and let $T(f)$ be as above. Then*

$$\tfrac{1}{12}m(b-a)^3 \le T(f) - \int_a^b f \le \tfrac{1}{12}M(b-a)^3.$$

Proof. Let $m \le f''(x) \le M$ for $x \in [a, b]$, and let $p(x)$ be as above. Apply Theorem 4 with $n = 1$: $q(x)$ is $(x-a)(x-b)$, so (reversing signs) we obtain

$$\tfrac{1}{2}m(x-a)(b-x) \le p(x) - f(x) \le \tfrac{1}{2}M(x-a)(b-x)$$

for $x \in [a, b]$. Writing $b - a = h$, we have

$$\int_a^b (x-a)(b-x)\,dx = \int_0^h y(h-y)\,dy = \left[\tfrac{1}{2}hy^2 - \tfrac{1}{3}y^3\right]_0^h = \tfrac{1}{6}h^3.$$

The stated inequalities follow. □

Because of the intermediate value property of derivatives (which does not require continuity), one can restate the result as follows: there exists $\xi \in [a, b]$ such that $T(f) - \int_a^b f = \frac{1}{12}(b-a)^3 f''(\xi)$.

Note that when $f''(x) \ge 0$ (so that f is convex) on $[a, b]$, this result reproduces the geometrically obvious fact that $T(f) \ge \int_a^b f$.

For the discussion of Simpson's rule, we denote the interval in question by $I = [a-h, a+h]$. The integral $\int_I f$ is approximated by $S(f,I) =: \frac{h}{3}[f(a-h)+4f(a)+f(a+h)]$. This equals the integral exactly if f is a quadratic or cubic polynomial, as one can easily check. So $S(f,I)$ is the integral of (a) the quadratic interpolating f at $a-h$, a, $a+h$, or (b) the cubic interpolating f at $a-h$, a, a, $a+h$. (It is *not* the integral of the cubic interpolating f at four equally spaced points!)

In general, one would expect the error estimate derived from (b) to be sharper. This is indeed the one presented as the standard result in most books. The statement is as follows:

PROPOSITION 6. *Let $I = [a-h, a+h]$. Suppose that $m \le f^{(4)}(x) \le M$ for $x \in I$, and let $S(f,I)$ be as above. Then*

$$\frac{1}{90} m h^5 \le S(f,I) - \int_I f \le \frac{1}{90} M h^5.$$

Proof. By considering $f_1(x) = f(x-a)$, we may assume that $a = 0$, so that $I = [-h, h]$. Then $S(f,I) = \int_I p$, where p is the cubic interpolating f at $-h$, 0, 0, h. Then

$$q(x) = (x+h)x^2(x-h) = x^2(x^2-h^2).$$

Note that $q(x) \le 0$ on I. By Theorem 4,

$$\tfrac{1}{24} m x^2(h^2-x^2) \le p(x) - f(x) \le \tfrac{1}{24} M x^2(h^2-x^2)$$

for $x \in I$. The statement now follows from the fact that

$$\int_{-h}^{h} x^2(h^2-x^2)\,dx = \tfrac{2}{3}h^5 - \tfrac{2}{5}h^5 = \tfrac{4}{15}h^5. \qquad \square$$

We remark that the literature contains proofs of this result that are decidedly more complicated!

Let us at least mention the estimate derived from (a). For this, we assume that $m' \le f^{(3)}(x) \le M'$ on I, and we have $q(x) = x(x^2-h^2)$, which is positive on $(-h, 0)$ and negative on $(0, h)$. By considering these intervals separately, one finds that

$$\left| S(f,I) - \int_I f \right| \le \tfrac{1}{24}(M'-m')h^4.$$

Example 5. Let $J = \int_0^2 2^x\,dx$. This equals $3/\log 2 \approx 4.3281$. The approximation S given by Simpson's rule is $\frac{1}{3}(1 + 4 \times 2 + 4) = 4\frac{1}{3}$, so $S - J \approx 0.0052$. In the notation of Proposition 6, $M = 4(\log 2)^4$ and $m = (\log 2)^4$, so the upper and lower estimates for

$S - J$ are $\frac{M}{90} \approx 0.0103$ and $\frac{m}{90} \approx 0.0026$. Meanwhile, the estimates derived from (a) are $\pm\frac{1}{8}(\log 2)^3 \approx \pm 0.0416$. (However, there are cases where one of the estimates derived from (a) is actually better; the reader can verify that this occurs for the integral $\int_1^3 (1/x)\, dx$.)

Newton's form of the polynomial and divided differences

Let $x_0, x_1, \ldots, x_n$ be any list of points (possibly with repetitions), and let p be any polynomial of degree n: $p(x) = c_0 + c_1 x + \cdots + c_n x^n$. Then $p(x)$ can be expressed in the form

$$p(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \ldots + a_n(x - x_0)(x - x_1)\ldots(x - x_{n-1}). \quad (3)$$

To demonstrate this, first equate coefficients of x^n to get $a_n = c_n$. Next, equate coefficients of x^{n-1}: we get $c_{n-1} = a_{n-1} - a_n(x_0 + \cdots + x_{n-1})$, which determines a_{n-1}. Continuing in the same way, we see that each a_k exists and is uniquely determined.

For the interpolating polynomial, the expression of the form (3) is called *Newton's form*. For now, we restrict the discussion to the case where the points x_j are distinct. As already mentioned, a_n is the coefficient of x^n, which is given by (2). At the opposite end, the value at x_0 shows that $a_0 = p(x_0) = f(x_0)$. Equating values at x_1, we then have $a_0 + a_1(x_1 - x_0) = p(x_1) = f(x_1)$, which determines a_1. We will now derive an expression for each a_k.

Newton's form has the following highly desirable property:

PROPOSITION 7. *Let the points x_j be distinct. Let $p(x)$ be as in (3), and let $p_k(x)$ be the sum of the first $k+1$ terms (that is, as far as the term containing a_k). Then $p_k(x)$ is the polynomial interpolating f at $x_0, x_1, \ldots, x_k$.*

Proof. The polynomial $p_k(x)$ has degree at most k and, for $0 \le j \le k$, we have $p_k(x_j) = p(x_j) = f(x_j)$, since all the subsequent terms in $p(x)$ have $(x - x_j)$ as a factor. □

It follows that each a_k in (3) is defined by (2), with k replacing n. We restate this more carefully. Our original notation N_j, q_j presupposed a fixed, unstated n. For clarity, we now adopt the following more precise notation. Let distinct points $x_0, x_1, \ldots, x_k$ (with $k \ge 1$) be given. For $0 \le j \le k$, let $N_{k,j} = \{0, 1, \ldots, k\} \setminus \{j\}$ and $q_{k,j}(x) = \prod_{r \in N_{k,j}} (x - x_r)$. Given a function f, the *divided difference* $f[x_0, x_1, \ldots, x_k]$ is defined by

$$f[x_0, x_1, \ldots, x_k] = \sum_{j=0}^{k} \frac{f(x_j)}{q_{k,j}(x_j)}. \quad (4)$$

For a single point, $f[x_0]$ is defined to be $f(x_0)$. What we have established is that when $p(x)$

is expressed as in (3), we have $a_k = f[x_0, x_1, \ldots, x_k]$ for each k.

Of course, Proposition 3 applies: $f[x_0, x_1, \ldots, x_k] = f^{(k)}(\xi)/k!$ for some $\xi \in I$.

With this notation in place, we can give a pleasantly simple expression for the difference between $f(x)$ and $p(x)$:

PROPOSITION 8. *Let $p(x)$ be the polynomial interpolating $f(x)$ at distinct points $x_0, x_1, \ldots, x_n$, and let $q(x) = \prod_{j=0}^{n}(x - x_j)$. Then for x different from all the x_j,*

$$f(x) - p(x) = f[x_0, x_1, \ldots, x_n, x]q(x).$$

Proof. Let x be given and write $x = x_{n+1}$. Let $p_{n+1}(x)$ be the polynomial interpolating f at $x_0, x_1, \ldots, x_n, x_{n+1}$. By Proposition 7,

$$p_{n+1}(x) = p(x) + f[x_0, \ldots, x_n, x_{n+1}]q(x)$$

(for all x). Apply this with $x = x_{n+1}$: since $p_{n+1}(x_{n+1}) = f(x_{n+1})$, we obtain the desired statement. $\square$

This gives a second proof of Theorem 4 (at least for distinct points).

Both from (4), and from the fact that the interpolating polynomial does not depend on the order in which the points are listed, we have:

PROPOSITION 9. *If $(y_0, y_1, \ldots, y_n)$ is a permutation of $(x_0, x_1, \ldots, x_n)$, then*

$$f[y_0, y_1, \ldots y_n] = f[x_0, x_1, \ldots, x_n]. \qquad \square$$

Example 6. Let $f(x) = x^k$, and let $x_0, x_1, \ldots, x_n$ be given, with $n \geq k$. Then the interpolating polynomial is x^k itself. So $f[x_0, x_1, \ldots, x_n]$ equals 0 if $n > k$ and 1 if $n = k$.

We now show how divided differences of order n can be derived from those of order $n - 1$. This is very useful for actual calculation. Numerous different proofs can be found in the literature. We present two of them.

LEMMA 3. *Let points $x_0, x_1, \ldots x_{n-2}, y, z$ be given (where $n \geq 2$). Denote as follows the polynomials interpolating f at the points stated:*

$p_y(x)$: *points* $x_0,\ x_1, \ldots,\ x_{n-2},\ y$;

$p_z(x)$: *points* $x_0,\ x_1, \ldots,\ x_{n-2},\ z$;

$p_{y,z}(x)$: *points* $x_0,\ x_1, \ldots,\ x_{n-2},\ y,\ z$.

Then

$$p_{y,z}(x) = \frac{(x - z)p_y(x) - (x - y)p_z(x)}{y - z}.$$

Proof. The stated polynomial has degree n and agrees with f at each x_j, y and z. $\square$

PROPOSITION 10. *We have*

$$f[x_0, x_1, \ldots, x_{n-1}, x_n] = \frac{f[x_0, \ldots, x_{n-2}, x_n] - f[x_0, \ldots, x_{n-2}, x_{n-1}]}{x_n - x_{n-1}}. \tag{5}$$

Note. Because of the symmetry of divided differences, we can present (5) in various alternative ways, for example

$$f[x_0, x_1, \ldots, x_n] = \frac{f[x_1, x_2, \ldots, x_n] - f[x_0, x_1, \ldots, x_{n-1}]}{x_n - x_0}. \tag{6}$$

Proof 1. In Lemma 3, take $y = x_n$ and $z = x_{n-1}$. Equating the coefficients of x^n, we obtain the stated identity. $\square$

Proof 2. We prove the statement in the form (6). Express $p(x)$ as in (3). But also, taking the points in reverse order, we can write

$$p(x) = b_0 + b_1(x - x_n) + \cdots + b_{n-1}(x - x_n)\ldots(x - x_2) + b_n(x - x_n)(x - x_{n-1})\ldots(x - x_1).$$

We have $a_n = b_n = f[x_0, x_1, \ldots, x_n]$, also

$$a_{n-1} = f[x_0, x_1, \ldots, x_{n-1}], \qquad b_{n-1} = f[x_n, x_{n-1}, \ldots, x_1].$$

But, equating the coefficients of x^{n-1}, we have

$$a_{n-1} - a_n(x_0 + x_1 + \cdots + x_{n-1}) = b_{n-1} - b_n(x_n + x_{n-1} + \cdots + x_1),$$

hence (noting that $b_n = a_n$) $a_n(x_n - x_0) = b_{n-1} - a_{n-1}$. $\square$

Some writers prefer to use (5), together with $f[x_0] = f(x_0)$, as a recursive *definition* of divided differences.

We can use Proposition 10 to calculate divided differences successively. For example,

$$f[x_1, x_2, x_3] = \frac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1}.$$

The calculations can be tabulated in the following way:

$$\begin{array}{llllll}
x_0 & f(x_0) & & & \\
 & & f[x_0,x_1] & & \\
x_1 & f(x_1) & & f[x_0,x_1,x_2] & \\
 & & f[x_1,x_2] & & f[x_0,x_1,x_2,x_3] \\
x_2 & f(x_2) & & f[x_1,x_2,x_3] & \\
 & & f[x_2,x_3] & & \\
x_3 & f(x_3) & & &
\end{array}$$

Example 7. To interpolate $f(x) = 2^x$ at 0, 1, 2, 3. We construct a table as above. The columns labelled f_2, f_3, f_4 are the divided differences over sets of 2, 3 and 4 points respectively.

$$\begin{array}{ccccc}
x_j & f(x_j) & f_2 & f_3 & f_4 \\
 & & & & \\
0 & 1 & & & \\
 & & 1 & & \\
1 & 2 & & \frac{1}{2} & \\
 & & 2 & & \frac{1}{6} \\
2 & 4 & & 1 & \\
 & & 4 & & \\
3 & 8 & & &
\end{array}$$

The successive divided differences involving x_0 are seen on the top sloping line: 1, 1, $\frac{1}{2}$, $\frac{1}{6}$, so the Newton form for the polynomial (with the points in this order) is

$$1 + x + \tfrac{1}{2}x(x-1) + \tfrac{1}{6}x(x-1)(x-2),$$

which one can rewrite (if desired) as $1+\frac{5}{6}x+\frac{1}{6}x^3$. The method of solving for coefficients would have been distinctly more laborious! The first three terms give the polynomial $1 + \frac{1}{2}x + \frac{1}{2}x^2$ interpolating 2^x at the points 0, 1, 2, as previously found in Example 1. Also, ignoring the point 0, we can read off the polynomial interpolating 2^x at 1, 2, 3: $2+2(x-1)+(x-1)(x-2)$.

We finish this section with a pleasant result on substitution of divided differences.

PROPOSITION 11. *Let the points* x_i $(0 \le i \le k)$ *and* y_j $(0 \le j \le r)$ *be distinct, and let* $g(x) = f[y_0, y_1, \ldots, y_r, x]$. *Then*

$$g[x_0, x_1, \ldots, x_k] = f[y_0, \ldots, y_r, x_0, \ldots, x_k].$$

Proof. Induction on k. The case $k = 0$ is the definition of g. Assume the statement true for $k-1$ (i.e. for sets of k points). By Proposition 10 and the induction hypothesis,

$$\begin{aligned}
g[x_0, x_1, \ldots, x_k] &= \frac{g[x_1, x_2, \ldots, x_k] - g[x_0, x_1, \ldots, x_{n-1}]}{x_k - x_0} \\
&= \frac{f[y_0, \ldots, y_r, x_1, x_2, \ldots, x_k] - f[y_0, \ldots, y_r, x_0, x_1, \ldots, x_{k-1}]}{x_k - x_0} \\
&= f[y_0, \ldots, y_r, x_0, \ldots, x_k]. \qquad \square
\end{aligned}$$

Repeated points

We have already defined what we mean by the interpolating polynomial in the case when points are repeated, and shown that it exists (Proposition 2). We now show how the notion of divided differences and the construction of the Newton form can be adapted, without too much trouble, to deal with this case.

First consider the extreme case when all the points coincide: $x_j = x_0$ for $0 \le j \le n$. Suppose that $f^{(n)}$ exists and is continuous at x_0. It is clear from Proposition 3 that if we define

$$f[x_0, x_0, \ldots, x_0] = \frac{f^{(n)}(x_0)}{n!}, \tag{7}$$

then we will have extended the definition of the divided difference in a way that makes it continuous at the point $(x_0, x_0, \ldots, x_0)$. Furthermore, the resulting Newton form

$$f[x_0] + f[x_0, x_0](x - x_0) + \cdots + f[x_0, x_0, \ldots, x_0](x - x_0)^n = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k,$$

is the polynomial interpolating f at $x_0, \ldots, x_0$.

Now suppose that the sequence $x_0, x_1, \ldots, x_n$ comprises k_0 repetitions of y_0 followed by k_1 repetitions of y_1, up to k_r repetitions of y_r. For the moment, it is important to keep the repeated terms together. Having defined expressions of the form $f[y_j, \ldots, y_j]$ by (7), we now complete a difference table as before. In other words, having defined divided differences of length k, those of length $k+1$ are defined by

$$f[x_0, x_1, \ldots, x_k] = \frac{f[x_1, x_2, \ldots, x_k] - f[x_0, x_1, \ldots, x_{k-1}]}{x_k - x_0} \tag{8}$$

whenever the x_j do not all coincide, so that $x_k \neq x_0$. It is clear that divided differences, defined this way, are *continuous* functions of the variables, given continuity of enough derivatives of f. Also, an easy induction shows that $f[x_k, x_{k-1}, \ldots, x_0] = f[x_0, x_1, \ldots, x_k]$.

LEMMA 4. *Proposition 7 still applies when there are repetitions.*

Proof. Let $p(x)$, expressed as in (3), be the polynomial interpolating f at $x_0, x_1, \ldots, x_n$, and let $p_{n-1}(x)$ be formed from $p(x)$ by leaving out the last term $a_n q_n(x)$, where $q_n(x) = (x - x_0)\ldots(x - x_{n-1})$. We show that $p_{n-1}(x)$ interpolates f at $x_0, x_1, \ldots, x_{n-1}$. Supppose that k_j of the terms $x_0, x_1, \ldots, x_n$ equal y_j. Then $f - p$ has a zero of order at least k_j at y_j. If $x_n \neq y_j$, then q_n has a zero of order at least k_j at y_j, and hence $f - p_{n-1}$ does so. If $x_n = y_j$, then q_n, and hence $f - p_{n-1}$, has a zero of order at least $k_j - 1$ at y_j. In both cases, this is what is required. $\square$

PROPOSITION 12. *Let $x_0, x_1, \ldots, x_n$ be any list of points, possibly with repetitions, in an interval I, and suppose that f has continuous nth derivative on I. Then the polynomial interpolating f at these points is given by (3), with $a_k = f[x_0, x_1, \ldots, x_k]$, as just defined.*

Proof. We prove the statement by induction. It is trivial for $n = 0$ (and almost trivial for $n = 1$). Assume that it is correct for $n - 1$, and let $x_0, x_1, \ldots, x_n$ be given. The case where $x_j = x_0$ for all j has been established above, so we suppose that this is not the case. Let $p(x)$, expressed as in (3), be the polynomial interpolating f at $x_0, x_1, \ldots, x_n$. We have to show that $a_n = f[x_0, x_1, \ldots, x_n]$.

We do this by adapting the second proof of Proposition 10. In the notation used there, we have again $a_n = b_n$ and, by Lemma 4 and the induction hypothesis,

$$a_{n-1} = f[x_0, x_1, \ldots, x_{n-1}],$$

$$b_{n-1} = f[x_n, x_{n-1}, \ldots, x_1] = f[x_1, x_2, \ldots, x_n].$$

Exactly as before, we have $a_n(x_n - x_0) = b_{n-1} - a_{n-1}$. By (8), it follows that $a_n = f[x_0, x_1, \ldots, x_n]$. $\square$

Example 8. To interpolate $f(x) = 1/x$ at 1, 2, 2, 2, 3. Then $f[2,2] = f'(2) = -\frac{1}{4}$ and $f[2,2,2] = \frac{1}{2}f''(2) = \frac{1}{8}$. Entering these values in the table, we obtain:

$$\begin{array}{cccccc}
x_j & f(x_j) & f_2 & f_3 & f_4 & f_5 \\
1 & 1 & & & & \\
 & & -\frac{1}{2} & & & \\
2 & \frac{1}{2} & & \frac{1}{4} & & \\
 & & -\frac{1}{4} & & -\frac{1}{8} & \\
2 & \frac{1}{2} & & \frac{1}{8} & & \frac{1}{24} \\
 & & -\frac{1}{4} & & -\frac{1}{24} & \\
2 & \frac{1}{2} & & \frac{1}{12} & & \\
 & & -\frac{1}{6} & & & \\
3 & \frac{1}{3} & & & &
\end{array}$$

So the polynomial is

$$1 - \tfrac{1}{2}(x-1) + \tfrac{1}{4}(x-1)(x-2) - \tfrac{1}{8}(x-1)(x-2)^2 + \tfrac{1}{24}(x-1)(x-2)^3.$$

The reader may care to repeat Example 3 in this style.

For Proposition 12, we did not need full-scale symmetry of divided differences, only reversal of the order. However, a simple continuity argument shows that full-scale symmetry still applies:

PROPOSITION 13. *Proposition 9 still applies when there are repetitions: if* $(y_0, y_1, \ldots, y_n)$ *is a permutation of* $(x_0, x_1, \ldots, x_n)$*, then*

$$f[y_0, y_1, \ldots y_n] = f[x_0, x_1, \ldots, x_n].$$

Proof. Given a point $\boldsymbol{x} = (x_0, x_1, \dots, x_n)$ of $\mathbb{R}^{n+1}$, there are clearly points $\boldsymbol{x}^{(k)}$ of $\mathbb{R}^{n+1}$ that converge to $\boldsymbol{x}$ as $k \to \infty$ and have all components distinct. Let $\boldsymbol{y}^{(k)}$ be formed from $\boldsymbol{x}^{(k)}$ by the corresponding permutation. By Proposition 9, we have (with obvious notation) $f[\boldsymbol{y}^{(k)} = f[\boldsymbol{x}^{(k)}]$ for each k. Since the divided difference is a continuous function on $\mathbb{R}^{n+1}$, we have $f[\boldsymbol{y}] = \lim_{k\to\infty} f[\boldsymbol{y}^{(k)}]$ and similarly for $\boldsymbol{x}$, so $f[\boldsymbol{y}] = f[\boldsymbol{x}]$. □

In the same way, Propositions 8 and 11 can be freed from the requirement that the points are distinct.

The integral expression for divided differences

(The reader is free to defer this section, or leave it out.) There is an explicit expression for divided differences in the form of a repeated integral. This is of interest in theory (especially in the case of repeated points), but distinctly less useful for actual computation than the process described above.

PROPOSITION 14. *Suppose that x_j $(0 \le j \le n)$ are points of an interval I (possibly repeated) , and that f has continuous nth derivative on I. Then $f[x_0, x_1, \dots, x_n]$ equals*

$$\int_0^1 dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} f^{(n)}[x_0 + (x_1 - x_0)t_1 + \cdots + (x_n - x_{n-1})t_n]\, dt_n. \tag{9}$$

Proof. We prove the statement for the case when the points are distinct. The case where there are repeated points then follows by continuity as in Proposition 13, since it is clear that the integral defines a continuous function of the variables x_j. First, we prove the case $n = 1$. The stated integral is then

$$I_1 = \int_0^1 f'[x_0 + (x_1 - x_0)t_1]\, dt_1.$$

The substitution $x_0 + (x_1 - x_0)t_1 = u$ gives

$$I_1 = \frac{1}{x_1 - x_0} \int_{x_0}^{x_1} f'(u)\, du = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = f[x_0, x_1].$$

Assume the statement correct for sets of n points. Take $x_0, x_1, \dots, x_n$ and let I_n be the integral stated. For the integration with respect to t_n, substitute

$$x_0 + (x_1 - x_0)t_1 + \cdots + (x_{n-1} - x_{n-2})t_{n-1} + (x_n - x_{n-1})t_n = u.$$

The limits of integration for u are

$$u_0 = x_0 + (x_1 - x_0)t_1 + \cdots + (x_{n-1} - x_{n-2})t_{n-1},$$

$$u_1 = x_0 + (x_1 - x_0)t_1 + \cdots + (x_n - x_{n-2})t_{n-1}.$$

The transformed integral is

$$\frac{1}{x_n - x_{n-1}} \int_{u_0}^{u_1} f^{(n)}(u)\, du = \frac{1}{x_n - x_{n-1}} [f^{(n-1)}(u_1) - f^{(n-1)}(u_0)].$$

By the induction hypothesis and Proposition 10, we now have

$$I_n = \frac{f[x_0, \ldots, x_{n-2}, x_n] - f[x_0, \ldots, x_{n-2}, x_{n-1}]}{x_n - x_{n-1}} = f[x_0, x_1, \ldots, x_{n-1}, x_n]. \qquad \square$$

This gives an alternative proof of Proposition 3, since

$$\int_0^1 dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n = \frac{1}{n!}.$$

There is also an explicit expression for divided differences (with repeated points) in terms of partial derivatives. We state it without proof: *Let the list $x_0, x_1, \ldots, x_n$ comprise $k_j + 1$ repetitions of y_j for $0 \le j \le r$. Then*

$$f[x_0, x_1, \ldots, x_n] = \frac{1}{k_0! \ldots k_r!} \frac{\partial^{k_0}}{\partial y_0^{k_0}} \frac{\partial^{k_r}}{\partial y_r^{k_r}} f[y_0, y_1, \ldots y_r].$$

Equally spaced points: forward differences

Suppose that the points x_j are equally spaced, so that (for some $h > 0$), $x_j = x_0 + jh$ for each j. The divided differences can then be expressed in terms of the *forward difference* operator Δ, defined as follows (for a chosen h):

$$(\Delta f)(x) = f(x+h) - f(x),$$

and $\Delta^n f = \Delta(\Delta^{n-1} f)$, so that

$$(\Delta^n f)(x) = (\Delta^{n-1} f)(x+h) - (\Delta^{n-1} f)(x).$$

Hence, for example, $(\Delta^2 f)(x) = f(x+2h) - 2f(x+h) + f(x)$. Clearly, $(\Delta f)(x_0) = hf[x_0, x_1]$, where $x_1 = x_0 + h$.

PROPOSITION 15. *Let $x_j = x + jh$ $(0 \le j \le n)$. Then*

$$(\Delta^n f)(x) = n! h^n f[x_0, x_1, \ldots, x_n].$$

Proof. By induction. The case $n = 1$ is immediate, as above. Assume the statement holds for $n - 1$. Then, by Proposition 10,

$$\begin{aligned} (\Delta^n f)(x) &= (\Delta^{n-1} f)(x_1) - (\Delta^{n-1} f)(x_0) \\ &= (n-1)! h^{n-1} (f[x_1, x_2, \ldots, x_n] - f[x_0, x_1, \ldots, x_{n-1}] \\ &= (n-1)! h^{n-1} n h f[x_0, x_1, \ldots, x_n] \\ &= n! h^n f[x_0, x_1, \ldots, x_n]. \end{aligned}$$

PROPOSITION 16. $(\Delta^n f)(x) = \sum_{j=0}^{n}(-1)^j \binom{n}{j} f(x_j)$, *where* $x_j = x + jh$.

Proof. By Proposition 15,

$$(\Delta^n f)(x) = n!h^n \sum_{j=0}^{n} \frac{f(x_j)}{q_j(x_j)}.$$

Now for fixed j, $\prod_{i<j}(x_j - x_i) = j!h^j$ and $\prod_{k>j}(x_j - x_k) = (-1)^{n-j}(n-j)!h^{n-j}$, so

$$q_j(x_j) = h^n(-1)^{n-j}j!(n-j)!.$$

The stated equality follows. □

Propositions 16 and 3 give at once:

PROPOSITION 17. *If f is n times differentiable on $[x, x + nh]$, then there exists ξ in $(x, x + nh)$ such that $(\Delta^n f)(x) = h^n f^{(n)}(\xi)$.* □

Alternative direct proof. By induction. The case $n = 1$ is the mean-value theorem. Assume the statement true for a certain n. Then $\Delta^{n+1}f = \Delta^n g$, where $g = \Delta f$, so $g(x) = f(x+h)-f(x)$. By the induction hypothesis, there exists η in $[x, x+nh]$ such that $(\Delta^n g)(x) = g^{(n)}(\eta) = f^{(n)}(\eta + h) - f^{(n)}(\eta)$. By the mean-value theorem again, this equals $f^{(n+1)}(\xi)$ for some ξ in $(\eta, \eta + h)$. □

One can give a direct proof of Proposition 16 in similar fashion.

The integral expression for forward differences is pleasantly simple: each integration is now on the fixed interval $[0, h]$:

PROPOSITION 18. *Suppose that f is n times differentiable on $[x, x + nh]$. Then*

$$(\Delta^n f)(x) = \int_0^h dt_1 \int_0^h dt_2 \ldots \int_0^h f^{(n)}(x + t_1 + \cdots + t_n)\, dt_n.$$

Proof. The case $n = 1$ is correct, since it says

$$\int_0^h f'(x + t_1)\, dt_1 = f(x + h) - f(x).$$

Assuming the statement correct for n, we have

$$(\Delta^{n+1} f)(x) = (\Delta^n f)(x + h) - (\Delta^n f)(x) = \int_0^h dt_1 \int_0^h dt_2 \ldots \int_0^h G(t_n)\, dt_n,$$

where

$$\begin{aligned} G(t_n) &= f^{(n)}(x + h + t_1 + \cdots + t_n) - f^{(n)}(x + t_1 + \cdots + t_n) \\ &= \int_0^h f^{(n+1)}(x + t_1 + \cdots + t_n + t_{n+1})\, dt_{n+1}. \end{aligned}$$

Substituting this, we obtain the required formula for the case $n+1$. $\square$

Uniform approximation and Chebyshev polynomials

This section of our notes has a slightly more advanced flavour, but most of it should still be accessible to readers with a basic grounding in Real Analysis (at one point, we use a standard result from Complex Analysis). Our starting point is the expression for the "error" $f(x)-p(x)$ in Theorem 4: $[1/(n+1)!]f^{(n+1)}(\xi)q(x)$, where $q(x)=\prod_{j=0}^{n}(x-x_j)$. In Example 2, we saw that when $f(x)$ is x^{n+1}, the error is exactly $q(x)$.

Now $q(x)$ depends on the points x_j. How can we choose these points so as to make $q(x)$ (and hence the error estimation) as "small" as possible (in some sense)? Different measures of "smallness" are possible, but we will adopt the simple-minded one of best uniform approximation, measured by the maximum deviation from 0 on the given interval $[a,b]$. In other words, the problem is to minimize $\|q\|_\infty$, where

$$\|q\|_\infty = \sup\{|q(x)| : a \le x \le b\}$$

(this is standard notation). For the moment, we take $[a,b]$ to be $[-1,1]$ and replace $n+1$ by n.

One might expect the objective to be achieved by taking equally spaced points, but this is not the case! Furthermore, it is rather remarkable that the desired $q(x)$ can be identified explicitly: a beautiful piece of reasoning shows that it is the *Chebyshev polynomial* $T_n(x)$. These polynomials are defined, for each n, by the identity

$$\cos nt = 2^{n-1}T_n(\cos t).$$

For example, since $\cos 3t = 4\cos^3 t - 3\cos t$, we have $T_3(x)=x^3-\frac{3}{4}x$. For present purposes, we do not need to know anything about these polynomials except that they exist, and that T_n is monic with degree n (easily proved by induction, using the identity $\cos(n+1)t + \cos(n-1)t = 2\cos nt\cos t$).

Every x in $[-1,1]$ is expressible as $\cos t$, and $|\cos nt|\le 1$ for all t, so $\|T_n\|_\infty = 2^{-(n-1)}$. (Of course, this says nothing about the behaviour of T_n outside $[-1,1]$.) Also, since $\cos k\pi = (-1)^k$ and $\cos(k+\frac{1}{2})\pi = 0$ for integers k, we have:

$$\text{if } y_k = \cos\frac{k\pi}{n}, \text{ then } T_n(y_k) = (-1)^k\, 2^{-(n-1)},$$

$$\text{if } x_j = \cos\frac{(j+\frac{1}{2})\pi}{n}, \text{ then } T_n(x_j)=0, \text{ hence } T_n(x)=(x-x_0)(x-x_1)\dots(x-x_{n-1}).$$

THEOREM 19. *Let* $\|q\|_\infty = \sup\{|q(x)| : -1 \le x \le 1\}$. *Among all monic polynomials* q *of degree* n, $\|q\|_\infty$ *is least when* $q = T_n$, *and then* $\|q\|_\infty = 2^{-(n-1)}$. *So if* $q(x)$ *is expressed as* $\prod_{j=0}^{n-1}(x - x_j)$, *then* $\|q\|_\infty$ *is least when* $x_j = \cos[(j + \frac{1}{2})\pi/n]$.

Proof. Write $2^{-(n-1)} = \alpha$, so that $T_n(y_k) = (-1)^k\alpha$ for $0 \le k \le n$. Note that the points y_k are in decreasing order, with $y_0 = 1$, $y_n = -1$. Suppose that $|q(y_k)| < \alpha$ for each k. Then $T_n(y_k) - q(y_k)$ is strictly positive for *even* k (since then $T_n(y_k) = \alpha$) and strictly negative for *odd* k. By the intermediate value theorem, it follows that $T_n - q$ has a zero in each open interval (y_{k+1}, y_k), hence at least n zeros in total. But this is impossible, since $T_n - q$ is a polynomial of degree at most $n - 1$ (the x^n term cancels). So in fact $|q(y_k)| \ge \alpha$ for some k, hence $\|q\| \ge \alpha = \|T_n\|_\infty$. $\square$

To transfer this result to a general interval $[a, b]$, perform the substitution $x = \frac{1}{2}(b - a)t + \frac{1}{2}(a + b)$: when x goes from a to b, t goes from -1 to 1. If x_j is the point corresponding to $t_j = \cos[(j + \frac{1}{2})\pi/n]$, then $x - x_j = \frac{1}{2}(b - a)(t - t_j)$, so

$$\prod_{j=0}^{n-1}(x - x_j) = \frac{1}{2^n}(b - a)^n T_j(t) :$$

denote this polynomial by $\tilde{T}_n(x)$: we call it the "transferred Chebyshev polynomial", and the points x_j the "Chebyshev points". The conclusion is:

COROLLARY. *Let* $\|q\|_\infty = \sup\{|q(x)| : a \le x \le b\}$. *Among all monic polynomials* q *of degree* n, $\|q\|_\infty$ *is least when* $q = \tilde{T}_n$, *and then* $\|q\|_\infty = (b - a)^n/2^{2n-1}$.

Inserted into Theorem 4 (still with $n + 1$ replaced by n), this gives at once:

THEOREM 20. *Suppose that* $|f^{(n)}(x)| \le M_n$ *on* $[a, b]$, *and let* p_{n-1} *be the polynomial interpolating* f *at the Chebyshev points* x_j $(0 \le j \le n - 1)$. *Let* $\|f\|_\infty = \sup\{|f(x)| : a \le x \le b\}$. *Then*

$$\|f - p_{n-1}\|_\infty \le \frac{(b - a)^n M_n}{2^{2n-1} n!}. \qquad \square$$

As the reader may know, one says that $p_n \to f$ as $n \to \infty$ *uniformly* on $[a, b]$ if $\|f - p_n\|_\infty \to 0$ as $n \to \infty$. By Theorem 20, this will occur, for any $[a, b]$, if, for some K, we have $M_n \le K^n$ for all n. However, bounds of this form for $f^{(n)}(x)$ do not commonly apply!

Example 9. Consider $f(x) = 1/x$ on the interval $[1, b]$. Then $f^{(n)}(x) = (-1)^{n-1} n!/x^{n+1}$, so $M_n = n!$ and for the p_n defined in Theorem 20, we have $\|f - p_{n-1}\|_\infty \le (b - 1)^n/2^{2n-1} = 2[(b - 1)/4]^n$. This tends to 0, implying uniform convergence, if $b - 1 < 4$.

In general, Real Analysis does not provide a pleasant estimation of bounds for higher derivatives. However, Complex Analysis does, in the form of the following standard result:

LEMMA 5. *Suppose that f is analytic on $\{z : |z - z_0| < R\}$. Suppose that $r < R$ and $|f(z)| \le M$ for $|z - z_0| = r$. Then $|f^{(n)}(z_0)| \le Mn!/r^n$ for all $n \ge 1$.* □

Armed with this, we can formulate a convergence theorem based on the nature of f as a complex function. Given a real interval $I = [a, b]$, let $E_r(I)$ be the set of points in the complex plane at distance no more than r from some point of I. This is comprised of the rectangle $\{x + iy : a \le x \le b,\ |y| \le r\}$ together with semicircles centred at a and b.

THEOREM 21. *Suppose that f is analytic on a complex region containing $E_r(I)$, where $I = [a, b]$ and $b - a < 4r$. Let p_n be defined as in Theorem 20. Then $p_n \to f$ uniformly on I as $n \to \infty$.*

Proof. Then $|f(z)|$ is bounded, say by M, on $E_r(I)$. By Theorem 20 and Lemma 5,

$$\|f - p_{n-1}\|_\infty \le \frac{M(b-a)^n}{2^{2n-1}r^n} = 2M\left(\frac{b-a}{4r}\right)^n,$$

which tends to 0 as $n \to \infty$ if $b - a < 4r$. □

This is not the strongest possible theorem of this type, but it is what follows naturally from this approach.

For comparison, we now give a brief account of the corresponding results for equally spaced points x_j, starting with an estimation of $|q(x)|$. To cater for the fact that q is zero at each x_j, define the function r^* by: $r^*(t) = t(1-t)$ for $0 \le t \le 1$ and $r^*(t+k) = r^*(t)$ for integers k. Since $t(1-t) = \frac{1}{4} - (t - \frac{1}{2})^2$, we have $0 \le r^*(t) \le \frac{1}{4}$ for all t.

PROPOSITION 22. *Let $x_j = x_0 + jh$ ($0 \le j \le n$), and let $q_n(x) = \prod_{j=0}^n (x - x_j)$. If $x_k \le x \le x_{k+1}$, then*

$$|q_n(x)| \le h^{n+1}(k+1)!(n-k)!r^*(t),$$

where $x = x_0 + th$. If q_n is defined this way with $[x_0, x_n] = [a, b]$, then

$$\|q_n\|_\infty \le \frac{n!}{4n^{n+1}}(b-a)^{n+1}.$$

Proof. The substitution $x = x_0 + th$ gives $q_n(x) = h^{n+1}\pi_n(t)$, where $\pi_n(t) = t(t-1)\dots(t-n)$. Also, $x_k \le x \le x_{k+1}$ translates to $k \le t \le k+1$. Clearly,

$$\prod_{j=0}^{k-1}(t-j) \le \prod_{j=0}^{k-1}(k+1-j) = (k+1)!,$$

$$\prod_{j=k+2}^{n}(j-t) \le \prod_{j=k+2}^{n}(j-k) = (n-k)!.$$

Hence $\pi_n(t) \le (k+1)!(n-k)!r^*(t)$.

Now $(k+1)!(n-k)! \le n!$ for each k (with equality when k is 0 or $n-1$) and $0 \le r^*(t) \le \frac{1}{4}$. Substituting $h = (b-a)/n$, we obtain the stated bound for $\|q_n\|_\infty$. $\square$

Clearly, if k is close to $n/2$, then $(k+1)!(n-k)!$ is much less than $n!$, so much smaller bounds for $q_n(x)$ apply near the middle of the interval.

We assume Stirling's theorem, which states that $n! \sim cn^{n+\frac{1}{2}}e^{-n}$ as $n \to \infty$, where $c = (2\pi)^{1/2}$. With this substitution, the bound for $\|q_n\|_\infty$ in Proposition 22 becomes $c'(b-a)^{n+1}/[n^{1/2}e^n]$, where $c' = \frac{1}{4}c$. The corresponding bound for the Chebyshev polynomial, after replacing n by $n+1$, was $\frac{1}{2}(b-a)^{n+1}/4^n$ (Corollary to Theorem 19). Essentially, it is better by having the dividing factor 4^n instead of e^n.

Inserting Proposition 22 into Theorem 4, we obtain:

PROPOSITION 23. *Suppose that $|f^{(n+1)}(x)| \le M_{n+1}$ on $[a,b]$, and let p_n be the polynomial interpolating f at $n+1$ equally spaced points x_j in $[a,b]$. Then*

$$\|f - p_n\|_\infty \le \frac{(b-a)^{n+1}M_{n+1}}{4(n+1)n^{n+1}}. \qquad \square$$

Applying Lemma 5, we obtain the analogue of Theorem 21:

PROPOSITION 24. *Suppose that f is analytic on a complex region containing $E_r(I)$, where $I = [a,b]$ and $b-a < er$. Let p_n be the polynomial interpolating f at $n+1$ equally spaced points x_j in $[a,b]$. Then $p_n \to f$ uniformly on I as $n \to \infty$.*

Proof. By Proposition 23 and Lemma 5, $\|f-p_n\|_\infty$ is bounded by

$$\begin{aligned}
\frac{(b-a)^{n+1}}{4(n+1)n^{n+1}}\,\frac{M(n+1)!}{r^{n+1}} &= \frac{M(b-a)^{n+1}n!}{(nr)^{n+1}} \\
&\sim \frac{M(b-a)^{n+1}}{(nr)^{n+1}}\,cn^{n+\frac{1}{2}}e^{-n} \quad \text{by Stirling's theorem} \\
&= \frac{c'}{n^{1/2}}\left(\frac{b-a}{er}\right)^{n+1}
\end{aligned}$$

for another constant c'. $\square$

We leave it as an exercise for the interested reader to show that, under these conditions, convergence will occcur at the mid-point of the interval if $b-a < 2er$.

It was shown by Runge in 1901 that for the function $f(x) = 1/(1+x^2)$, the sequence of interpolating polynomials for equally spaced points does not converge pointwise to $f(x)$. We finish with a proof of this fact. It involves some fairly detailed estimations, but the following account offers at least a modest degree of simplification compared with some.

Recall Proposition 8: $f(x) - p(x) = f[x_0, x_1, \ldots, x_n, x]q(x)$.

LEMMA 6. *Let $f(x) = 1/(1+x^2)$ and let x_j $(-n \le j \le n)$ be distinct points with $x_{-j} = -x_j$ for each j. Then*

$$f[x_0, x_{-1}, x_1, \ldots, x_{-n}, x_n, x] = (-1)^{n+1} x f(x) \prod_{j=1}^{n} \frac{1}{1+x_j^2}.$$

Proof. We show first that

$$f[x_{-1}, x_1, \ldots, x_{-n}, x_n, x] = (-1)^n f(x) \prod_{j=1}^{n} \frac{1}{1+x_j^2}. \tag{10}$$

First we consider the case $n = 1$. Since $x_{-1} = -x_1$ and $f(x_{-1}) = f(x_1)$, we have, by (2):

$$\begin{aligned} f[x_{-1}, x_1, x] &= \frac{f(x)}{x^2 - x_1^2} + \frac{f(x_1)}{2x_1(x_1 - x)} + \frac{f(x_1)}{2x_1(x_1 + x)} \\ &= \frac{f(x) - f(x_1)}{x^2 - x_1^2} \\ &= \frac{1}{x^2 - x_1^2}\left(\frac{1}{1+x^2} - \frac{1}{1+x_1^2}\right) \\ &= -\frac{1}{(1+x^2)(1+x_1^2)}. \end{aligned}$$

Assume now that (10) holds for a certain n, and denote the LHS by $g(x)$. By Proposition 11 and the case $n = 1$, we then have

$$\begin{aligned} f[x_{-1}, x_1, \ldots, x_{-n}, x_n, x_{-(n+1)}, x_{n+1}, x] &= g[x_{-(n+1)}, x_{n+1}, x] \\ &= (-1)^n \prod_{j=1}^{n} \frac{1}{1+x_j^2} f[x_{-(n+1)}, x_{n+1}, x] \\ &= (-1)^{n+1} \prod_{j=1}^{n+1} \frac{1}{1+x_j^2} f(x). \end{aligned}$$

This establishes (10), by induction. Also, since $x_0 = 0$ and $f(x_0) = 1$, we have

$$f[x_0, x] = \frac{f(x) - 1}{x} = -\frac{x}{1+x^2} = -xf(x).$$

By Proposition 10 again,

$$\begin{aligned} f[x_0, x_{-1}, x_1, \ldots, x_{-n}, x_n, x] &= g[x_0, x] \\ &= (-1)^n \prod_{j=1}^{n} \frac{1}{1+x_j^2} f[x_0, x] \\ &= (-1)^{n+1} x f(x) \prod_{j=1}^{n} \frac{1}{1+x_j^2}. \qquad \square \end{aligned}$$

PROPOSITION 25. *Let $f(x) = 1/(1+x^2)$. Let p_n be the polynomial interpolating f at $n+1$ equally spaced points through $[-a, a]$. If $a > 12\frac{1}{2}$, then the sequence $[p_n(\frac{1}{2}a)]$ does not converge.*

Proof. Let $h = a/n$ and $x_j = jh$ for $-n \le j \le n$. By Proposition 8 and Lemma 6,

$$|f(x) - p_{2n}(x)| = |x| f(x) |q_n(x)| \prod_{j=1}^{n} \frac{1}{1+j^2h^2},$$

where $q_n(x) = \prod_{j=-n}^{n}(x - jh)$. Write $A_n = |q_n(\frac{1}{2}a)|$ and $B_n = \prod_{j=1}^{n}(1 + j^2h^2)$. Our statement will follow if we can show that $A_n/B_n \to \infty$ when n tends to infinity through odd values. To do this, we will estimate $\log A_n$ from below and $\log B_n$ from above by comparison with the corresponding integrals.

Let $n = 2r - 1$, so that $\frac{1}{2}a = (r - \frac{1}{2})h$. Then

$$\begin{aligned} \log A_n &= \sum_{j=-2r+1}^{2r-1} \log|r - \tfrac{1}{2} - j| + (2n+1)\log h \\ &= \sum_{k=1}^{r} \log(k - \tfrac{1}{2}) + \sum_{k=1}^{3r-1} \log(k - \tfrac{1}{2}) + (4r-1)\log h. \end{aligned}$$

Now $\log x$ is a concave function (the second derivative is negative) and for any concave function g, one has $\int_{k-1}^{k} g \le g(k - \frac{1}{2})$ (this is geometrically obvious, and easy to prove formally from the mean-value theorem). So

$$\sum_{k=1}^{r} \log(k - \tfrac{1}{2}) \ge \int_0^r \log x \, dx = x \log x - x,$$

(remark: this is the the main element of the proof of Stirling's formula). Hence also

$$\begin{aligned} \sum_{k=1}^{3r-1} \log(k - \tfrac{1}{2}) &\ge \sum_{k=1}^{3r} \log(k - \tfrac{1}{2}) - \log 3r \\ &\ge (3r-1)\log 3r - 3r \\ &\ge 3r(\log r + \log 3 - 1) - \log 3r. \end{aligned}$$

Further,

$$(4r-1)\log h = (4r-1)[\log a - \log(2r-1)] > (4r-1)\log a - 4r(\log r + \log 2).$$

Together, these inequalities give

$$\log A_{2r-1} > 4r(\log a - 1) + (3\log 3 - 4\log 2)r - \log r - c_1 \tag{11}$$

for a certain constant c_1.

We turn to the estimaton of $\log B_n$. It equals $\sum_{j=1}^{n} g(jh)$, where $g(x) = \log(1+x^2)$. Since g is an increasing function, standard integral comparison gives $h\sum_{j=1}^{n-1} g(jh) \leq I$, where $I = \int_0^a g(x)\,dx$. Integrating by parts, with 1 as one factor, we find

$$\begin{aligned} I &= a\log(1+a^2) - 2\int_0^a \left(1 - \frac{1}{1+x^2}\right)\,dx \\ &< a\log(1+a^2) - 2a + \pi. \end{aligned}$$

Hence

$$\log B_n \leq \frac{I}{h} + g(a) = \frac{nI}{a} + g(a) < (n+1)\log(1+a^2) - 2n + \frac{\pi n}{a}.$$

Now

$$\log(1+a^2) - 2\log a = \int_{a^2}^{1+a^2} \frac{1}{t}\,dt < \frac{1}{a^2},$$

hence (with $n = 2r-1$) we have

$$\begin{aligned} \log B_n &< 2n(\log a - 1) + \left(\frac{\pi}{a} + \frac{1}{a^2}\right)n + c_2 \\ &< 4r(\log a - 1) + 2\left(\frac{\pi}{a} + \frac{1}{a^2}\right)r + c_2 \end{aligned} \tag{12}$$

for a certain constant c_2. By (11) and (12), it is clear that we will have $\log A_{2r-1} - \log B_{2r-1} \to \infty$ as $r \to \infty$ provided that

$$2\left(\frac{\pi}{a} + \frac{1}{a^2}\right) < 3\log 3 - 4\log 2 \approx 0.5232,$$

which is true for $a \geq 12\frac{1}{2}$. □