Numerical Integration   Numerical Analysis, Fall 2020

Daksh Aggarwal   December 24, 2020

It is well-known that integration is an essential tool for mathematics and its applications. Numerous physical laws require some kind of integration to state and apply. Integration in the complex plane plays an important role in number theory. In topology and geometry, we are interested in defining a notion of integration in spaces different from Euclidean space. Integration even enters the discrete area of combinatorics through the powerful probabilistic method. However, many naturally occurring functions have no simple closed-form integral. In other cases, we might only have the values the function takes at a finite set of points. When faced with integrating such functions, we have to rely on numerical methods. Our goal in this short article is to explore one of the most common ways of numerically integrating a function: approximate the function by a polynomial and then integrate the resulting polynomial. One of the ways of approximating a function by a polynomial is known as *interpolation*. There are different ways of applying interpolation to numerically integrate a function, each with its strengths and weaknesses. Here we will primarily be focusing on Chebyshev interpolation and its application to numerical integration.

Let us begin by reviewing interpolation and how it might be used to integrate a function. Suppose we have a function $f : [a, b] \to \mathbb{R}$ that we know to pass through the points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. We want to approximate f by a polynomial p. In what sense do we wish to "approximate" f using p? We require that $p(x_i) = f(x_i) = y_i$ for each $i = 1, \ldots, n$; moreover, we want p to have the lowest possible degree. A simple way of creating such a polynomial is given by *Lagrange interpolation*. For $1 \leq i \leq n$, consider the polynomial

$$L_i(x) = \prod_{1 \leq j \neq i \leq n} \frac{x - x_j}{x_i - x_j}. \tag{1}$$

Note that $L_i(x_i) = 1$ and $L_i(x_j) = 0$ for $j \neq i$. Thus, if we let

$$p(x) = \sum_{i=1}^{n} y_i L_i(x), \tag{2}$$

we see that $p(x_i) = y_i$ for each i. Further, note that since each $L_i(x)$ has degree exactly $n - 1$, the polynomial $p(x)$ has degree at most $n - 1$, which means that $p(x)$ is the *unique* minimal degree polynomial that interpolates the points $\{(x_i, y_i)\}_{1\le i\le n}$. Indeed, suppose $q(x)$ is any polynomial such that $q(x_i) = y_i$ with degree at most $n-1$. Consider the polynomial $r(x) = p(x) - q(x)$; since both $p(x)$ and $q(x)$ have degree at most $n-1$, their difference $r(x)$ has degree at most $n-1$ too. Further, for each $i = 1, \ldots, n$, $r(x_i) = p(x_i) - q(x_i) = y_i - y_i = 0$, and so $r(x)$ has n roots. Since r has degree at most $n-1$, $r(x)$ must be identically zero and so $q(x) = p(x)$.

Now, suppose we want to integrate f on an interval $[a, b] \subseteq \mathbb{R}$, where we know that f passes through points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, with all $x_i \in [a, b]$. Then, taking p to be the Lagrange interpolation of f described in (2), we can approximate

$$\int_a^b f(x)\,dx \approx \int_a^b p(x)\,dx = \int_a^b \left(\sum_{i=1}^n y_i L_i(x)\right) dx = \sum_{i=1}^n y_i \cdot \int_a^b L_i(x)\,dx,$$

thus allowing us to leverage the fact that any polynomial can be easily integrated. A natural question arises: how good is this approximation? Clearly, the answer will depend on how good an approximation the polynomial p is for the function f. We now prove a result about the error $f(\alpha) - p(\alpha)$ for $\alpha \in [a, b]$. We will require the following generalization of Rolle's theorem.

Theorem 1 (Generalized Rolle's theorem)**.** *Let $f \in C^{n-1}([a, b])$ such that $f^{(n)}$ exists on (a, b). Suppose that $f(x_1) = f(x_2) = \cdots = f(x_{n+1})$ for some $x_i \in [a, b], i = 1, \ldots, n+1$. Then there exists $c \in (a, b)$ such that $f^{(n)}(c) = 0$.*

Theorem 2 (Interpolation error)**.** *Let $f \in C^{n-1}([a, b])$ such that $f^{(n)}$ exists on (a, b). Suppose $f(x_i) = y_i$ for $i = 1, \ldots, n$ and let $p(x)$ be the Lagrange interpolating polynomial of the points $\{(x_i, y_i)\}_{1\le i\le n}$. Let $\alpha \in [a, b]$. Then there exists $c \in (a, b)$ such that*

$$f(\alpha) - p(\alpha) = \frac{f^n(c)}{n!} \prod_{i=1}^n (\alpha - x_i).$$

Proof. Let $L_i(x)$ for $i = 1, \ldots, n$ be defined as in (1), so that

$$p(x) = \sum_{i=1}^n y_i L_i(x).$$

Let $q(x)$ Lagrange interpolation of the $n+1$ points $\{(x_i, y_i)\}_{1\le i\le n} \cup (\alpha, f(\alpha))$. Note that we can write $q(x)$ as

$$q(x) = \sum_{i=1}^{n} y_i \frac{L_i(x)(x-\alpha)}{x_i - \alpha} + f(\alpha) \prod_{i=1}^{n} \frac{x - x_i}{\alpha - x_i}.$$

Define the function $r(x) = f(x) - q(x)$. By the construction of $q(x)$, we know that

$$r(x_i) = f(x_i) - q(x_i) = 0, \text{ for } i = 1, \ldots, n,$$

and $r(\alpha) = f(\alpha) - q(\alpha) = 0$. Thus, the function r has $n + 1$ zeroes in $[a, b]$. By the generalized Rolle's theorem, we infer that there exists a $c \in (a, b)$ such that $r^{(n)}(c) = 0$. Now,

$$\begin{aligned} r^{(n)}(x) &= f^{(n)}(x) - q^{(n)}(x) \\ &= f^{(n)}(x) - \sum_{i=1}^{n} y_i \cdot \frac{n!}{(x_i - \alpha) \prod_{1 \le j \ne i \le n} (x_i - x_j)} - f(\alpha) \cdot \frac{n!}{\prod_{i=1}^{n} (\alpha - x_i)}, \end{aligned}$$

where have used the fact that each $L_i(x)(x - \alpha)$ is an n-degree polynomial, and so differentiating it n times will yield $n!$ times the coefficient of x^n in $L_i(x)(x - \alpha)$ (and similarly for the last term). Since $r^{(n)}(c) = 0$, solving for $f(\alpha)$ we get

$$\begin{aligned} f(\alpha) &= \frac{f^{(n)}(c)}{n!} \prod_{i=1}^{n} (\alpha - x_i) - \sum_{i=1}^{n} y_i \frac{\prod_{j=1}^{n} (\alpha - x_j)}{(x_i - \alpha) \prod_{1 \le j \ne i \le n} (x_i - x_j)} \\ &= \frac{f^{(n)}(c)}{n!} \prod_{i=1}^{n} (\alpha - x_i) + \sum_{i=1}^{n} y_i \prod_{1 \le j \ne i \le n} \frac{\alpha - x_j}{x_i - x_j} \qquad \text{[Cancelling } (\alpha - x_i)\text{]} \\ &= \frac{f^{(n)}(c)}{n!} \prod_{i=1}^{n} (\alpha - x_i) + \sum_{i=1}^{n} y_i L_i(\alpha) \\ &= \frac{f^{(n)}(c)}{n!} \prod_{i=1}^{n} (\alpha - x_i) + p(\alpha), \end{aligned}$$

from which we obtain the desired result

$$f(\alpha) - p(\alpha) = \frac{f^{(n)}(c)}{n!} \prod_{i=1}^{n} (\alpha - x_i).$$

□

Notice that in the error $f(\alpha) - p(\alpha)$, the quantity that we can potentially control is the product

$\prod_{i=1}^{n}(\alpha - x_i)$. Therefore, since the error in the integral approximation directly depends on the interpolation error, we should choose the points $x_1, \ldots, x_n \in [a, b]$ so that the quantity $\prod_{i=1}^{n}(x - x_i)$, for $x \in [a, b]$, is minimized. This leads us to the idea of Chebyshev nodes, which are the roots of Chebyshev polynomials, and ensure the optimality we seek.

The family of *Chebyshev polynomials* is defined by

$$T_n(x) = \cos(n \arccos(x)), \text{ for } x \in [-1, 1],$$

for each $n \in \mathbb{N}$. Yes, they *are* polynomials! Indeed, we can inductively see this is the case. The first two, $T_0(x) = \cos(0) = 1$ and $T_1(x) = \cos(\arccos(x)) = x$, are certainly polynomials. Now, assume that $T_i(x)$ are polynomials for all $i \leq n$ with some $n \in \mathbb{N}_{n \geq 1}$. For each $x \in [-1, 1]$, let $y_x = \arccos(x)$. Then,

$$T_{n+1}(x) = \cos(ny_x + y_x) = \cos(ny_x)\cos(y_x) - \sin(ny_x)\sin(y_x),$$

and

$$T_{n-1}(x) = \cos(ny_x - y_x) = \cos(ny_x)\cos(y_x) + \sin(ny_x)\sin(y_x).$$

So, adding the two relations,

$$\begin{aligned} T_{n+1}(x) + T_{n-1}(x) &= 2\cos(ny_x)\cos(y_x) \\ &= 2\cos(n \arccos(x)) \cdot x \\ &= 2xT_n(x). \end{aligned}$$

Thus,

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \tag{3}$$

and so, by the inductive hypothesis, we conclude that $T_{n+1}(x)$ is a polynomial too. From the recurrence relation (3), we also see that T_n has degree n since T_1 and T_0 have degrees 1 and 0 respectively. We collect a few other simple observations about $T_n(x)$:

1. Since $T_n(x)$ is equal to the cosine of some real number (namely, ny_x), we have $|T_n(x)| \leq 1$.

2. From the relation (3), we deduce that the coefficient of x^n in $T_n(x)$ is 2^{n-1}.

3. Let $\beta_i = \cos\left(\frac{(2i-1)\pi}{2n}\right)$ for $i = 1, \ldots, n$. Then, $T_n(\beta_i) = 0$ for $i = 1, \ldots, n$. Combining with the previous observation, we conclude that

$$T_n(x) = 2^{n-1} \prod_{i=1}^{n} (x - \beta_i)\,.$$

These roots β_i of $T_n(x)$ are known as *(nth degree) Chebyshev nodes.*

4. $T_n\Big(\cos\left(\frac{i\pi}{n}\right)\Big) = (-1)^n$ for $i = 0, \ldots, n$. In other words, $T(x)$ alternates $n+1$ times between 1 and -1. Combining with observation 1, we also have that $\max\limits_{x\in[-1,1]} |T_n(x)| = 1$.

We are now ready to prove the optimality of Chebyshev nodes in the sense that

$$\max_{x\in[-1,1]} \prod_{i=1}^{n} (x - x_i)$$

is minimized when we take x_i to be Chebyshev nodes.

Theorem 3. *Let $n \in \mathbb{N}^+$. Then,*

$$\min_{\{x_i\}\in[-1,1]^n} \max_{x\in[-1,1]} \left| \prod_{i=1}^{n} (x - x_i) \right| = \frac{1}{2^{n-1}},$$

where the optimal value is achieved when $\{x_i\}$ are the n roots of $T_n(x)$.

Proof. Let β_i be the n roots of $T_n(x)$. By observations 3 and 4 above, we have

$$\max_{x\in[-1,1]} \left| \prod_{i=1}^{n} (x - \beta_i) \right| = \max_{x\in[-1,1]} \left| \frac{1}{2^{n-1}} T_n(x) \right| = \frac{1}{2^{n-1}}.$$

Suppose, for a contradiction, that $Q(x) = \prod_{i=1}^{n} (x - \gamma_i)$ is such that

$$\max_{x\in[-1,1]} |Q(x)| < \frac{1}{2^{n-1}}.$$

Then, consider the polynomial $R(x) = \frac{1}{2^{n-1}} T_n(x) - Q(x)$. By observation 4, as x ranges over $[-1, 1]$, $\frac{1}{2^{n-1}} T_n(x)$ alternates $n+1$ times between $\frac{1}{2^{n-1}}$ and $-\frac{1}{2^{n-1}}$. As a result, $R(x)$ changes sign at least n times in $[-1, 1]$. Being the difference of two polynomials, $R(x)$ is continuous, and so by the Intermediate Value theorem, $R(x)$ has n roots in $[-1,1]$. However, since $R(x)$ has degree at most $n-1$ (being the difference of two *monic* degree n polynomials), $R(x)$ must be identically zero. Thus, $Q(x) = \frac{1}{2^{n-1}}T_n(x)$, which is a contradiction since $\left|\frac{1}{2^{n-1}}T_n(x)\right|$ achieves a maximum value of $\frac{1}{2^{n-1}}$ while we assumed $|Q(x)|$ was strictly less than that on $[-1,1]$. $\square$

Let us consider an example. Suppose that

$$f(x) = \frac{x \sin x \cos x}{4\cos^2 x + \sin^2 x}, \tag{4}$$

and we want to interpolate $f(x)$ on $[0, \frac{\pi}{2}]$ using three points. One possibility for the nodes would be the evenly spaced $x_1 = 0, x_2 = \frac{\pi}{4}$, and $x_3 = \frac{\pi}{2}$, with corresponding y values $y_1 = 0, y_2 = \frac{\pi}{20}$, and $y_3 = 0$. The resulting interpolation would be

$$\begin{aligned} p_1(x) &= 0 \cdot L_1(x) + \frac{\pi}{20} \cdot \frac{x(x-\frac{\pi}{2})}{\frac{\pi}{4}(\frac{\pi}{4}-\frac{\pi}{2})} + 0 \cdot L_3(x) \\ &= \frac{2x}{5} - \frac{4}{5\pi}x^2. \end{aligned}$$

Next, let us use Chebyshev nodes to interpolate f; this is known as *Chebyshev interpolation*. One immediate problem is that Chebyshev nodes are defined on the interval $[-1,1]$ and so we have to map them from $[-1,1]$ to $[0, \frac{\pi}{2}]$ using a linear transformation to preserve their optimality. As a fun application of interpolation, the desired linear mapping of $[-1,1]$ onto some interval $[a,b]$ is given by the interpolating polynomial of the points $(-1,a)$ and $(1,b)$, which is

$$T(x) = \frac{b-a}{2}x + \frac{a+b}{2}. \tag{5}$$

In our example, we have $T(x) = \frac{\pi}{4}x + \frac{\pi}{4}$, and so we use the nodes $\frac{\pi}{4}\beta_i + \frac{\pi}{4}$ for $i = 1,2,3$, where β_i are the roots of the Chebyshev polynomial $T_3(x)$. The resulting polynomial is

$$p_2(x) = -0.176x^2 + 0.384x - 0.036.$$

The two interpolating polynomials $p_1(x)$ and $p_2(x)$ along with $f(x)$ are plotted in Figure 1. As expected, at most points $p_2(x)$ shows relatively less error than $p_1(x)$ in approximating $f(x)$.

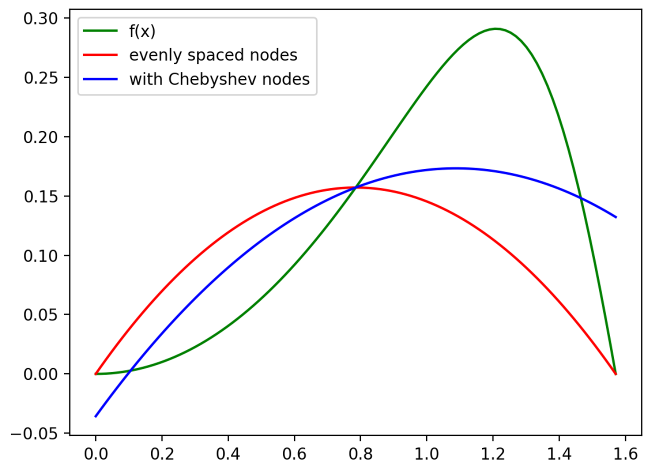

Figure 1: $f(x), p_1(x)$, and $p_2(x)$

Let us now understand a little more carefully the error when we interpolate using Chebyshev nodes. Let $T(\beta_1), \ldots, T(\beta_n) \in [a, b]$ be the transformed nth degree Chebyshev nodes that we are using the interpolate f on $[a, b]$. By Theorem 2, for $\alpha \in (a, b)$, there exists $c \in [a, b]$ such that

$$f(\alpha) - p(\alpha) = \frac{f^{(n)}(c)}{n!} \prod_{i=1}^{n} (\alpha - T(\beta_i)).$$

Since T is clearly invertible, we can fix $\hat{\alpha} \in [-1, 1]$ such that $T(\hat{\alpha}) = \alpha$. Then,

$$\begin{aligned} |f(\alpha) - p(\alpha)| &= \left| \frac{f^{(n)}(c)}{n!} \prod_{i=1}^{n} (T(\hat{\alpha}) - T(\beta_i)) \right| \\ &= \frac{|f^{(n)}(c)|}{n!} \prod_{i=1}^{n} \left| \frac{b-a}{2}\hat{\alpha} - \frac{b-a}{2}\beta_i \right| \\ &= \frac{|f^{(n)}(c)|}{n!} \left(\frac{b-a}{2}\right)^n \left| \prod_{i=1}^{n} (\hat{\alpha} - \beta_i) \right| \\ &\leq \frac{|f^{(n)}(c)|}{n!} \left(\frac{b-a}{2}\right)^n \frac{1}{2^{n-1}}. && \text{[By Theorem 3]} \end{aligned}$$

We record this result as

Lemma 4 (Chebyshev interpolation error)**.** *Let $f \in C^n([a, b])$ and let $\{\beta_i\}_{1 \leq i \leq n}$ be the nth degree Chebyshev nodes. If p is the interpolating polynomial of the points $\{(\beta_i, f(\beta_i))\}_{1 \leq i \leq n}$, then for each*

$\alpha \in [a,b]$, there exists $c \in (a,b)$ such that

$$|f(\alpha) - p(\alpha)| \le \frac{\left|f^{(n)}(c)\right|}{2^{n-1} n!} \left(\frac{b-a}{2}\right)^n.$$

All this study was motivated with the goal of finding a good polynomial to approximate the integral of f. The next result shows the error we can expect when we use Chebyshev interpolation for numerical integration.

Theorem 5. *Let $f \in C^n([a,b])$ and let p be the interpolating polynomial on n Chebyshev nodes that interpolates f on $[a,b]$. Then there exists $c \in (a,b)$ such that*

$$\left|\int_a^b f(x)\,dx - \int_a^b p(x)\,dx\right| \le \frac{\left|f^{(n)}(c)\right|}{2^{2n-1} n!}(b-a)^{n+1}.$$

Proof. For each $x \in (a,b)$, by Lemma 4, there exists $c_x \in (a,b)$ such that

$$|f(x) - p(x)| \le \frac{\left|f^{(n)}(c_x)\right|}{2^{n-1} n!} \left(\frac{b-a}{2}\right)^n.$$

This, along with general properties of the definite integral, imply

$$\begin{aligned}
\left|\int_a^b f(x)\,dx - \int_a^b p(x)\,dx\right| &\le \int_a^b |f(x) - p(x)|\,dx \\
&\le \int_a^b \frac{\left|f^{(n)}(c_x)\right|}{2^{n-1} n!} \left(\frac{b-a}{2}\right)^n dx \\
&= \frac{1}{2^{n-1} n!} \left(\frac{b-a}{2}\right)^n \int_a^b \left|f^{(n)}(c_x)\right|\,dx.
\end{aligned}$$

Since $f^{(n)}$ is continuous and c_x depends continuously on $x \in [a,b]$, by the Mean Value theorem for integrals, there exists $c \in (a,b)$ such that

$$\int_a^b \left|f^{(n)}(c_x)\right|\,dx = \left|f^{(n)}(c)\right|(b-a),$$

and so

$$\left|\int_a^b f(x)\,dx - \int_a^b p(x)\,dx\right| \le \frac{1}{2^{n-1} n!} \left(\frac{b-a}{2}\right)^n \left|f^{(n)}(c)\right|(b-a).$$

□

Let us use Chebyshev interpolation to approximate the integral of our example $f(x) = \frac{x \sin x \cos x}{4\cos^2 x + \sin^2 x}$ on the interval $[0, \frac{\pi}{2}]$. Using the three-node Chebyshev interpolation $p_2(x)$ we constructed, we obtain $\int_0^{\pi/2} f(x)\,dx \approx 0.1897671487$. How do we find or bound the error in this approximation? The specific form of $f(x)$ is not amenable to easy differentiation and estimation. Moreover, $\int_0^{\pi/2} f(x)\,dx$ does not appear to have an evaluation using any methods utilizing simple substitutions or techniques[1]. Fortunately, the definite integral can indeed be evaluated using complex integration and the Cauchy Residue theorem in a fairly standard way, though we omit the calculation here. We obtain

$$\int_0^{\pi/2} \frac{x \sin x \cos x}{4\cos^2 x + \sin^2 x}\,dx = \frac{\pi}{6} \ln \frac{3}{2} \approx 0.2123010342.$$

So, Chebyshev interpolation has given us a quadratic whose integral approximates the integral of the not-very-tame $f(x)$ with an absolute error of 0.02253 – not too bad! When we instead use the interpolation $p_1(x)$ constructed from evenly spaced nodes, we get an error of 0.04781, more than twice as much error. In Table 1 we have tabulated approximations using 4 to 16 Chebyshev nodes along with the corresponding error; by 15 nodes we have an approximation that is correct up to 8 decimal places.

# of nodes	Chebyshev approximation	Absolute error
4	0.2098535681	0.0024474660
5	0.2134277497	0.0011267155
6	0.2125554443	0.0002544101
7	0.2123315763	0.0000305421
8	0.2123139090	0.0000128748
9	0.2122963678	0.0000046663
10	0.2122987410	0.0000022932
11	0.2123004356	0.0000005985
12	0.2123008319	0.0000002023
13	0.2123011051	0.0000000709
14	0.2123010739	0.0000000398
15	0.2123010295	0.0000000046

Table 1: Approximating $\int_0^{\pi/2} \frac{x \sin x \cos x}{4\cos^2 x + \sin^2 x}\,dx$ using Chebyshev interpolation

While numerical integration using Chebyshev nodes certainly works wonderfully, we should emphasize that it is not the best technique. A major factor that limits the performance of Chebyshev interpolation when used for integration is that it attempts to construct a single polynomial approx-

[1] At least the author thinks this is the case. The author would be interested in knowing if the reader finds a simple way.

imation for f on the whole interval $[a, b]$, having the least possible error in the *worst* case. So, a Chebyshev interpolation might have a high *average* error, thus degrading the integral approximation. One potential way to ameliorate this problem might be to partition the interval $[a, b]$ into smaller subintervals and then sum the results of approximating the integral of f using Chebyshev interpolation on those subintervals. We will not pursue this particular idea but rather briefly discuss Gaussian quadrature, to illustrate that the optimality of Chebyshev nodes for interpolation does not translate to optimality in numerical integration.

Like Chebyshev interpolation, Gaussian quadrature takes advantage of a special family of polynomials known as Legendre polynomials. For each $n \in \mathbb{N}$, we define the *n-th Legendre polynomial* as

$$G_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}[(x^2 - 1)^n], \text{ for } x \in \mathbb{R}.$$

Note that the degree of G_n is precisely n. Another primary property of Legendre polynomials is that they are orthogonal on $[-1, 1]$. In other words,

$$\int_{-1}^{1} G_m(x) G_n(x)\, dx = \begin{cases} 0 & m \neq n \\ \text{nonzero} & m = n. \end{cases}$$

As a result the first $n+1$ Legendre polynomials $\{G_0, G_1, \ldots, G_n\}$ form a basis for the vector space $\mathcal{P}_n$ of polynomials with degree at most n and coefficients in $\mathbb{R}$. Indeed, the linear independence is observed by the fact that if

$$\sum_{j=0}^{n} c_j G_j(x) = 0, \text{ for some } c_j \in \mathbb{R},$$

then, for each $i = 0, \ldots, n$,

$$0 = \int_{-1}^{1} G_i(x) \left(\sum_{j=0}^{n} c_j G_j(x) \right) dx = \sum_{j=0}^{n} c_j \int_{-1}^{1} G_i(x) G_j(x)\, dx = c_i \int_{-1}^{1} G_i^2(x)\, dx,$$

with the last equality following from orthogonality. Since the last integral is nonzero, we conclude that $c_i = 0$ for $i = 0, \ldots, n$. The fact that $\{G_0, G_1, \ldots, G_n\}$ spans $\mathcal{P}_n$ follows from a straightforward induction, which we omit here.

Thus, if we have any polynomial $p(x)$ with degree at most $n - 1$, then we can find a set of real coefficients $b_0, \ldots, b_{n-1}$ such that

$$p(x) = \sum_{i=0}^{n-1} b_i G_i(x),$$

and so, again by orthogonality,

$$\int_{-1}^{1} G_n(x)p(x)\,dx = \sum_{i=0}^{n-1} b_i \int_{-1}^{1} G_n(x)G_i(x)\,dx = 0. \tag{6}$$

We can leverage this observation to integrate any polynomial with degree at most $2n-1$ without any error using the nth Legendre polynomial as follows. Let $p \in \mathcal{P}_{2n-1}$. By the Division algorithm for polynomials, we know that there exists a polynomial $r \in \mathcal{P}_{n-1}$ and a polynomial q (possibly zero) such that

$$p(x) = G_n(x)q(x) + r(x).$$

Moreover, since p has degree at most $2n-1$ and G_n has degree precisely n, q must have degree at most $n-1$. Thus, by our observation in (6),

$$\int_{-1}^{1} p(x)\,dx = \int_{-1}^{1} G_n(x)q(x)\,dx + \int_{-1}^{1} r(x)\,dx = 0 + \int_{-1}^{1} r(x)\,dx.$$

One problem remains: how do we determine r? Fortunately, we do not have to actually carry out the actual division of p by G_n. Instead, we make use of yet another useful property of $G_n(x)$: it can be shown that it has n real roots in the interval $(-1, 1)$. Thus, if we denote these n roots by $l_1, \ldots, l_n$, then for each i,

$$p(l_i) = G_n(l_i)q(l_i) + r(l_i) = r(l_i).$$

In other words, the n points $(l_1, p(l_1)), \ldots, (l_n, p(l_n))$ all lie on the at most $n-1$ degree polynomial r. So, by the uniqueness of the Lagrange interpolating polynomial, r can be recovered as

$$r(x) = \sum_{i=1}^{n} p(l_i) \prod_{1 \leq j \neq i \leq n} \frac{x - l_j}{l_i - l_j} = \sum_{i=1}^{n} p(l_i) L_i(x).$$

Therefore, the *Gaussian quadrature* of the polynomial $p \in \mathcal{P}_{2n-1}$ on the interval $[-1, 1]$ is

$$\int_{-1}^{1} p(x)\,dx = \sum_{i=1}^{n} p(l_i) \int_{-1}^{1} L_i(x)\,dx,$$

which can be extended to approximate the integral of any function f as

$$\int_{-1}^{1} f(x)\,dx = \sum_{i=1}^{n} f(l_i) \int_{-1}^{1} L_i(x)\,dx.$$

We can use this to integrate f over any interval $[a,b]$ by applying our mapping T (5) from $[-1,1]$ to $[a,b]$:

$$\begin{aligned}\int_{a}^{b} f(x)\,dx &= \int_{-1}^{1} f(T(x)) \cdot \frac{dT(x)}{dx}\,dx \\ &= \int_{-1}^{1} f(T(x)) \cdot \frac{b-a}{2}\,dx \\ &= \left(\frac{b-a}{2}\right) \sum_{i=1}^{n} f(T(l_i)) \int_{-1}^{1} L_i(x)\,dx.\end{aligned}$$

We had earlier hinted that Gaussian quadrature is, in some sense, optimal for numerical integration. Notice that by using interpolation on just the n roots of G_n, Gaussian quadrature can integrate any polynomial $\mathcal{P}_{2n-1}$ with zero error. Formally, this property is equivalent to saying that Gaussian quadrature has *degree of precision* $2n-1$. Now, note that any integration technique using interpolation on n points will essentially be of the form

$$\int_{a}^{b} f(x)\,dx \approx \sum_{i=1}^{n} c_i \int_{a}^{b} L_i(x)\,dx,$$

where $c_1, \ldots, c_n \in \mathbb{R}$ and the L_i's form the interpolation through some chosen nodes $x_1, \ldots, x_n$ as in (1). Therefore, in choosing these unknowns, we have $2n$ degrees of freedom, which is the dimension of the vector space $\mathcal{P}_{2n-1}$. In other words, we heuristically expect that an integration technique using interpolation on n nodes can have a degree of precision at most $2n-1$, a bound which is met by Gaussian quadrature.

On the other hand, since $p^{(n)}(x)$ is zero for only $\deg(p) \leq n-1$, by Theorem 5, we know that integration using Chebyshev interpolation has degree of precision $n-1$. Therefore, in this respect, Gaussian quadrature is far more superior than integration using Chebyshev interpolation. We have tabulated the approximations given by Gaussian quadrature for our example in Table 2. Indeed, Gaussian quadrature obtains the correct answer up to 8 decimal places with only 12 nodes and up to 10 with 14 nodes.

# of nodes	Gaussian quadrature	Absolute error
4	0.2111538158	0.0011472184
5	0.2123844274	0.0000833932
6	0.2123361880	0.0000351538
7	0.2122985628	0.0000024713
8	0.2122999640	0.0000010702
9	0.2123011066	0.0000000725
10	0.2123010666	0.0000000325
11	0.2123010320	0.0000000021
12	0.2123010332	0.0000000010
13	0.2123010342	0.0000000001
14	0.2123010342	0.00000000003

Table 2: Approximating $\int_0^{\pi/2} \frac{x \sin x \cos x}{4\cos^2 x + \sin^2 x}\, dx$ using Gaussian quadrature

Thus, here we conclude our exploration of Chebyshev interpolation and its application to integration. Although, the optimality of Chebyshev nodes for interpolation does not yield an optimal integration method, we saw that the underlying theory of Chebyshev polynomials is interesting in itself. We also briefly discussed Gaussian quadrature as an example of an optimal method from the perspective of integration, but there is still much we have left unsaid here. Moreover, there many other numerical integration techniques, which have equally rich underlying theories, that the reader might like to explore next