1. Gauss-Jacobi quadrature and Legendre polynomials

Simpson's rule for evaluating an integral $\int_a^b f(t)dt$ gives the correct answer with error of about $O(n^{-4})$ (with a constant that depends on f, in particular, it depends on the size of the fourth derivative of f). However, we can do much better than Simpson's rule for evaluating integrals.

Suppose w is defined and integrable on $[a, b]$ and we want to evaluate

$$\int_a^b p(t)w(t)dt,$$

for $p \in \mathcal{P}_n$ (the polynomials of degree n). Think of $w = 1$ or $w(t) = (t-a)^\alpha$ as the main examples. If we are given any $n+1$ distinct points $\{x_k\}_0^n \subset [a, b]$ then p is determined by its values at these points, i.e., the map

$$p \to \{p(x_0), \dots p(x_n)\}$$

is an invertible map $\mathcal{P}_n \to \mathbb{R}^n$. Thus there must be real numbers w_k so that

$$\int_a^b p(t)w(t)dt = \sum_{k=0}^n w_k p(x_k), \tag{1}$$

holds for all $p \in \mathcal{P}_n$.

What are these weights more explicitly? Given the point set $\{x_k\}_0^n$ define the Lagrange polynomials

$$L_k(x) = \prod_{0 \le j \le n, j \ne k} \frac{x - x_j}{x_k - x_j}.$$

This is equal to 1 at x_k and equal to 0 at the other x_j's. We must have

$$p(x) = \sum_{k=0}^n p(x_k) L_k(x),$$

for any $p \in \mathcal{P}_n$, since both sides are degree n polynomials that agree at $n+1$ points. Thus

$$\int_a^b p(x)w(x)dx = \int_a^b \sum_{k=0}^n p(x_k)L_k(x)w(x)dx = \sum_{k=0}^n p(x_k)[\int_a^b L_k(x)w(x)dx].$$

Thus (1) holds with $w_k = \int_a^b L_k(x)w(x)dx$.

We can simplify this further by noting that if

$$p_n(x) = (x - x_1)\cdots(x - x_n) = \prod_{k=1}^n (x - x_k),$$

then

$$L_k(x) = \prod_{0 \le j \le n, j \ne k} \frac{x - x_j}{x_k - x_j} = \frac{p_n(x)}{(x - x_k)p_n'(x_k)},$$

since both sides are degree n polynomials that are 1 at x_k and 0 at x_j, $j \ne k$ (this derivation uses l'Hopital's rule). Thus

$$w_k = \int_a^b \frac{p_n(x)}{(x - x_k)p_n'(x_k)} w(x)dx. \tag{2}$$

When $w(t) = 1$ is the constant function, then these weights are integrals of polynomials and can be computed exactly.

The discussion so far assumes that we are given the points $\{x_k\}$. If we are allowed to choose these points, then we have $n+1$ additional degrees of freedom, so we might hope to correctly evaluate integrals for even higher degree polynomials. In fact, we can choose $n+1$ points $\{x_k\}_0^n$ so that (1) holds for all polynomials of degree $\le 2n+1$.

The secret is to choose a polynomial p of degree $n+1$ which is orthogonal to every polynomial q of lesser degree, i.e., so that

$$\langle p, q \rangle_w = \int_a^b p(t)q(t)w(t)dt = 0,$$

for all $q \in \mathcal{P}_n$. Now let $\{x_k\}$ be the zeros of p and let $\{w_k\}$ be the weights which make (1) true for polynomials of degree $\le n$. If f is a polynomial of degree $\le 2n+1$, then long division of polynomials shows that we can write $f = a + bp$ where a, b are polynomials of degree $\le n$. Thus

$$\begin{aligned}
\int_a^b f(t)w(t)dt &= \int_a^b a(t)w(t)dt + \int_a^b b(t)p(t)w(t)dt \\
&= \sum_k w_k a(x_k) + 0 \\
&= \sum_k w_k f(x_k),
\end{aligned}$$

where the last line holds since $f = a$ on the zeros of p.

To see that it is not possible to increase the degree of f to $2n+2$, consider the function $\prod_{k=0}^n (t - x_k)^2$. It vanishes at the points $\{x_k\}$ so $\sum_k w_k f(x_k) = 0$, but $\int_a^b f(t)w(t)dt > 0$, at least if $w > 0$ since $f > 0$ except at $n+1$ points.

LEMMA 1. *Given a sequence of orthonormal polynomials* $\{p_k\}_{k=1}^n$ *we have*

$$w_k = \frac{k_n}{k_{n-1}} \frac{\langle p_{n-1}, p_{n-1}\rangle}{p_{n-1}(x_k)p_n'(x_k)},$$

where $\{x_k\}_1^n$ *are the zeros of* p_n *and* k_n *is the leading coefficient of* p_n *(i.e., the coefficient of* x^n *in* p_n*).*

To prove this we need two preliminary results. The first is:

LEMMA 2. *Let*

$$K_n(x,y) = \sum_{k=0}^n p_k(x)p_k(y).$$

Suppose $K(x,y)$ *is a polynomial of degree* n *in both* x *and* y*. Then*

$$\langle p(x), K(x,y)\rangle_{w(x)} = p(y),$$

holds for every polynomial p *of degree* n *iff* $K = K_n$*.*

PROOF. If p is polynomial of degree $\leq n$ then it has a n expansion in terms of the basis $p(x) = \sum a_m p_m(x)$, so

$$\begin{aligned}
\langle p(x), K_n(x)\rangle_w &= \langle \sum a_m p_m(x), \sum p_k(x)p_k(y)\rangle_w \\
&= \sum_{m,k} a_m p_k(y)\langle p_m(x), p_k(x)\rangle_w \\
&= \sum_k a_k p_k(y) \\
&= p(y),
\end{aligned}$$

so the equality holds when $K = K_n$. Conversely, some equality holds for K and all p. Fix w and choose $p(x) = K_n(x,w)$. Then

$$\langle K_n(x,w), K(x,y)\rangle_w = K_n(y,w).$$

But by our earlier calculation k_n has the reproducing property so

$$\langle K(x,w), K_n(x,y)\rangle_w = K(y,w).$$

Since the two left hand sides equal the same integral, we deduce $K(y,w) = K_n(y,w)$ for any y, w, which proves the lemma. □

The second preliminary result we need is:

THEOREM 3 (Christoffel-Darboux). *With notation as above,*

$$K_n(x,y) = \frac{k_n}{k_{n+1}} \frac{p_{n+1}(x)p_n(y) - p_n(x)p_{n+1}(y)}{x-y}$$

PROOF. Let $K(x,y)$ denote the right hand side above. The numerator is a polynomial in x of degree $\leq n+1$ and vanishes when $x = y$, so $K(x,y)$ is actually a polynomial in x of degree $\leq n$. Similarly for y. Thus to show $K = K_n$ we only have to show it has the reproducing property of the previous lemma.

A bit of expanding and using $\langle p_n, p_{n+1}\rangle_w = 0$ shows

$$\begin{aligned} \langle p(x)m, K(x,y)\rangle_w &= \frac{k_n}{k_{n+1}} \langle (p_{n+1}(x)p_n(y) - p_n(x)p_{n+1}(y)), \frac{p(x)-p(y)}{x-y}\rangle_w \\ &+ \frac{k_n}{k_{n+1}} p(y) \langle p_{n+1}(x), \frac{p_n(y)-p_n(x)}{x-y}\rangle_w \\ &+ \frac{k_n}{k_{n+1}} p(y) \langle p_n(x), \frac{p_{n+1}(x)-p_{n+1}(y)}{x-y}\rangle_w \end{aligned}$$

Note that $(p(x)-p(y))/(x-y)$ has degree $\leq n-1$ as a polynomial in x and hence is orthogonal to p_n. Thus the first inner product is 0. Similarly for the second inner product. To compute the third inner product, write

$$\begin{aligned} \frac{k_n}{k_{n+1}} p(y) \frac{p_{n+1}(x)-p_{n+1}(y)}{x-y} &= k_n[\frac{y^{n+1}-x^{n+1}}{y-x} + \ldots] \\ &= k_n x^n + \ldots \\ &= p_n(x) + q(x,y), \end{aligned}$$

where q is a polynomial of degree $\leq n-1$ and hence orthogonal to p_n. Thus the third inner product equals $\langle p_n, p_n\rangle_w = k_{n+1}/k_n$, and hence

$$\langle p(x)m, K(x,y)\rangle_w = p(y).$$

By the previous lemma this implies $K = K_n$, as desired. $\square$

LEMMA 1. We already know from (2) that

$$w_k = \frac{1}{p_n'(x_k)} \int_a^b \frac{p_n(x)}{x-x_k} w(x)dx$$

Since $p_n(x_k) = 0$,

$$\begin{aligned}\int_a^b \frac{p_n(x)}{x - x_k} w(x)dx &= \frac{1}{p_{n-1}(x_k)} \int \frac{p_n(x)p_{n-1}(x_k)}{x - x_k} w(x)dx \\ &= \frac{1}{p_{n-1}(x_k)} \int \frac{p_n(x)p_{n-1}(x_k) - p_{n-1}(x)p_n(x_k)}{x - x_k} w(x)dx \\ &= \frac{1}{p_{n-1}(x_k)} \frac{k_n}{k_{n-1}} \int K_n(x, x_k) w(x)dx \\ &= \frac{k_n}{k_{n-1} p_{n-1}(x_k)}.\end{aligned}$$

□

For more general functions, the difference between our discrete estimate and the actual integral can be bounded as follows:

$$E_n(f) = \int_a^b f(t)w(t)dt - \sum_k w_k f(x_k) = \frac{f^{(2n)}(\zeta)}{(2n)! k_n^2},$$

where ζ is some point in (a, b) and k_n is the coefficient of the power t^n in $p(t)$. For Schwarz-Christoffel integrals, the most relevant case is when w is a Jacobi weight

$$w(t) = (1 - x)^\alpha (1 + x)^\beta,$$

when this estimate is known to be (see []),

$$E_n(f) = f^{(2n)}(\zeta) \frac{2^{2n+\alpha+\beta+1}\Gamma(n + \alpha + 1)\Gamma(n + \beta + 1)\Gamma(n + \alpha + \beta + 1)n!}{\Gamma(2n + \alpha + \beta + 1)\Gamma(2n + \alpha + \beta + 2)(2n)!}.$$

If $\alpha = \beta = 0$ then $w = 1$ and p is a Legendre polynomial. Then the error bound simplifies to

$$E_n(f) = f^{(2n)}(\zeta) \frac{2^{2n+1}(n!)^4}{(2n + 1)((2n)!)^3}.$$

Consider a simple case like $f(t) = e^t$. Then all the derivatives of f are bounded on $[-1, 1]$ and using Stirling's formula

$$n! \sim n^n e^{-n} \sqrt{2\pi n},$$

we see that

$$E_n(e^t) = O(n^{-2n}).$$

On the other hand, the nth order Taylor series for e^t only approximates it to within $1/n! \gg n^{-n}$ on $[-1, 1]$. Thus the numerical integration using n points should give about twice as many correct digits as term-by-term integration of the nth order power series. Of course, we can just double the number of terms in the power series to obtain the same accuracy. Even for fairly small n, both error estimates will be less than machine precision.

So very efficient numerical integration is possible if we can

(1) find $p_{n+1} \in \mathcal{P}_{n+1}$ so that $p_{n+1} \perp \mathcal{P}_n$ and $\|p_{n+1}\|_w = \langle p_{n+1}, p_{n+1}\rangle_w = 1$,
(2) find the zeros $\{x_k\}$ of p_{n+1}
(3) find the weights $\{w_k\}$.

The first step is the main difficulty. Once we have the polynomial p, we can use Newton's method to find the roots of p_{n+1} and the weights are given by

$$w_k = -\frac{k_{n+1}}{k_n}\frac{1}{p_{n+1}(x_k)p_n'(x_k)}.$$

Suppose $\{p_k\}_0^n$ are orthonormal polynomials of degree k and the coefficient of x^k in p_k is c_k. We can find a polynomial (orthogonal to $\mathcal{P}_n$, but not necessarily of unit norm) p_{n+1} by taking any $(n+1)$st degree polynomial p and subtracting always its orthogonal projection onto each of the 1-dimensional subspaces corresponding to these vectors, i.e.,

$$p_{n+1}(x) = p(x) - \sum_{k=0}^{n} p_k(x)\langle p, p_k\rangle_w.$$

Since we get to choose p, we take $p = xp_n$, so that

$$\begin{aligned} p_{n+1}(x) &= xp_n(x) - \sum_{k=0}^{n} p_k(x)\langle xp_n, p_k\rangle_w \\ &= xp_n(x) - p_n(x)\langle xp_n, p_n\rangle_w - p_{n-1}(x)\langle xp_n, p_{n-1}\rangle_w - \sum_{k=0}^{n-2} p_k(x)\langle p_n, xp_k\rangle_w \\ &= xp_n(x) - p_n(x)\langle xp_n, p_n\rangle_w - p_{n-1}(x)\langle xp_n, p_{n-1}\rangle_w \\ &= p_n(x)(x - \langle xp_n, p_n\rangle_w) - p_{n-1}(x)\langle xp_n, p_{n-1}\rangle_w \\ &= p_n(x)(x - a_n) - p_{n-1}(x)b_n \end{aligned}$$

We have used the facts that $\langle xf, g\rangle_w = \langle f, xg\rangle_w$ and that p_n is perpendicular to xp_k if $k < n-1$. The polynomial constructed is not necessarily of unit norm, but we can fix this by replacing p_{n+1} by

$$\frac{p_{n+1}}{\|p_{n+1}\|_w}.$$

To implement the method we have to be able to compute the recursion coefficients

$$\begin{aligned} a_n &= \langle xp_n, p_n\rangle_w \\ b_n &= \langle xp_n, p_{n-1}\rangle_w \\ c_n &= \|p_{n+1}\|_w = \|p_n(x-a_n) - p_{n-1}b_n\|_w. \end{aligned}$$

Recall that each of these inner products is an integral of the form

$$\int_a^b f(t)w(t)dt.$$

We already know p_n (by induction) so we could find its roots and use these to exactly evaluate such integrals for polynomials of degree $\leq 2n-1$. However, the inner products above involve polynomials of degree up to $2n+1$, and using the roots of p_n will definitely give a wrong answer for $\int_a^b tp_n^2(t)w(t)dt$. Therefore these coefficients should be computed by other means.

Here we will only consider the case of Jacobi weights with a singularity at one endpoint (or possibly neither endpoint), i.e., weights of the form $w(x) = (x-a)^\alpha$ on the interval $[a, b]$. The most important case is when $\alpha = 0$ and $w(x) = 1$ is constant. We can compute an integral of the form

$$\int_a^b (\sum_{k=0}^n a_k x^k)(x-a)^\alpha dx,$$

using the following observation. A polynomial $p(x) = \sum_{k=0}^n a_k x^k$ has a Taylor expansion around any point, including the point a. This Taylor expansion, must also be a polynomial of degree n. Thus we can write

$$\sum_{k=0}^n b_k(x-a)^k = p(x) = \sum_{k=0}^n a_k x^k.$$

Then

$$\begin{aligned} \int_a^b (\sum_{k=0}^n a_k x^k)(x-a)^\alpha dx &= \int_a^b \sum_{k=0}^n b_k(x-a)^k(x-a)^\alpha \\ &= \sum_{k=0}^n b_k \int_a^b (x-a)^{k+\alpha} dx \\ &= \sum_{k=0}^n b_k \frac{(b-a)^{k+\alpha+1}}{k+\alpha+1}. \end{aligned}$$

So now we have to compute the $\{b_k\}$ from the $\{a_k\}$. Note that

$$\sum_{k=0}^{n} a_k x^k = \sum_{k=0}^{n} a_k (x-a+a)^k = \sum_{k=0}^{n} a_k [\sum_{j=0}^{k} (x-a)^j a^{k-j} \binom{k}{j}],$$

so we get $\mathbf{b} = (b_0, \ldots, b_n)$ we just have to apply the matrix

$$M = (m_{jk}) = a^{k-j} \cdot \binom{k}{j},$$

to the vector $\mathbf{a} = (a_0, \ldots, a_n)$.

In the special case $\alpha = 0$, $w(t) = 1$, the Gauss-Jacobi polynomials specialize to the Legendre polynomials. These are generated by the recursion

(3) $$p_0 = 1, \quad p_{n+1} = ((2n+1)xp_n - np_{n-1})/(n+1).$$

Here are the first ten Legendre polynomials for the interval $[a, b] = [-1, 1]$, generated by this recursion:

$$\begin{aligned}
P_1(x) &= x, \\
P_2(x) &= -\left(\frac{1}{2}\right) + \frac{3\,x^2}{2} \\
P_3(x) &= \frac{-3\,x}{2} + \frac{5\,x^3}{2} \\
P_4(x) &= \frac{3}{8} - \frac{15\,x^2}{4} + \frac{35\,x^4}{8} \\
P_5(x) &= \frac{15\,x}{8} - \frac{35\,x^3}{4} + \frac{63\,x^5}{8} \\
P_6(x) &= -\left(\frac{5}{16}\right) + \frac{105\,x^2}{16} - \frac{315\,x^4}{16} + \frac{231\,x^6}{16} \\
P_7(x) &= \frac{-35\,x}{16} + \frac{315\,x^3}{16} - \frac{693\,x^5}{16} + \frac{429\,x^7}{16} \\
P_8(x) &= \frac{35}{128} - \frac{315\,x^2}{32} + \frac{3465\,x^4}{64} - \frac{3003\,x^6}{32} + \frac{6435\,x^8}{128} \\
P_9(x) &= \frac{315\,x}{128} - \frac{1155\,x^3}{32} + \frac{9009\,x^5}{64} - \frac{6435\,x^7}{32} + \frac{12155\,x^9}{128} \\
P_{10}(x) &= -\left(\frac{63}{256}\right) + \frac{3465\,x^2}{256} - \frac{15015\,x^4}{128} + \frac{45045\,x^6}{128} - \frac{109395\,x^8}{256} + \frac{46189\,x^{10}}{256}
\end{aligned}$$

4	1.999984228457721944767532072144696487557194483115
5	2.000000110284471879766230094981509385528232424409
6	1.999999999477270715570406402679408076715703270262
7	2.00000000000179047139889795228027253968825895516
8	1.99999999999999536042661896677198535555733701582
9	2.0000000000000000094136064072597719396414294942
10	1.9999999999999999999846379297653491184960575953
11	2.0000000000000000000000206013457173278936238709
12	1.999999999999999999999999976892865748089102133
13	2.000000000000000000000000000021998288119455387
14	1.999999999999999999999999999999982001465659210
15	2.00000000000000000000000000000000001279196248
16	1.99999999999999999999999999999999999999202885
17	2.00000000000000000000000000000000000000000439
18	2.0000000000000000000000000000000000000000000

Table 1. Approximating $\frac{\pi}{2}\int_{-1}^{1}\cos(\frac{\pi}{2}t)dt$ using the roots of the nth Legendre polynomial.

4	1.098570353649360421369450714823175319789315274643
5	1.098609241812471960520412741139524450426200089726
6	1.098612068116940643764150014765232798503789743085
7	1.09861227273834560823704233014754244788917670177
8	1.09861228751917825294586526806601222503431516024
9	1.0986122885853231560758415201023000760438081814
10	1.0986122886621485872861135030048483168226650251
11	1.0986122886676806754603956463038864607321813301
12	1.0986122886680788273422876748043133557118525407
13	1.098612288668107471652784971526472129102334449
14	1.098612288668109531789219669531325192092724242
15	1.09861228866810967992129366358635127998592044
16	1.09861228866810969057052073640076421999813943
17	1.09861228866810969133597337241914282263302985
18	1.0986122886681096913909859076206421219228913
19	1.0986122886681096913949391857585049768764517
20	1.0986122886681096913952232475480128000949082
21	1.098612288668109691395243657121154867973554
22	1.098612288668109691395245123430128760261856
23	1.09861228866810969139524522876968496262020
24	1.09861228866810969139524523633688431079351
25	1.09861228866810969139524523688045909892622
26	1.0986122886681096913952452369195041669804
27	1.0986122886681096913952452369223086820056
28	1.0986122886681096913952452369225101173798
29	1.098612288668109691395245236922524585148
30	1.098612288668109691395245236922525624245

Table 2. Approximating $\log(3) = \int_{-1}^{1} \frac{1}{2+t} dt$ using the roots of the nth Legendre polynomial. *Mathematica* gives the first 50 digits of $\log 3$ as 1.0986122886681096913952452369225257046474905578227

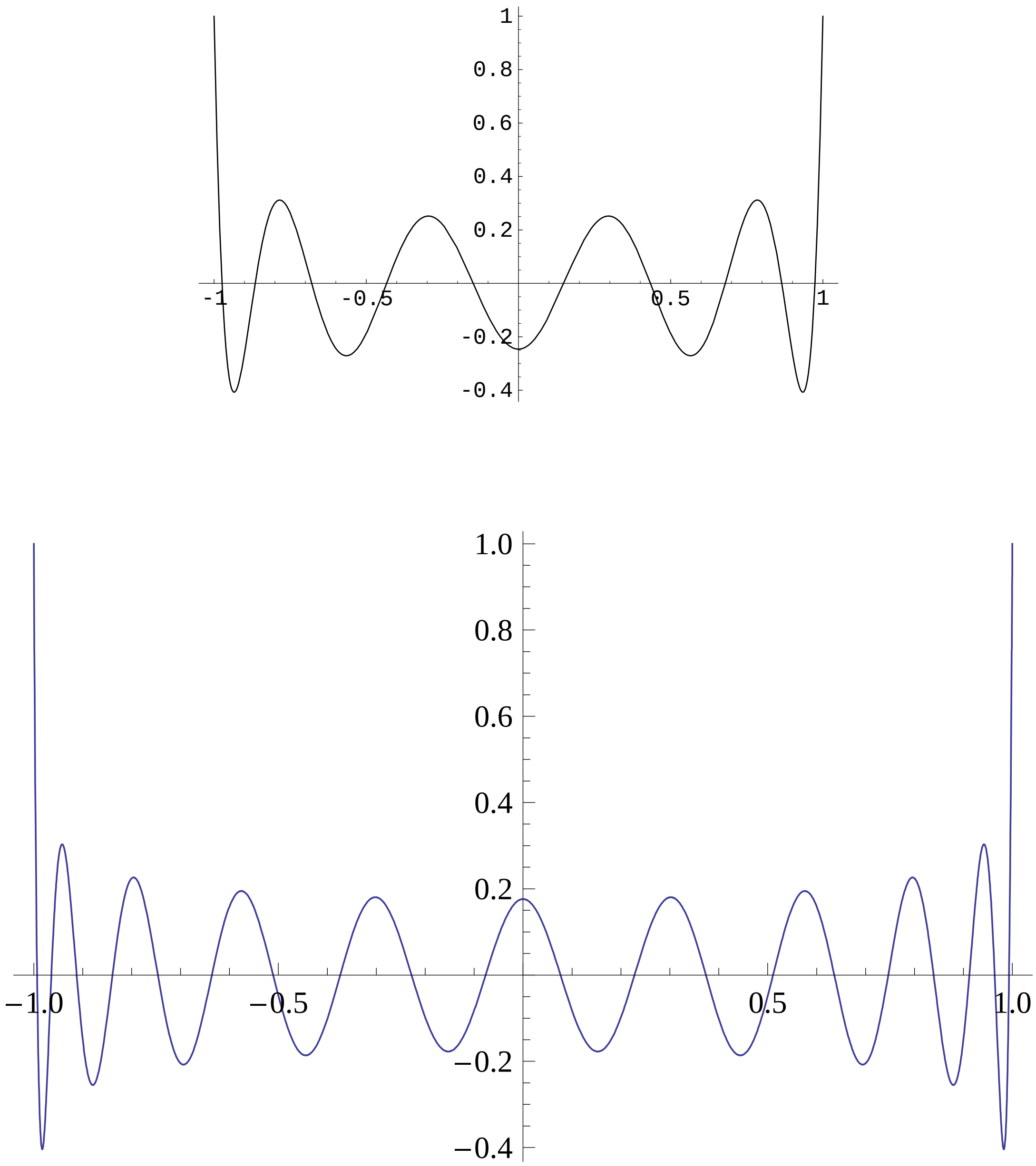

Figure 1. Some example of Legendre polynomials. The roots of P_n are the optimal n points to sample to compute an integral of the form $\int_{-1}^{1} f(t)dt$ in the sense that they will give the correct answer if f if a polynomial of degree at most $2n+1$. Shown are $n = 10, 20$.