**Accuracy of Gaussian Quadrature** PhL 1.15.04

Contents:

**Theorem 1:** The Gaussian Quadrature n+1 point formula is valid for an arbitrary polynomial of degree n where any distinct $x_i$ may be used.

**Theorem 2.** If we *assume* that the n+1 point GQ formula is valid for polys of some degree M, then for such polys we have two equivalent ways to state the Gauss formula, and either implies the other:

$$\int dx\, w(x)f(x) = \Sigma_{i=0}^{n} A_i\, f(x_i) \qquad (1)$$

$$\Sigma_{i=0}^{n} A_i\, p_k(x_i) = \delta_{k,0}\, (h_0/k_0) \qquad k = 0,1,2....M \qquad (2) \qquad \text{"sum rule"}$$

**Theorem 3**. If $x_i$ are the zeros of $p_{n+1}(x_i)$, then the above "sum rule" is valid for a larger range of k:

$$\Sigma_{i=0}^{n} A_i\, p_k(x_i) = \delta_{k,0}\, (h_0/k_0) \qquad k = 0,1,2....2n+1 \qquad (2)$$

**Theorem 4:** If the $x_i$ are chosen as the zeros of the orthog poly $p_{n+1}(x)$, then the n+1 point Gaussian quadrature formula is true for any poly of degree 2n+1. [ Otherwise, it is only true up to degree n. ] this theorem follows at once from Theorem 3 and Theorem 2.

**Appendix A Theorem:** There is a 1-to-1 correspondence between degree n polynomials written in the form $\Sigma_{i=0}^{n} a_i x^i$ and those written in the form $\Sigma_{i=0}^{n} L_i(x)f(x_i)$ for any set of n+1 distinct points $\{ x_i \}$.

---

**Theorem 1:** The Gaussian Quadrature n+1 point formula is valid for an arbitrary polynomial of degree n.

Proof: Start with the formula where we use an arbitrary set of n+1 distinct points $x_i$ in our interval:

$$\int dx\, w(x)f(x) = \Sigma_{i=0}^{n} A_i\, f(x_i) \qquad \text{" Gaussian formula"} \quad \text{with n+1 points}$$

where

$$A_i = \int dx\, w(x)L_i(x)$$

We know from Appendix A that we can represent an *arbitrary* degree n poly as (it has n+1 params)

$$f(x) = \Sigma_{i=0}^{n} L_i(x)f(x_i).$$

Then here is a proof of the Gauss formula for f(x) a arbitrary degree n poly:

$$\int dx\, w(x)\, f(x) = \int dx\, w(x)\, \{\Sigma_{i=0}^{n}\ L_i(x)\, f(x_i)\} = \Sigma_{i=0}^{n} [\int dx\, w(x)\, L_i(x)\,]f(x_i) = \Sigma_{i=0}^{n} A_i\, f(x_i)$$

*Any* distinct set $\{x_i\}$ of n+1 distinct points may be used. I don't think the $x_i$ have to be within the integration interval (a,b), but I might be wrong about that. Later when we decide to use the zeros of $p_{n+1}(x)$, they *will* be in the interval because that is where all n+1 zeros of $p_{n+1}$ are located.

---

**Theorem 2.** If we *assume* that the Gauss formula is valid for polys of some degree M, then for such polys we have two equivalent ways to state the Gauss formula, and either implies the other:

$$\int dx\ w(x)f(x) = \Sigma_{i=0}^{n} A_i\ f(x_i) \qquad (1)$$

$$\Sigma_{i=0}^{n} A_i\ p_k(x_i) = \delta_{k,0}\ (h_0/k_0) \qquad k = 0,1,2....M \qquad (2)$$

So far, we know only that (1) is true for $M \le n$, as shown in Theorem 1, but this fact does not interfere with our statement and proof of Theorem 2. In the proof that follows, $p_n(x)$ is the unique set of orthogonal polynomials (apart from normalization) that is defined by the interval (a,b) and weight function w(x). The integral in (1) is over this interval. We also use the scalar product notation (f,g).

Proof:

**=> First, we will show that (1) => (2):**

Using the Erdelyi notation and $p_0(x) = k_0$ and $h_n = (p_n, p_n)$ we know that

$$\int dx\ w(x)\ p_0(x) = (1,p_0) = (1/p_0)(p_0, p_0) = h_0/k_0$$

Since the $p_k$ are orthogonal, we know this is true for all k,

$$\int dx\ w(x)\ p_k(x) = (1,p_k) = \delta_{k,0}\ (h_0/k_0) = \qquad k = 0,1,2.....\infty$$

If we now assume that (1) is true, we can apply (1) in particular to $f(x) = p_k(x)$ for k = 0..M to get

$$\Sigma_{i=0}^{n} A_i\ p_k(x_i) = \int dx\ w(x)p_k(x) = \delta_{k,0}\ (h_0/k_0) \qquad k = 0,1,2....M$$

and we have thus shown that (1) => (2) QED.

**=> Second, we will show that (2) => (1):**

We expand an arbitrary degree M poly f(x) in the $p_k(x)$ as follows,

$$f(x) = \Sigma_{k=0}^{M} b_k\ p_k(x) \qquad \text{// this is a poly of degree} \le M$$

We then integrate this to get an expression for the LHS of (1),

$$\int dx\ w(x)f(x) = \int dx\ w(x)\{\Sigma_{k=0}^{M} b_k\ p_k(x)\} = \Sigma_{k=0}^{M} b_k\ (1,p_k) = \Sigma_{k=0}^{M} b_k\ \delta_{k,0}\ (h_0/k_0) = b_0(h_0/k_0)$$

and this is true for *any positive integer M*. So we have shown that LHS of (1) = $b_0(h_0/k_0)$.

Now let's examine the RHS of (1):

$$RHS = \Sigma_{i=0}^{n} A_i f(x_i)$$

Insert the same expansion above for f(x) of degree M,

$$RHS = \Sigma_{i=0}^{n} A_i \{ \Sigma_{k=0}^{M} b_k p_k(x_i)\} = \Sigma_{k=0}^{M} b_k [ \Sigma_{i=0}^{n} A_i p_k(x_i) ]$$

If we now assume (2) is valid for polys of degree M, we may make this replacement

$$[\Sigma_{i=0}^{n} A_i p_k(x_i)] = \delta_{k,0} (h_0/k_0) \qquad k = 0,1,2....M$$

and we find that

$$RHS = \Sigma_{k=0}^{M} b_k [ \Sigma_{i=0}^{n} A_i p_k(x_i) ] = \Sigma_{k=0}^{M} b_k \delta_{k,0} (h_0/k_0) = b_0(h_0/k_0)$$

We have thus shown that RHS (1) = LHS (1) = $b_0(h_0/k_0)$, so (1) is true. QED.

---

**Theorem 3**. Consider the above version (2) of the n+1 point Gauss formula:

$$\Sigma_{i=0}^{n} A_i p_k(x_i) = \delta_{k,0} (h_0/k_0) \qquad k = 0,1,2....M \qquad (2)$$

We know from Theorem 1 that (2) is true for $M \le n$. The current theorem claims that, *if* we select the points $x_i$ to be the zeros of $p_{n+1}(x_i)$, then (2) is true for the extended range $M \le 2n+1$.

Proof: This proof uses the following ingredients:

(a) the fact that $p_{n+1}(x_i) = 0$ since we are selecting the n+1 points $x_i$ to be the n+1 zeros of $p_{n+1}(x)$
(b) the recursion formula which schematically reads $p_{k+1} = xp_k + p_k + p_{k-1}$
(c) the "angular momentum" theorem which says schematically that $xp_k = p_{k+1} + p_k + p_{k-1}$

By "schematically" we mean we are not showing various constants like $A_n$, $B_n$, $C_n$ in the recursion relation, or coefficients in the product expansion. We set all constants to 1 because we are only going to be concerned with whether terms are zero or non-zero; we won't be caring about exact values of non-zero terms.

We start with a derivation of item (c) above. The "angular momentum" product rule (see my Erdelyi Addendum notes, this result does not appear in Erdelyi's book) states:

$$p_{j1}(x)p_{j2}(x) = \sum_{\mathbf{j=|j1-j2|}}^{\mathbf{j1+j2}} \mathbf{a_j}\, p_j(x)$$

and the proof of this fact you will see there is quite trivial and is based on nothing but the basic properties of the orthogonal polys, in particular, their orthogonality. We use integers $j_1$, $j_2$ and j for label names as a

reminder that this is very similar to the "angular momentum addition theorem" in SO(3) group theory. Knowledge of group theory of course is irrelevant to what we are doing here, and the reader is welcome to ignore the phrase "angular momentum". The above expansion is just a simple fact. We now apply this fact to claim that

$$x\, p_k = (p_0 + p_1)p_k = (p_0 p_k) + (p_1 p_k) = (p_k) + (p_{k+1} + p_k + p_{k-1}) = p_{k+1} + p_k + p_{k-1}$$

where again we are setting all constants to 1. Certainly we can write x as a linear combination of $p_0$ and $p_1$with some specific coefficients. Clearly in this notation we have $p_0 p_k = p_k$ by our theorem. And the second term is what gives the interesting part, $p_1 p_k = p_{k+1} + p_k + p_{k-1}$. So the effect of applying x to $p_k$ is toe create a "spread" of unit in each direction.

Now let's indicate by $\{ p_a \rightarrow p_b \}$ a descending range of $p_n$ . For example, $p_{k+1}+p_k+p_{k-1} = \{ p_{k-1} \rightarrow p_{k-1} \}$. The range need not be "full" (as it is in our example), we just mean that a is the largest index and b is the smallest index in a sum of terms. Then rule (c) $x p_k = p_{k+1} + p_k + p_{k-1}$ tells us that

$$x\{ p_a \rightarrow p_b \} = \{ p_{a+1} \rightarrow p_{b-1} \} \qquad (d)$$

so application of x "stretches the range" by one unit in at each end.

Now we begin our process of expanding the range for which (2) is true. As noted, we already know it is true for k = 0,1.2..n.

$$\Sigma_{i=0}^{n} A_i\, p_k(x_i) = \delta_{k,0}\, (h_0/k_0) \qquad k = 0,1,2....n \qquad (2)$$

It is obviously true for k = n+1 because by (a) we know $p_{n+1}(x_i) = 0$.

For $p_{n+2}$ we use the recursion (b) which says $p_{n+2} = x_i p_{n+1} + p_{n+1} + p_n$ where we have $x = x_i$ everywhere. This simplifies to $p_{n+2} = p_n$ since $p_{n+1} = 0$ and we already know that (2) is true for k=n, so now we know that it is true for k = n+2 as well. There is one exception, however. If we have n=0, meaning a 1-point Gaussian formula, then $p_{n+2} = p_n \Rightarrow p_2 = p_0$, and we know that (2) $\neq 0$ for $p_0$, and therefore (2) is non-zero for $p_2$ in this case. This agrees with the claim that the result is valid only for $k \leq 2n+1 = 1$. So we now assume that $n > 0$ and continue:

[ By the way, the above case is the ONLY place we need to use the fact that $p_{n+1}(x_i) = 0$. It allows us to "get started". Accordingly, we will not delete $p_{n+1}$ terms in the following, just to make this fact clear. ]

For $p_{n+3}$ the recursion now says $p_{n+3} = x_i p_{n+2} + p_{n+2} + p_{n+1}$. But from above we know that $p_{n+2} = p_n$, so we get $p_{n+3} = x_i p_n + p_{n+2} + p_{n+1} = \{ p_{n+1} \rightarrow p_{n-1} \} + p_{n+2} + p_{n+1} = \{ p_{n+2} \rightarrow p_{n-1} \}$ where we have used rule (d) to expand $x p_n$ . The "danger" in this expression is the term $p_{n-1}$. If we have n=1, then this is $p_0$ and (2) is non zero for $p_{n+3} = p_4$. So if n = 1, we would conclude that our theorem was not true for $p_4$, and this agrees with our claimed validity limit of 2n+1 = 3. So we now assume that $n > 1$ and continue:

For $p_{n+4}$ the recursion now says $p_{n+4} = x_i p_{n+3} + p_{n+3} + p_{n+2}$ . We insert $p_{n+3} = \{ p_{n+2} \rightarrow p_{n-1} \}$ to get that $p_{n+4} = x_i \{ p_{n+2} \rightarrow p_{n-1} \} + p_{n+3} + p_{n+2} = \{ p_{n+3} \rightarrow p_{n-2} \}$. The higher terms all give 0 in (2), but now we are concerned about the $p_{n-2}$ term. But let's now jump to the general case.

Here is the general pattern of the recursion formulas. The first three in this list we have explicitly derived. We see that the "range" increases by 1 in each direction for each step,

| | | |
|---|---|---|
| $p_{n+2} = \{ p_{n+1} \rightarrow p_{n-0} \}$ | 1 = 2-1 | 0 = 2-2 |
| $p_{n+3} = \{ p_{n+2} \rightarrow p_{n-1} \}$ | 2 = 3-1 | 1 = 3-2 |

$p_{n+4} = \{ p_{n+3} \rightarrow p_{n-2} \}$ 3 = 4-1 2 = 4-2

.....

$p_{n+m} = \{ p_{n+(m-1)} \rightarrow p_{n-(m-2)} \}$ (m-1) = m-1 (m-2) = m-2

and the last line will be "invalid" when m gets up to a value such that $p_{n-(m-2)} = p_0$ . Again, this is so because (2) $\neq 0$ when k =0. By "invalid" we mean that (2) will be not true when we insert $p_{n+m}$ for $p_k$.

The last line is clear from the pattern, but we nevertheless give an induction proof below. Assuming for now it is correct, we find that our largest valid value of m occurs when $p_{n-(m-2)} = p_1$ which says m = n+1 so that n+m = 2n+1. Thus, the largest polynomial for which (2) = 0 is $p_{n+m} = p_{2n+1}$ . We have then shown that (2) is true up to this point, that is,

$$\Sigma_{i=0}^{n} A_i \, p_k(x_i) = \delta_{k,0} \, (h_0/k_0) \qquad k = 0,1,2....(2n+1) \qquad (2)$$

and our theorem is proved QED.

Here is our induction proof of the $p_{n+m}$ expression above. Assume it is true for m:

$$p_{n+m} = \{ p_{n+(m-1)} \rightarrow p_{n-(m-2)} \} \qquad (3)$$

Then for m+1 we have

$$p_{n+(m+1)} = x_i p_{n+m} + p_{n+m} + p_{n+m-1} = x_i \{ p_{n+(m-1)} \rightarrow p_{n-(m-2)} \} + p_{n+m} + p_{n+m-1}$$

$$= \{ p_{n+m} \rightarrow p_{n-(m+1-2)} \} + p_{n+m} + p_{n+m-1} \qquad \text{// where we used rule (d)}$$

$$= \{ p_{n+m} \rightarrow p_{n-(m+1-2)} \} = \{ p_{n+(m+1-1)} \rightarrow p_{n-(m+1-2)} \}$$

so we have shown that if (3) is true for m, it is also true for m+1, and of course we have shown already that (3) is true explicitly for m = 2,3 and 4.

---

**Theorem 4:** If the $x_i$ are chosen as the zeros of the orthog poly $p_{n+1}(x)$, then the Gaussian quadrature formula is true for any poly of degree 2n+1. [ Otherwise, it is only true up to degree n. ]

Proof: We have shown in Theorem 1 that the Gaussian formula is generally true for degree n. We have shown in Theorem 3 that, when the $x_i$ are chosen such that $p_{n+1}(x_i) = 0$, the alternate form (2) of the Gauss formula is true for k = 0,1,2 up to (2n+1), that is to say

$$\Sigma_{i=0}^{n} A_i \, p_k(x_i) = \delta_{k,0} \, (h_0/k_0) \qquad k = 0,1,2....(2n+1) \qquad (2)$$

Theorem 2 then tells us that, since (2) is true for M = 2n+1, then (1) is also true for M = 2n+1, and (1) is the official Gaussian quadrature formula. QED.

---

**Comment #1**: Sometimes one uses the n-point Gaussian formula (instead of our n+1 point one). In that case, the general Gauss formula is true for polys up to degree n-1 (instead of n), and the extended formula is then true up to 2(n-1)+1 = 2n-1 (instead of 2n+1 ). Also, the $x_i$ are then the n zeros of $p_n$ (not $p_{n+1}$).

**Comment #2:** The formula $\Sigma_{i=0}^{n} A_i\, p_k(x_i) = \delta_{k,0}\, (h_0/k_0)$ for k = 0,1,2....M (with $A_i = \int dx\, w(x) L_i(x)$ ) is really a statement about the orthogonal polynomial set { $p_k$}. It has nothing to do with any function we are integrating. Notice that the $x_i$ appear inside $L_i(x)$ (see appendix below). If we change from one general set of $x_i$ to another, the $A_i$ all change, but then of course so do the $p_k(x_i)$, and these changes occur in a way that the sum $\Sigma_{i=0}^{n} A_i\, p_k(x_i)$ is unchanged and still equals $\delta_{k,0}\, (h_0/k_0)$. For general $x_i$ this peculiar property of the $\{p_k\}$ is true for $k = 0 \rightarrow n$, but for the specific $x_i$ which are the zeros of $p_{n+1}$ it is true for the much larger range $k = 0 \rightarrow 2n+1$. Therefore, the accuracy of the Gaussian formula applied as an *approximation* to an arbitrary function f(x) is obviously going to be much improved when we use $x_i$ which are these zeros. You can "fit" an arbitrary f(x) much better with a 2n+1 poly compared to an n poly. Scheid gives detailed error formulas which show that n+1 point error is proportional to $f^{(2n+2)}(\zeta)$ where $\zeta$ is some point in (a,b). Clearly for a polynomial of degree 2n+1, the error formula then gives 0 error. Deriving that error formula is another way to get the result we have shown above, but we wanted here to show directly how it is that those $x_i$ work with the recursion formula to expand the range of the formula.

---

**Appendix A:**

**Theorem:** There is a 1-to-1 correspondence between degree n polynomials written in the form $\Sigma_{i=0}^{n} a_i\, x^i$ and those written in the form $\Sigma_{i=0}^{n} L_i(x) f(x_i)$ for any set of n+1 distinct points { $x_i$ }.

Proof: First, we show that $g(x) = \Sigma_{i=0}^{n} L_i(x) f(x_i)$ is in fact a polynomial of degree n. The reason is that each factor $L_i(x)$ is a poly of degree n, and we are then adding up polys of degree n to get an overall poly of degree n. To see that $L_i(x)$ is degree n, just write it out:

$\pi(x) = (x - x_0) \ldots (x - x_n)$ = degree n+1

$L_i(x) = \pi(x)/[(x-x_i)\pi'(x_i)]$ = degree n

Ignoring constants, $L_i(x)$ is just the product of n factors $(x - x_k)$ where the $i^{th}$ factor $(x - x_i)$ is missing.

Second, it is pretty obvious that $\Sigma_{i=0}^{n} a_i\, x^i$ is also a poly of degree n.

Now, assume $f(x) = \Sigma_{i=0}^{n} a_i\, x^i$ . Suppose we specify n+1 distinct $x_k$ points, so that $y_k \equiv f(x_k) = \Sigma_{i=0}^{n} a_i\, (x_k)^i$ or $\Sigma_{i=0}^{n} (x_k)^i\, a_i = y_k$ . We then have this matrix equation where X is n+1 square

$$X\mathbf{a} = \mathbf{y} \qquad \text{where } X_{ki} = (x_k)^i$$

As I have shown in the Erdelyi Addendum notes (and see page 158 top), $\det X \neq 0$ as long as the n+1 points are all different, a fact that is not obvious. Therefore, we can invert to get

$$\mathbf{a} = X^{-1}\, \mathbf{y}$$

So there is a 1-1 correspondence between the two ways you can parameterize a n degree poly:
{ $a_i$ } $\leftrightarrow$ { $y_k$ }. Each set of course contains n+1 numbers. If the points are not distinct, then $\det X = 0$ and then X has a nullspace and we might get $X\mathbf{a} = 0$ for some choices of **a**, so then **y** cannot describe an arbitrary polynomial.

**Corollary**: An "arbitrary" polynomial f(x) of degree n can be written as $\Sigma_{i=0}^{n} L_i(x_k)f(x_i)$.

The corollary just says this: pick an arbitrary deg n polynomial by picking coefficients $\{ a_i \}$. We can then compute $\mathbf{y} = X\mathbf{a}$, and then we know how to write that arbitrary deg n poly in the form $\Sigma_{i=0}^{n} L_i(x_k)y_i$ .