Erdelyi note on orthogonal polys PhL 1.11.05

The time has now come to clean up this subject. I think Erdelyi's presentation looks good, and I can supplement it as needed. I did take some earlier notes while reading some Jim Ball papers.

In these many pages I review in great detail the opening "general" pages of this chapter, and I skip the details of the specific functions described, such as Jacobi's etc. The main function of this document is to have a repository for the derivations of some quite messy formulas that I have never derived before. And in going through all this, I got much more familiar with the notation used, especially for the many constants.

In a separate document I give an "executive overview" of what I learned in this chapter, and what I did while reading it. This document is mostly just fine details. Things are done in book section order.

10.1 Systems of Orthogonal Functions

Given an interval (a,b) and a non-negative weight function w(x), we can define a scalar product $\langle f|g\rangle$ and thus a Hilbert space. The weight function has to be non-negative so that $\langle f|f\rangle = ||f||^2$ will be positive for any function f. The integral here is associated by Erdelyi with the name Stieltjes.

Comment: Web says that the usual integrals we deal with are Riemann integrals, and that Lebesgue integrals are not part of the ordinary world. The Stieltjes integral derives from this strange sum idea:

$$\Sigma_i \; f(x_i) \, [\, g(x_{i+1}) - g(x_i) \,] \text{ generalizing } \quad \Sigma_i \; f(x_i) \, [\, x_{i+1} - x_i \,]$$

where x_i^* is some arbitrary point in the interval (x_i, x_{i+1}). In the limit, this is $\int f(x)dg$, and so you can think of this as a generalization of $\int f(x)dx$. If things are reasonable, you write $dg = g'(x)\,dx = w(x)\,dx$. Now suppose you use $g(x) = w_i\,\theta(x - x_i) + w_{i+1}\,\theta(x - x_{i+1}) + w_{i+2}\theta(x - x_{i+2}) + \ldots$ which is just a sum of steps of height w_i. and width $[\, x_{i+1} - x_i \,]$. Then $w(x) = g'(x) = w_i\,\delta(x - x_i) + w_{i+1}\,\delta(x - x_{i+1}) + w_{i+2}\,\delta(x - x_{i+2}) \ldots$ Then you get $\int f(x)dg = \Sigma_i\, w_i\, f(x_i)$. So the claim is that the Stieltjes idea encompasses both continuous and discrete scalar products. I think in the second case, g(x) is a distribution rather than a function.

I think the scalar product as defined above gives the L_2 Hilbert Space of square integrable functions, and we can take the powers of x as a basis for a space of infinite dimension, or we can stop at some power and have a finite dimensional Hilbert Space. The basis functions are always linearly independent, perhaps not orthogonal. You can orthonormalize using G-S and this is mentioned bottom of page 154 where you start with functions ψ which are not orthonormal, and you create functions ϕ which are.

Lemma 1.1 Consider the following weighted and summed determinant expression

$$Q = \Sigma_i \; f_i \begin{vmatrix} a_1 & a_2 \\ b_1(i) & b_2(i) \end{vmatrix}$$

We claim that we can move the $\Sigma_i\; f_i$ into the bottom row and let it act on the terms there, so we have

$$Q = \begin{vmatrix} a_1 & a_2 \\ \Sigma_i\; f_i b_1(i) & \Sigma_i\; f_i b_2(i) \end{vmatrix}$$

Proof: To see this, just write the thing out as follows

$Q = \Sigma_i\ f_i\ [\ \varepsilon_{\ell m}\ a_\ell\ b_m(i)] = \ \varepsilon_{\ell m}\ a_\ell\ [\ \Sigma_i\ f_i\ b_m(i)]$ = as shown in second expression above.

Corollary 1.1 We could also state this replacing $\Sigma_i\ f_i\ = \int dx\ f(x)$ in which case we get:

$$Q = \int dx\ f(x) \begin{vmatrix} a_1 & a_2 \\ b_1(x) & b_2(x) \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ \int dx\ f(x)b_1(x) & \int dx f(x)b_2(x) \end{vmatrix}$$

These things help us to understand the comments on page 155. Suppose instead of doing the G-S process as described bottom of 154 and top of 155, we construct the $\phi_n(x)$ as shown in (10), where we are using a 2x2 matrix inside the det for $\phi_1(x)$ and a 3x3 for $\phi_2(x)$ and an (n+1) x (n+1) for $\phi_n(x)$.

WARNING: Notice that this author uses ϕ and φ as different things. These are both the same Greek letter "phi", so that is quite confusing. The ϕ are given by (10), whereas the φ are going to be normalized!

At this point, each statement really requires work, so here we go:

Claim 1.1: The $\phi_n(x)$ given by (10) is orthogonal to ψ_k for $k < n$.

Proof: If we compute (ψ_k, ϕ_n) using our Corollary 1.1 above, the bottom row of the determinant becomes the following: $(\psi_k, \psi_0), (\psi_k, \psi_1), (\psi_k, \psi_2)$.... . But this duplicates row k of the determinant, as long as we restrict to k = 0, 1, 2...n-1. QED.

Claim 1.2: The $\phi_n(x)$ given by (10) is orthogonal to ϕ_k for $k < n$.

Proof: ϕ_k for $k < n$ is given by a lincom of ψ_k for $k < n$. This is what (10) says. But according to Claim 1, we know that $\phi_n(x)$ is orthogonal to *all* these ψ_k for $k < n$. Thus, $\phi_n(x)$ is orthogonal to any lincom you can make of the ψ_k for $k < n$, and therefore $\phi_n(x)$ is orthogonal to ϕ_k for $k < n$.

Definition 1.1: Define $\mathbf{a_{jk}} = (\psi_j, \psi_k) = \int dx\ w(x)\ \psi_j(x)\psi_k(x)$

Definition 1.2: The **Gram Determinant** is $G_n = \det(a_{jk})$ where we let j,k range 0 to n. Thus, G_n is the determinant of a matrix of dimension n+1.

Claim 1.3: This matrix G_n is the upper left n+1 size subdeterminant of the matrix for ϕ_{n+1}. It is therefore the cofactor of the lower right element which is ψ_{n+1}.

Proof: Cross out the last row and column of the matrix for ϕ_{n+1}, and look at the subdeterminant so formed. The top row has a_{00} through a_{0n} . The first column has a_{00} through a_{n0}. This matrix is exactly G_n.

Corollary 1.3: In the matrix expression (10) for $\phi_n(x)$, the coefficient of $\psi_n(x)$ is G_{n-1}. Later when we choose the specific form $\psi_n(x) = x^n$, we will have that $G_{n-1}$is the coefficient of the leading power x^n in the matrix expression for $\phi_n(x)$.

Claim 1.4: The expression shown with the ξ coefficients is a quadratic form.

Proof: Write it in the form $\Sigma_{jk}\, \xi_j^T\, a_{jk}\, \xi_k = \boldsymbol{\xi}^T\, a\, \boldsymbol{\xi}$ where matrix a is n+1 x n+1.

Claim 1.5: G_n is a discriminant of this quadratic form.

Proof: According to M&M page 323, this matrix and all submatrices are called discriminants of a quadratic form, to G_n is the largest one.

Theorem 1.1: The $\| \phi_k \|^2 = G_{k-1}G_k$. From this, it follows that $\| \varphi_k \|^2 = 1$ as shown in (11).

Proof: This took me a while to see, but now it's obvious. Consider:

$$\| \phi_k \|^2 = (\phi_k, \phi_k)$$

Now look at the expansion (10) for $\phi_k = \xi_0\psi_0 + \xi_1\psi_1 + ... + \xi_k\psi_k$ where the ξ are the messy cofactors which arise from doing the Laplace expansion of the determinant along the bottom row. However, we know that $\xi_k = G_{k-1}$ from Claim 3 above. We also know from Claim 1 that ϕ_n is orthogonal to the lower ψ_k. Therefore:

$$\| \phi_k \|^2 = (\phi_k, \phi_k) = (\xi_0\psi_0 + \xi_1\psi_1 + ... + G_{k-1}\psi_k, \phi_k) = G_{k-1}\, (\psi_k, \phi_k).$$

But, consider (ψ_k, ϕ_k) which we can form by applying $(\psi_k,$ to expansion (1) for ϕ_k, using the idea of Corollary 1.1 above. We then find that $(\psi_k, \phi_k) = G_k$! That is, we get the entire regular matrix a_{xy} which is of dimension k+1. Thus we have shown that $\| \phi_k \|^2 = G_{k-1}\, G_k$. QED.

Comment: Expansion (10) says that $\phi_0 = \psi_0$. Our equation (11) then says that $\| \varphi_0 \|^2 = \| \phi_0 \|^2 / (G_0 G_{-1})$ We take $G_{-1} = 1$ as they suggest, and $G_0 = a_{00} = (\psi_0,\psi_0) = (\phi_0,\phi_0)$, so things work right for the 0 case. That is to say, $\| \varphi_0 \|^2 = \| \phi_0 \|^2 / (\phi_0,\phi_0) = 1$.

I am going to ignore equations (12) and (13) unless I find they are needed later!

Theorem 1.2: The interval (a,b) and the weight w(x) determine a unique set of orthonormal functions.

The dimension N has not really been mentioned. But suppose we start with the non-orthogonal powers x^n, then we can just start orthonormlizing them as far as we want to go. I think I downtick Erdelyi here for being vague about Hilbert Space and the dimension of that space, and so on. Maybe yet to come.

He ends this section by commenting that there are a few ways people decide to normalize functions. You might want to have them be monic, or to have a specified value at some point.

10.2 The Approximation Problem: Fitting a function with a lincom of orthog polys.

The question brought up here is this: We want to "fit" a general function f(x) as a lincom of ortho polys with some coefficients. How do we select the coefficients to make a best least squares fit? This takes us back to the nice discussion in Scheid on pages 242-243, Chapter 21. If you are trying to best fit a vector f

in a Hilbert space with a vector q in a subspace (eg, in a space of some finite number of ortho polys for your basis), then the best you can do is get to the point p in the subspace which is closest to f, as expressed in terms of the norm in the space, that is, minimize || f - q ||. This point of closest approach to f is reached (in the case of an orthog basis) when you use the Fourier coefficients (f,ϕ_k) as your lincom coefficients. Remember this whole business that, if you use *ortho* polys, then you improve your fit by adding more terms, and you never have to alter any of the earlier coefficients, and recall how this is not the case if you use non-orthog polys.

Erdelyi shows explicitly how choosing the Fourier coefficient minimizes your error, which is those extra terms in the equation after (2). We then get (4) which implies (5) which is the famous **Bessel's Inequality** $\Sigma\, a_i^2 \le \int dx\, w(x)\, f^2(x)$. He shows this with the sum going to infinity. But we know that for any "reasonable" basis functions (that is, **closed** basis), when you add up all the pieces you get equality, which is called **Parseval's Formula** (which I am familiar with in normal Fourier integral analysis), in the case f(x) is in L_2. In this case, our poly approx to f(x) is "**converging in the mean**". I think this has to do with the space being **complete** and therefore really a Hilbert Space, and we associate the **Riesz-Fischer** theorem with this concept. We also have a **Weierstrauss** Theorem on this subject. When the interval (a,b) is finite, you always have things being "closed" and complete. Infinite intervals are more worrisome.

Just in passing, recall from Scheid that if you attempt a least squared fit with a lincom of x^k powers (which is a non orthog basis), you end up with those "**normal equations**", as on page 241 there. You have to solve these normal equations to find your coeffs, and as you add more terms, all the coeffs "move".

10.3 General Properties of Orthog Polys

Our theorem is restated: interval and weight determine the polys uniquely apart from constant in each. Now in (1) we define the moments of a weight function, $c_n = \int dx\, w(x)x^n$. Equation (2) is trivially true for our case $\psi_k(x) = x^k$. We can now apply our general theory of Section 10.1 to get G_n as shown, since we have that $a_{jk} = (\psi_j, \psi_k) = c_{j+k}$. So to get the subscript on c in (3), just add row and column. For now, I ignore the right equation shown in (3).

Equation (4) is now just (10) of the first section for $\psi_k(x) = x^k$, where we have added a particular normalizing factor k_n / G_{n-1}. We are free to select k_n for different purposes, BUT notice the following fact: k_n is always the coefficient of x^n in $p_n(x)$. This fact is used later in deriving the recursion formula.

Theorem: $p_n(x) = (k_n/G_{n-1})$ * matrix where k_n is the coefficient of x^n. [This is equation (4).]

Proof: if we look at only the matrix part, we know that coeff of x^n is precisely G_{n-1}. This cancels the factor shown above, and you are then left with k_n as the coefficient of x^n.

Fact: Notice that $p_0(x) = k_0$ and the entire matrix determinant equals 1. Think of k_0x^0 if you want.

Theorem 3.1: Equation (6) for $\| p_k \|^2$ is true with $p_n(x)$ as normalized in (4).

Proof: In (4) the determinant is ϕ_n and we shows above in Theorem 1.1 that $\| \phi_k \|^2 = G_{k-1}G_k$, therefore we have that $\| p_k \|^2 = (k_k/G_{k-1})^2 \| \phi_k \|^2 = (k_k/G_{k-1})^2 G_{k-1}G_k = k_k^2\, G_k/G_{k-1}$.

Corollary 3.1. If we want to have $\| p_k \|^2 = 1$, we need to choose $k_k^2\, G_k/G_{k-1} = 1$ or $k_k^2 = G_{k-1}/G_k$.

Theorem 3.2. All n **zeros** of $p_n(x)$ lie inside the interval (a,b), and between each consecutive pair of zeros of $p_n(x)$ lies exactly one zero of $p_{n+1}(x)$.

The first claim is proved in a way I don't really like much, I am too lazy to write the whole thing up using my Erdelyi decoding method. The second claim is only claimed, refer to Szego he says. I think I have a simpler and clearer proof of the first claim elsewhere using Rolle's Theorem or some such.

Theorem 3.3 (Recursion Formula) The recursion formula shown in (7) page 158 is true for the $p_n(x)$ as defined in (4), where the leading terms of $p_n(x)$ are $k_n x^n + k_n' x^{n-1}$, and where A,B,C are as given on top of page 159 (and these are restated below).

Proof**:** This proof is long and tedious. You have your set of $p_n(x)$ and you construct this combination:

$$f_n(x) = p_{n+1}(x) - A_n \text{ x } p_n(x) \qquad A_n = k_{n+1} / k_n$$

where the k_n is the coefficient of the highest power in $p_n(x)$. I agree completely that $f_n(x)$ must be of degree n or less, because we have canceled out the x^{n+1} term by construction.

$$f_n(x) = \{ k_{n+1} x^{n+1} + \text{lower terms}\} - (k_{n+1} / k_n) \text{ x } \{ k_n x^n + \text{lower terms}\}$$

Now, I also agree that you can therefore write $f_n(x)$ as a lincom of the $p_n(x)$ as shown. For the moment, let's skip the claim that all $\gamma_i = 0$ except γ_0 and γ_1 and come back to this claim very soon.

Meanwhile, I do agree with the next line on page 159 and here is my proof:

$(p_{n-1}, f_n) = (p_{n-1}, p_{n+1} - A_n \text{ x } p_n) = - A_n (p_{n-1}, \text{ x } p_n)$ on the one hand, but then
$(p_{n-1}, f_n) = \gamma_1 (p_{n-1}, p_{n-1})$ on the other hand QED

Suppose I try doing this with some lower power on p ?

$(p_{n-2}, f_n) = (p_{n-2}, p_{n+1} - A_n \text{ x } p_n) = - A_n (p_{n-2}, \text{ x } p_n)$ on the one hand, but then
$(p_{n-2}, f_n) = \gamma_2 (p_{n-2}, p_{n-2})$ on the other hand

The trick is now to write $(p_{n-2}, \text{ x } p_n) = (xp_{n-2}, p_n)$, that is, we just move the x over to the other factor in the integral. But of course xp_{n-2} is a poly of degree n-1 or less which can be expanded in p_r or order n-1 or less, and all these terms will be orthog to p_n. Thus, we have $0 = (xp_{n-2}, p_n) = (p_{n-2}, \text{ x } p_n) \Rightarrow \gamma_2 = 0$. This line or reasoning applies for n-3 and γ_3, and to n-4 and γ_4 and so on. What about γ_0 ?

$(p_n, f_n) = (p_n, p_{n+1} - A_n \text{ x } p_n) = - A_n (p_n, \text{ x } p_n)$ on the one hand, but then
$(p_n, f_n) = \gamma_0 (p_n, p_n)$ on the other hand

This one does not vanish! Therefore we have arrived at this point:

$$p_{n+1} - A_n \text{ x } p_n = \gamma_0 p_n + \gamma_1 p_{n-1}$$

so right now we have the basic proof, we rewrite the above as

$p_{n+1} = p_n (\gamma_0 + A_n x) + \gamma_1 p_{n-1}$

and now we just have to find the expressions for γ_0 and γ_1. The claim is that

(a) $B_n = \gamma_0$ and (b) $C_n = -\gamma_1$.

Let's now prove these two facts.

(b) From above we have that $-A_n (p_{n-1}, x\, p_n) = \gamma_1 (p_{n-1}, p_{n-1}) = \gamma_1 h_{n-1}$. Now imitate the first expansion as follows:

$$f_{n-1}(x) = p_n(x) - A_{n-1}\, x\, p_{n-1}(x)$$

Therefore $(p_{n-1}, x\, p_n) = (x\, p_{n-1}, p_n) = (p_n - f_{n-1}, p_n) / A_{n-1} = (p_n, p_n) / A_{n-1} = h_n/A_{n-1}$. This tells us that $\gamma_1 h_{n-1} = -A_n * h_n/A_{n-1}$ or $\gamma_1 = -(A_n/A_{n-1}) * (h_n/h_{n-1}) = -C_n$. That is, $\boldsymbol{\gamma_1 = -C_n}$. We note that there is a sign error in his statement of this fact on page 159, but he has the sign correct in (7).

(a) From above we have that $-A_n (p_n, x\, p_n) = \gamma_0 (p_n, p_n) = \gamma_0 h_n$. Now go back to an earlier line where we wrote out $f_n(x) = p_{n+1}(x) - A_n\, x\, p_n(x)$,

$$f_n(x) = \{ k_{n+1} x^{n+1} + k'_{n+1} x^n + \text{lower terms}\} - (k_{n+1} / k_n) \times \{ k_n x^n + k'_n x^{n-1} + \text{lower terms}\}$$

where now we show the second term in each expansion. We already saw that the leading terms cancel, now what is the coefficient of the next term?

$$f_n(x) = [\, k'_{n+1} - (k_{n+1} / k_n) k'_n \,]\, x^n + \text{lower terms.}$$

Now as our first step, write $x\, p_n = (p_{n+1} - f_n)/A_n$ so that we have $(p_n, x\, p_n) = -(p_n, f_n)/A_n$. Now in the remaining scalar product, only the leading x^n term above survives (by the usual argument), so we have

$$(p_n, x\, p_n) = -(1/A_n) [\, k'_{n+1} - (k_{n+1} / k_n) k'_n \,] (p_n, x^n)$$

But if we expand x^n we get $x^n = \alpha p_n + \beta p_{n-1} + \ldots$, and comparing powers of x^n gives $1 = \alpha k_n$ so we conclude that $\alpha = 1/k_n$ so we have then

$$\begin{aligned} (p_n, x\, p_n) &= -(1/A_n) [\, k'_{n+1} - (k_{n+1} / k_n) k'_n \,] (p_n, p_n) / k_n \\ &= -(1/A_n) [\, k'_{n+1} - (k_{n+1} / k_n) k'_n \,] h_n/k_n \end{aligned}$$

But this is equal to $-\gamma_0 h_n/A_n$ from our first line above, so we conclude that

$$\begin{aligned} -\gamma_0/A_n &= -(1/A_n) [\, k'_{n+1} - (k_{n+1} / k_n) k'_n \,] (1/k_n) = -k_n / k_{n+1} [\, k'_{n+1} - (k_{n+1} / k_n) k'_n \,] (1/k_n) \\ &= -[\, k'_{n+1} - (k_{n+1} / k_n) k'_n \,] (1/k_{n+1}) \\ &= -[\, (k'_{n+1}/k_{n+1}) - (k'_n/k_n) \,] = -[\, r_{n+1} - r_n \,] \end{aligned}$$

and therefore $\gamma_0 = A_n [\, r_{n+1} - r_n \,] = B_n$ as claimed!

Therefore, our recursion relation is:

$$p_{n+1} = p_n (\gamma_0 + A_n x) + \gamma_1 p_{n-1} = p_n(B_n + A_n x) - C_n p_{n-1} \qquad \text{// agrees with (7) QED}$$

Comments: This recursion relation is a VERY general result. We have not assumed any particular interval (a,b) or weight w(x). We have the p_n normalized to $\|p_n\|^2 = h_n$ and in this normalization, we are required to know the leading two coefficients k_n and k_n' . We have not had to say how many polys we have in a set of polys.

Only one fact is needed to show the recursion relation, and that one fact is the orthogonality of the functions p_n . The factor x is in there because you need it so you can subtract out the highest order term in the proof. Only x can do this, not some other g(x). So ortho polys always have "x" in the recursion relation. This is not necessarily true for other special functions.

Comment on $p_{-1}(x)$. If we make up this item and set it = 0, then we don't have to alter the recursion formula for application to n = 1. There is no such function in our set, or course, since we defined the set to have indices 0,1,2.... and on upwards.

Theorem 3.4 (The Christoffel-Darboux formulas). The first claim is this:

$$\sum_{i=0}^{n} (1/h_i)\, p_i(x)p_i(y) = (k_n/k_{n+1})\,(1/h_n)\,(1/(x-y))\,[\, p_{n+1}(x)p_n(y) - p_{n+1}(y)p_n(x)\,]$$

Once you have the above, you can write y = x+dx so that (x-y) = -dx in the denominator, and the numerator becomes $[\, p_{n+1}(x)p_n(x+dx) - p_{n+1}(x+dx)p_n(x)\,] = dx\,[\, p_{n+1}(x)p_n'(x) - p'_{n+1}(x)p_n(x)\,]$, the leading terms canceled, and we get the second term first due to the -dx downstairs, so result is

$$\sum_{i=0}^{n} (1/h_i)\,[p_i(x)]^2 = (k_n/k_{n+1})\,(1/h_n)\,[\, p_n(x)\, p'_{n+1}(x) - p_n'(x)\, p_{n+1}(x)\,]$$

Notice that the upper limit of the sum on the left matches the ortho poly labels on the right. So this is some kind of fancy "sum rule" that will no doubt get used a lot.

Proof of the first claim above: Start with the recursion formula which is this:

$$p_{n+1} = p_n\,(B_n + A_n\,x) - C_n p_{n-1}$$

Isolate the xp_n term to get

$$A_n\, x\, p_n = p_{n+1} - B_n p_n + C_n p_{n-1}$$

Now write this explicitly for x and y:

$$A_n\, x\, p_n(x) = p_{n+1}(x) - B_n p_n(x) + C_n p_{n-1}(x)$$

$$A_n\, y\, p_n(y) = p_{n+1}(y) - B_n p_n(y) + C_n p_{n-1}(y)$$

Multiply first by $p_n(y)$ and second by $p_n(x)$

$$A_n\, x\, p_n(x)\, p_n(y) = p_{n+1}(x)\, p_n(y) - B_n p_n(x)\, p_n(y) + C_n p_{n-1}(x)\, p_n(y)$$
$$A_n\, y\, p_n(x) p_n(y) = p_{n+1}(y) p_n(x) - B_n p_n(y) p_n(x) + C_n p_{n-1}(y) p_n(x)$$

Subtract

$$A_n\,(x-y)\, p_n(x)\, p_n(y) = [\, p_{n+1}(x)\, p_n(y) + C_n p_{n-1}(x)\, p_n(y)] - [\, p_{n+1}(y) p_n(x) + C_n p_{n-1}(y) p_n(x)]$$

where the two middle terms canceled. Regroup terms on the right to get

$$A_n\,(x-y)\, p_n(x)\, p_n(y) = [\, p_{n+1}(x)\, p_n(y) - p_{n+1}(y) p_n(x)] - C_n[\, p_{n-1}(x)\, p_n(y) - p_{n-1}(y) p_n(x)]$$

Reorder factors on the right to put n first,

$$A_n\,(x-y)\, p_n(x)\, p_n(y) = [\, p_{n+1}(x)\, p_n(y) - p_{n+1}(y) p_n(x)] - C_n[\, p_n(y)\, p_{n-1}(x) - p_n(x)\, p_{n-1}(y)]$$

We want $1/h_n$ on the left, so multiply both sides by $1/(h_n A_n)$

$$(1/h_n)\,(x-y)\, p_n(x)\, p_n(y)$$

$$= [\, p_{n+1}(x)\, p_n(y) - p_{n+1}(y) p_n(x)]/(h_n A_n) - C_n[\, p_n(y)\, p_{n-1}(x) - p_n(x)\, p_{n-1}(y)]/(h_n A_n)$$

Now compute: $C_n/(h_n A_n) = 1/(h_{n-1} A_{n-1})$. Then we have

$$(1/h_n)\,(x-y)\, p_n(x)\, p_n(y)$$

$$= [\, p_{n+1}(x)\, p_n(y) - p_{n+1}(y) p_n(x)]/(h_n A_n) - [\, p_n(y)\, p_{n-1}(x) - p_n(x)\, p_{n-1}(y)]/(h_{n-1} A_{n-1})$$

But write this as

$$(1/h_n)\,(x-y)\, p_n(x)\, p_n(y) = F_{n+1} - F_n \text{ where } F_{n+1} = [\, p_{n+1}(x)\, p_n(y) - p_{n+1}(y) p_n(x)]/(h_n A_n)$$

Notice that $F_0 = [\, p_0(x)\, p_{-1}(y) - p_0(y) p_{-1}(x)]/(h_0 A_0) = 0$ since $p_{-1} = 0$ as discussed above.

Now change index to i to be consistent with theorem:

$$(1/h_i)\,(x-y)\, p_i(x)\, p_i(y) = F_{i+1} - F_i$$

Sum both sides i = 1 to n to get a simple telescoping result,

$$\sum_{i=0}^{n} (1/h_i)\,(x-y)\, p_i(x)\, p_i(y) = F_{n+1} - F_0 = F_{n+1}$$

Move the (x-y) to the right to get

$$\sum_{i=0}^{n} (1/h_i)\, p_i(x)\, p_i(y) = (1/(x-y))\, [\, p_{n+1}(x)\, p_n(y) - p_{n+1}(y) p_n(x)]/(h_n A_n)$$

Finally, compute $1/(h_n A_n) = k_n/(h_n k_{n+1})$ to get

$$\sum_{i=0}^{n} (1/h_i)\, p_i(x)\, p_i(y) = (k_n/k_{n+1})\, (1/h_n)(1/(x-y))\, [\, p_{n+1}(x)\, p_n(y) - p_{n+1}(y) p_n(x)]$$

and this is our claimed result. No rocket science, just lots of fiddling. Hildebrand p 342 was my guide for this proof. QED.

The rest of this section I don't think is useful to me right now, so we move on.

1.4. Mechanical Quadrature (meaning Gaussian Quadrature).

In this section, we have statements of facts that I am now very familiar with from my Scheid reading. The "weights" which Scheid calls A_i are here called $\lambda_{\upsilon n}$ where the first index is like i, and the second reminds you that you are fitting n points in your interval (a,b). All expressions from (1) through (6) on page 161 apply to an arbitrary set of x-values x_υ (as they are called here). However, it is noted in mid-page that the exactness of the fit J to the integral I greatly improves if you select the x_υ as the zeros of $p_n(x)$. The Gaussian quadrature weights $\lambda_{\upsilon n} = A_\upsilon$ are here called **the Christoffel numbers**.

The results (6) and (7) are only true when we select the x_υ as the zeros of $p_n(\upsilon)$, because then we can write $p_n(x) = \text{constant}_n * \pi_n(x)$. By the way, we know that $p_n(x)$ has n zeros in the interval (a,b) and we thus know the zeros are all real, so $p_n(x)$ must factor completely over R. These two (6) and (7) results appear in Scheid page 129 where the integrands are really just L_i and L_i^2 . It is interesting that both integrals are the same. (Of course in Scheid they contain π instead of p).

The Big Result of this section, however, is equation (8) for the weights, again only true when x_υ are the zeros of $p_n(x)$. You very rarely see this formula stated in the general case, but here it is!

Theorem 4.1: (weight formula) The Gaussian quadrature weights are given by

$$\lambda_{\upsilon n} = -\, h_n\, (k_{n+1}/k_n) / [\, p_n'(x_\upsilon)\, p_{n+1}(x_\upsilon)\,] = +\, h_{n-1} (k_n/k_{n-1})\, (1/[p_n'(x_\upsilon) p_{n-1}(x_\upsilon)])$$

Remember that $h_n = (p_n, p_n)$ = the normalization factor, and k_n is the coefficient of x^n in $p_n(x)$. The second form does not appear in Erdelyi, but I derive it below.

Proof: You can see that the leading factors on the right are the same as appear in the C-D formula, so we expect that to be used in the proof! To start off, write the C-D formula setting $y = x_\upsilon$.

$$\sum_{i=0}^{n} (1/h_i)\, p_i(x)\, p_i(x_\upsilon) = (k_n/k_{n+1})\,(1/h_n)(1/(x-x_\upsilon))\,[\, p_{n+1}(x)\, p_n(x_\upsilon) - p_{n+1}(x_\upsilon)p_n(x)]$$

$$= -(k_n/k_{n+1})\,(1/h_n)(1/(x-x_\upsilon))p_{n+1}(x_\upsilon)p_n(x)$$

because the first term on the right vanishes since $p_n(x_\upsilon)= 0$. We next apply $\int w(x)dx$ to both sides. We can think of this as $(1/k_0) \int dx\, w(x)\, p_0(x)dx$. When we apply this to the LHS, all terms except i=0 vanish because $(p_i, p_0) = 0$. For i=0, the term is $(p_0, p_0)/h_0 = 1$. So we get this result:

$$1 = -(k_n/k_{n+1})\,(1/h_n)p_{n+1}(x_\upsilon) \int dx\, w(x)\, (1/(x-x_\upsilon))p_n(x)$$

so we have by this means evaluated the (certainly non-obvious) integral on the right to be:

$$\int dx\, w(x)\, (1/(x-x_\upsilon))p_n(x) = -[\, h_n/p_{n+1}(x_\upsilon)\,]\;(k_{n+1}/k_n)$$

and this is true for all x_υ which are zeros of $p_n(x)$. This is the result we now need because we have from (7) that

$$\lambda_{\upsilon n} = \int dx\, w(x)\, L_\upsilon(x) = (1/p_n'(x_\upsilon)) \int dx\, w(x)\, (1/(x-x_\upsilon))p_n(x)$$

and the integral here is just the one evaluated, so we have

$$\lambda_{\upsilon n} = -(1/[p_n'(x_\upsilon)p_{n+1}(x_\upsilon)])h_n\,(k_{n+1}/k_n)$$

QED. Now, from our recursion formula on page 158 we have, when x_υ = a zero,

$$p_{n+1}(x_\upsilon) = -C_n\, p_{n-1}(x_\upsilon)$$

since the middle term vanishes. We can then rewrite our formula as:

$$\lambda_{\upsilon n} = +(1/[p_n'(x_\upsilon)p_{n-1}(x_\upsilon)])(h_n/C_n)\,(k_{n+1}/k_n)$$

but now compute $(h_n/C_n)\,(k_{n+1}/k_n) = h_{n-1}(k_n/k_{n-1})$, so we get this simpler result

$$\lambda_{\upsilon n} = h_{n-1}(k_n/k_{n-1})\,(1/[p_n'(x_\upsilon)p_{n-1}(x_\upsilon)])$$

Comment: for the Classical Poly cases, we get formulas for $p_n'(x)$ which let us replace $p_n'(x_\upsilon)$ with $p_{n-1}(x_\upsilon)$ times other factors, and then we get a formula for the above where we don't need any derivatives of $p_n(x)$ in the weight formula. However, I don't think there is a simple general formula, but we shall see later!

We now skip the rest of this section.

10.5 Continued Fractions

I don't even understand the very first equation (1)'s notation with the vertical bar, nor do I care about continued fractions right now, so skip this whole section.

10.6 The "Classical" Orthogonal Polynomials.

Start by *assuming* the R formula (1) on page 164 and see what we get. Note that X(x) = degree k. We find that the only possibilities are k = 0,1,2 as he now shows.

Case 0: [k = 0 => Hermite polynomials.] Suppose $X = X_0$ a constant, meaning k = 0. Assume that $p_1(x) = k_1x + k_1'$. Then (2) tells us that

$$K_1 (k_1x + k_1') = X_0 w'(x)/w(x) \quad => \quad w'(x)/w(x) = (K_1/X_0) (k_1x + k_1')$$

Now define a change of variable $y = \alpha x + \beta$ so that $dy = \alpha dx$. Then define

$$W(y) = w(x) \quad => \quad W'(y) = dW/dy = dw/dy = dw/dx * dx/dy = w'(x) (1/\alpha).$$

Then we have $W'(y)/W(y) = w'(x)/w(x) * (1/\alpha) = (K_1/\alpha X_0) (k_1x + k_1') = -2y$

We want to select α and β so that this last expression is in fact -2y. Thus, we have

$$y = -(K_1/2\alpha X_0) (k_1x + k_1')$$

from which we can read off that $\alpha = -(k_1K_1/2\alpha X_0)$ and $\beta = -(K_1/2\alpha X_0) k_1'$. The first equation says that

$$\alpha^2 = -(k_1K_1/2X_0) \quad => \quad \alpha = i \text{ sqrt}(k_1K_1/2X_0) \quad \text{and then } \beta = \text{whatever it is.}$$

The point is that we have change variables from x to y, and we end up with $W'(y)/W(y) = -2y$ and this then tells us $dW/W = -2ydy$ and then $\ln W(y) = -y^2$ and then $W(y) = \exp(-y^2)$.

Our change of variables is linear but a bit ugly, but it does seem to lead to the Hermite weight. We then write everything in terms of the y variable, including the Rodriguez formula. You can see that for the Hermite's, the R formula is exact as Erdelyi gives it on page 164, but we write it in y, and we are done if we set X = 1 and $w(y) = \exp(-y^2)$ and call the functions $p_n(y) = H_n(y)$. This shows up on page 785 of A&S, where they use α_n for Erdelyi's K_n. Of course A&S nail down the value of α_n.

Case 1: [k = 1 => Laguerre polynomials.] Here we have to do some kind of more complex variable change (I don't think it is just linear as he says), to get to a form $W'(y)/W(y) = -1 + \alpha/y$.

Here from (2) we get (3) and we can then write

$$w'(x)/w(x) = (Ax+B)/(\alpha x+\beta) = c [- 1+ d/(x+e)] \text{ with apropo constants.}$$

As our first change of variables, let z = fx so dz = f dx and then $W'(z)/W(z) = [w'(x)/w(x)] (1/f)$. Then we get $W'(z)/W(z) = (1/f) c [- 1+ d/((z/f)+e)]$. So choose f = c and get $W'(z)/W(z) = -1 + g/(z+h)$. Now do

a second change of variables from z to y of the form y = z+h and get W'(y)/W(y) = -1 + g/y, and this is the desired form, but he writes it W'(y)/W(y) = -1 + α/y. If you solve this equation dW/W = (-1 + α/y)dy by integrating to lnW = -y + αlny then expo to get W = exp(-y)exp(αlny) = exp(-y) y^{α} and we have then arrived at the Laguerre poly world. Without loss of generality, you can pick X(y) = y. As before, Rodriguez comes out exactly as you think, see A&S page 785.

Case 2: [k = 2 => Hypergeometric (Jacobi) polynomials.]

I don't follow things here but accept the results. Looking at the table on page 164, we can see that the second line is just a special case of this case, and the first is a special case of the second line. Authors show on page 165 (but I don't follow) that if k > 2, you find that $p_2(x)$ is not of degree 2, so things are inconsistent, so the only allowed (and distinct) Rodriguez choices are X = 1, x and $(1-x)^2$.

10.7 General Properties of the classical ortho polys.

This section opens with this claim

Theorem 7.1: The quantity $D^k[w(x)X^n]$ vanishes at a and b for the three lower cases of the table on page 164 provided $\alpha > -1$ and $\beta > -1$ and $\alpha > -1$ in the Laguerre case, as long as k = 0,1,2...up to (n-1).

Proof: Let's first look at the Jacobi case. There we have $D^k [(1-x)^{\alpha+n} (1+x)^{\beta+n}]$ as our quantity of interest. There are many terms in this expansion, but each time D hits something like $(1-x)^{\alpha+n}$, it lowers the power by one. So the worse case is a term like $D^{n-1} (1-x)^{\alpha+n} = (\alpha+n)(\alpha+n-1)\ldots (1-x)^{\alpha+1}$. This vanishes at x = 1 as long as $\alpha > -1$, as they claim. Same idea on the β side for x = -1, so I accept the Theorem as true for this case.

What about the Hermite case? There we have $D^k [e^{-x*x}]$ which is just e^{-x*x} x poly, and this vanishes at the two endpoints which are infinite.

What about the Laguerre case? Here we have $D^k [x^{\alpha+n} e^{-x}]$. As in our first case, the worse case for x=0 occurs when we grind down the first exponent, but we will have $(x)^{\alpha+1}$ again, and we will be OK then for $\alpha > -1$. The infinite side will take care of itself since e^{-x} will be in every term.

QED.

Theorem 7.2: The set of polynomials defined by the Rodriguez formula forms an orthonormal system.

Proof: Using the above theorem, author shows that $(f,p_n) = 0$ for any poly f of degree < n, where p_n is the Rodriguez poly. But since all p_k for k < n are candidates for this f, we know that $(p_k,p_n) = 0$ k < n, and this means that we have an orthonormal system!

Comment: This is a very nice result. First we saw that we could only have k = 0,1,2 in the Rodriguez and we took standard forms for each case to get our three cases. Now we see that in each case, we get an orthogonal system on (a,b) with the w(x) and X shown in the table on page 164. This is of course only a small subset of the universe of possible orthogonal systems. One wonders if there is a more general R formula that can generate ALL orthogonal systems?

Lemma 7.3. Derive the 4th equation from the bottom of page 166 using Leibniz' Formula.

Proof: That formula is this:

$D^n(fg) = \Sigma_{i=0}^{n}\ (n,i)\ D^i f * D^{n-i} g$ where (n,i) is the binomial coefficient.

Our application is going to have one higher power, so write

$D^{n+1}(fg) = \Sigma_{i=0}^{n+1}\ (n+1,i)\ D^i f * D^{n+1-i} g$

Now we take f = X and g = $D(wX^n)$. Notice that $D^{n-1}g = D^n(wX^n) = K_n w p_n$ from R formula. Now in the sum, since K is at most degree 2, only the first three terms will survive. They are:

$(n+1,0)\ X\ D^{n+1}g\ + (n+1,1)\ X'\ D^n g\ + (n+1,2)\ X''\ D^{n-1}g$

$= [\ 1*X*D^2\ + (n+1)*X'\ D + n(n+1)/2 * X''\ ]\ (D^{n-1}g)$

$= [\ X\ D^2\ + (n+1)\ X'\ D + n(n+1)/2\ X''\ ]\ (K_n w p_n)$

and this is the claimed result, QED.

Lemma 7.4. Derive the 3rd equation from the bottom of page 166.

Proof: This equation has two lines. The first line is proven if we can show that

$XD(wX^n) = (wX^n)[\ K_1 p_1 + (n-1)X'\ ]$

First, do the left side to get

$XD(wX^n) = X\ [\ wnX^{n-1} + w'\ X^n\ ] = wnX^n X' + w'\ X^{n+1} = wX^n\ [\ nX' + (w'/w)X\ ]$

Now use 10.6.3 which says $(w'/w) = (K_1 p_1 - X')/X$ to get

$XD(wX^n) = wX^n\ [\ nX' + (K_1 p_1 - X')\ ] = (wX^n)\ [\ K_1 p_1 + (n-1)X'\ ]$ QED.

Now go back to our Leibniz formula with n+1

$D^{n+1}(fg) = \Sigma_{i=0}^{n+1}\ (n+1,i)\ D^i f * D^{n+1-i} g$

and this time take f = $[\ K_1 p_1 + (n-1)X'\ ]$ and g = (wX^n). Since f is of degree 1 at most, only the first two terms survive and we get

$(n+1,0)\ f\ D^{n+1}g\ + (n+1,1)\ Df\ D^n g = [\ 1*f*\ D\ + (n+1)*Df\ ]\ D^n g$

$= \{\ f\ D + (n+1)\ (Df)\ \}\ (K_n w p_n)$

$= \{\ [\ K_1 p_1 + (n-1)X'\ ]\ D\ + (n+1)\ [\ K_1 p_1' + (n-1)X''\ ]\ \}\ (K_n w p_n)$

and this is the second line QED.

****

Theorem 7.3. Show that the polynomial $p_n(x)$ satisfies the ODE shown as (1) on bottom of page 166 where λ_n is given by (2) in which k_1 is the leading coefficient of $p_1(x)$, that is, $p_1(x) = k_1x + k_1'$.

Proof: It all works, but a HUGE amount of algebra is required! I don't know why it is so messy. Perhaps another day I can let Maple do this.

STEP 1. Compare results of previous equations as they say to get:

$$[\, X D^2 + (n+1)\, X' D + n(n+1)/2\, X'' \,] (K_n w p_n)$$
$$= \{\, [\, K_1 p_1 + (n-1)X' \,] D + (n+1) \, [\, K_1 p_1' + (n-1)X'' \,] \,\} (K_n w p_n)$$

Cancel the K_n and gather like terms to the left side:

$$[\, X D^2 + \{\, (n+1)\, X' - [\, K_1 p_1 + (n-1)X' \,] \,\} D + \{\, n(n+1)/2\, X'' - (n+1) \, [\, K_1 p_1' + (n-1)X'' \,] \,\} \,] (w p_n) = 0$$

Simplify some factors

$$\{\, (n+1)\, X' - [\, K_1 p_1 + (n-1)X' \,] \,\} = \{\, (n+1)\, X' - K_1 p_1 - (n-1)X' \,] \,\}$$
$$= \{\, 2X' - K_1 p_1 \}$$

$$\{\, n(n+1)/2\, X'' - (n+1) \, [\, K_1 p_1' + (n-1)X'' \,] \,\} = (n+1)\{\, n/2\, X'' - [\, K_1 p_1' + (n-1)X'' \,] \,\}$$

$$= (n+1)\{\, n/2\, X'' - K_1 p_1' - (n-1)X'' \,\} = (n+1)/2\{\, n\, X'' - 2K_1 p_1' - 2(n-1)X'' \,\}$$
$$= (n+1)/2\{\, (-n+2)\, X'' - 2K_1 p_1' \,\}$$
$$= -(n+1)\{\, (n-2)\, X'' + 2K_1 p_1' \,\}/2$$

Therefore we get

$$[\, X D^2 + \{2X' - K_1 p_1 \} D - (n+1) \, \{\, (n-2)/2\, X'' + K_1 p_1' \,\} \,] (w p_n) = 0$$

STEP 2: This is still a pretty messy result! Now do this:

$$D(w p_n) = w' p_n + w D p_n$$

$$D^2(w p_n) = w'' p_n + w' D p_n + w D^2 p_n + w' D p_n = w'' p_n + 2w' D p_n + w D^2 p_n$$

Jam this in and get

$$X\{\, w'' p_n + 2w' D p_n + w D^2 p_n \} + \{2X' - K_1 p_1 \} \{w' p_n + w D p_n \} - (n+1)\{\, (n-2)/2\, X'' + K_1 p_1' \,\} w p_n = 0$$

Now group terms:

$XwD^2p_n + [\ 2Xw' + \{2X' - K_1p_1\}w]\ Dp_n + [\ Xw'' + \{2X' - K_1p_1\}w' - (n+1)\{\ (n-2)/2\ X'' + K_1p_1'\}w]p_n = 0$

Now divide by w to make the first term match:

$XD^2p_n + [\ 2Xw'/w + \{2X' - K_1p_1\}]\ Dp_n$
$+ [\ Xw''/w + \{2X' - K_1p_1\}w'/w - (n+1)\{\ (n-2)/2\ X'' + K_1p_1'\}]p_n = 0$

STEP 3: The first term now has the right form. To verify the second term, we need to show that:

$[\ 2Xw'/w + \{2X' - K_1p_1\}] = K_1p_1$

We can use (3) to replace $2Xw'/w = 2(K_1p_1 - X')$ and then we have

$LHS = [\ 2(K_1p_1 - X') + \{2X' - K_1p_1\}]\ = K_1p_1$ hurray!

STEP 4: Now it only remains to show that the third term is right. It's factor is this

$[\ Xw''/w + \{2X' - K_1p_1\}w'/w - (n+1)\{\ (n-2)/2\ X'' + K_1p_1'\}]\ = LHS$

Let's differentiate (2) page 164 to get

$K_1p_1' = X'' + [\ -X(w'/w)^2 + Xw''/w + X'w'/w\ ]$

Put K_1p_1' in for the last factor and we get

$LHS = [\ Xw''/w + \{2X' - K_1p_1\}w'/w - (n+1)\{\ (n-2)/2\ X'' + (X'' + [\ -X(w'/w)^2 + Xw''/w + X'w'/w\ ])\}]$

$= [\ Xw''/w + \{2X' - K_1p_1\}w'/w - (n+1)\{\ (n-2)/2\ X'' + X'' - X(w'/w)^2 + Xw''/w + X'w'/w\}]$

Now replace $K_1p_1 = X' + Xw'/w$ in the first term to get $\{2X' - K_1p_1\} = X' - Xw'/w$, then we have

$LHS = [\ Xw''/w + \{X' - Xw'/w\}w'/w - (n+1)\{\ (n-2)/2\ X'' + X'' - X(w'/w)^2 + Xw''/w + X'w'/w\}]$

If we now group factors like (w''/w) and so on [I did this on scratch paper] we simplify to:

$LHS = -n\ \{\ X(w''/w) + X'(w'/w) - X(w'/w)^2 + (n+1)X''/2\ \}$

Looking now at the predicted answer: we have to show that

$-n\ \{\ X(w''/w) + X'(w'/w) - X(w'/w)^2 + (n+1)X''/2\ \} = -n\ \{\ k_1K_1 + (n-1)/2X''\ \}$

which is to say, we must show that

$X(w''/w) + X'(w'/w) - X(w'/w)^2 + (n+1)X''/2 = k_1K_1 + (n-1)/2X''$

or we must show that

$$X(w''/w) + X'(w'/w) - X(w'/w)^2 + X'' = k_1K_1$$

where I presume k_1 is the leading coefficient in p_1 (the only way we have ever used this symbol). Now suppose we write $p_1(x) = k_1x + k_1'$, then the R formula says

$$k_1x + k_1' = (1/K_1w)\, D[wX]$$

Differentiate both sides then to get:

$$K_1k_1 = D\{ w^{-1} D[wX] \}$$

$$= D\{ X' + (w'/w)X \} = X'' + X'(w'/w) + (w''/w)X - (w'/w)^2X$$

But this is exactly what we want, so we are finally done!!! QED.

Comments: In the first part of this chapter, we talk about orthogonal polynomials in general. There is no differential equation at all. Only when we start talking about the restricted class of ortho polys given by the Rodriguez formula do we end up with an ODE. Deriving the form of the ODE is non-trivial, but we have done it in full detail.

****

Lemma 7.5. The expression $Xp_n' - (n/2)X'' \times p_n = f$ where poly f is of degree $\leq n$.

Proof:
- For X = zeroth order = constant (k=0), we have p_n' = degree n-1 and X'' = 0, so deg(f) = n-1.
- For X = first order, Xp_n' has a leading power of x^n and again X'' = 0, so deg(f) = n.
- For X = second order, things are a little more complicated. Assume that $X = ax^2 + o(x)$, so that X'' = 2a. Our expression is then $[ax^2 + o(x)]\ p_n' - na \times p_n$. If we write $p_n = k_nx^n$ + lower, then we have the leading power of the first term being $ax^2 * nk_n\ x^{n-1} = ank_nx^{n+1}$. But the second term has $-na\ k_nx^{n+1}$ so these two terms cancel and the leading term must then be order p^n. Notice that our Lemma is true for any quadratic form of X, such as $X = (1-x)^2$ or $1 - x^2$.

This shows that the lemma is true in all three cases k = 0,1 and 2 (where k is the power of X in Rodriguez). Notice

****

Lemma 7.6. Show that $I = (Xp_n',\ p_{n-k}) = 0$ where k = 2,3,4...n.

Proof: Write this out as $\int dx\ w\ p_{n-k}\ X\ (Dp_n)$ and do parts once, parts vanish hopefully with same conditions as in the parts done at the top of page 66. This gives

$$I = \int dx\ w\ p_{n-k}\ X\ (Dp_n) = -\int dx\ D(wX\ p_{n-k})p_n$$

Now write $D(wX\ p_{n-k}) = [\ D(wX)\ p_{n-k} + (wX)\ p_{n-k}'\ ]$ so we how have

$$I = -\int dx \ [\ D(wX)\ p_{n-k} + (wX)\ \ p_{n-k}'\ ]\ p_n$$

The second term has the form $(Xp_{n-k}'\ ,\ p_n)$. The left factor has degree n-k -1+2 = n -k+1 at most. Since we are dealing here only with $k \geq 2$, left factor has degree n-1 at most. This scalar product thus vanishes since we can write this factor as a sum of p_n of degree n-1 max. We are then left with

$$I = -\int dx \ [\ D(wX)\ p_{n-k}\ ]\ p_n$$

We can now write $D(wX) = K_1\ w\ p_1$ from Rod, so we have

$$I = -\ K_1 \int dx\ w\ p_1\ p_{n-k}\ p_n \ \ = \ -K_1\ (p_1\ p_{n-k},\ p_n)$$

Now the product $p_1\ p_{n-k}$ is at most of degree n-k+1 which is the same situation as before with the "left factor" and by the same argument we get 0 for this scalar product. Therefore I = 0, QED.

****

Theorem 7.4. Derive the differentiation formula stated on page 167 as (4).

Proof: Note: This proof assumes only that X is a quadratic with leading term $X = ax^2$.

STEP 1: (Show $\gamma_i = 0$) Start as suggested and we then have that

$$Xp_n' - (n/2)X''\ x\ p_n \ = f(x) = \alpha_n p_n + \beta_n p_{n-1} + \ \text{sum}(i=2,n)\ \gamma_i\ p_{n-i}$$

We get to this point because we know f(x) is degree n or less, and we know we can expand any such poly in terms of the full set of p_n as shown. The coefficients are given strange names, fine. So right off the bat we have the following version of our formula:

$$Xp_n' \ = [\alpha_n + \ (n/2)X''\ x\ ]\ p_n \ + \beta_n p_{n-1} + \ \text{sum}(i=2,n)\ \gamma_i\ p_{n-i}$$

So our task is to evaluate α_n and β_n and to show that the $\gamma_i = 0$. Let's first show that $\gamma_2 = 0$. Close both sides of second last equation with (,p_{n-2}) to get

$$(\ Xp_n' - (n/2)X''\ x\ p_n,\ p_{n-2}) \ = \ \gamma_2\ (p_{n-2},\ p_{n-2})$$

Write the LHS as

$$(\ Xp_n',\ \ p_{n-2}\ ) - (n/2)X''\ (x\ p_n,\ p_{n-2})$$

but we know that $(x\ p_n,\ p_{n-2}) = (p_n,\ x\ p_{n-2}) = 0$ because xp_{n-2} is of degree n-1 and can be expanded in terms of p_n for n-1 and below, all of which are orthogonal to p_n. So we are then left with

$$(\ Xp_n',\ \ p_{n-2}\ ) = \ \gamma_2\ (p_{n-2},\ p_{n-2})$$

and we want to show that the LHS = 0. But this is what our Lemma 7.6 above shows, so we have $\gamma_2 = 0$. The same discussion causes all the $\gamma_k = 0$ for $k \geq 2$. The thing we get is in that case

$$(Xp_n',\ p_{n-k}) = \gamma_k (p_{n-k,}\, p_{n-k}) \quad k \geq 2$$

and our Lemma applies to all these cases at once, so $\gamma_k = 0$. We then have this result so far:

$$Xp_n' = [\alpha_n + (n/2)X'' x\ ]\ p_n + \beta_n p_{n-1}$$

which shows the essential fact that the derivative mixes p_n with one adjacent poly. We now have to isolate the two parameters and evaluate them!

STEP 2: (Solve for β_n) If we close with p_{n-1} we get

$$(Xp_n', p_{n-1}) = (n/2)X'' (xp_n, p_{n-1}) + \beta_n(p_{n-1}, p_{n-1})$$

and we have something we might solve for β_n. We can evaluate this:

$(xp_n, p_{n-1}) = (p_n, xp_{n-1}) = (p_n, x\, k_{n-1}x^{n-1} + o(x^{n-1}))$ where we can drop lower terms in the expansion

$= k_{n-1} (p_n, x^n) = k_{n-1} (p_n, (1/k_n)p_n + o(x^{n-1}))$ and again drop extra terms
$= (k_{n-1}/k_n)\ (p_{n,}\, p_n)$ OK

So we then have

$$I = (Xp_n', p_{n-1}) = (n/2)X'' (k_{n-1}/k_n)\ (p_{n,}\, p_n) + \beta_n(p_{n-1}, p_{n-1})$$

and we have now to deal with the LHS. Following the start of our Lemma above, we can do parts to get

$$- I = \int dx\ [\ D(wX)\, p_{n-1} + (wX)\ p_{n-1}'\ ]\ p_n = I_1 + I_2$$

Use the Rod formula $D(wX) = K_1\, w\, p_1$ in the first term to get

$I_1 = K_1 (p_1, p_{n-1}p_n) = K_1 (p_n, p_{n-1}p_1) = K_1 (p_n, p_{n-1} [k_1x + o(x^0)]\ ) = K_1k_1(p_n, xp_{n-1})$
$= K_1k_1 (xp_n, p_{n-1})$

but we just evaluated this bracket above so we can go on to say

$$I_1 = K_1k_1 (k_{n-1}/k_n)\ (p_{n,}\, p_n)$$

Now consider

$$I_2 = (Xp_{n-1}' ,p_n) = (X\ \{(n-1)k_{n-1}\, x^{n-2} + o(x^{n-3})\}, p_n) = (n-1)\, k_{n-1}\, (X\, x^{n-2}, p_n)$$

The rightmost bracket is nonzero only if X is the k=2 case. In that case we have

$$(X\, x^{n-2}, p_n) = (\, [\, ax^2 + o(x)]\, x^{n-2}, p_n) = a(x^n, p_n) = a(\, p_n/k_n + o(x^{n-1})\, , p_n) = (a/k_n)(p_n, p_n)$$

Therefore we have shown that $I_2 = a(n-1)(k_{n-1}/k_n)(p_n, p_n)$ for this one case, and it is zero for the lower cases. Thus, since X" = 0,0,2a for our 3 cases, we can add a factor (X"/2a) and say
$I_2 = (n-1)(k_{n-1}/k_n)(p_n, p_n)(X"/2)$ and we have then covered all three cases. So we have now learned that

$$I = (Xp_n', p_{n-1}) = -I_1 - I_2 = -\ K_1k_1\,(k_{n-1}/k_n)\ (p_n, p_n)\ -(n-1)(k_{n-1}/k_n)(p_n, p_n)\ (+\ X"/2)$$

$$=\ -\,(k_{n-1}/k_n)\,[\ K_1k_1 + (n-1)(X"/2)\ ]\ (p_n, p_n)$$

but from above we also have that $I = (n/2)X"\ (k_{n-1}/k_n)\ (p_n, p_n) + \beta_n(p_{n-1}, p_{n-1})$. Therefore:

$$-\,(k_{n-1}/k_n)\,[\ K_1k_1 + (n-1)(X"/2)\ ]\ (p_n, p_n) = (n/2)X"\ (k_{n-1}/k_n)\ (p_n, p_n) + \beta_n(p_{n-1}, p_{n-1})$$

which we can write as

$$-\,(k_{n-1}/k_n)\,[\ K_1k_1 + (n-1)(X"/2)\ ]\ h_n = n\ (X"/2)\ (k_{n-1}/k_n)\ h_n + \beta_n h_{n-1}$$

and we can group the h_n terms to get

$$\beta_n h_{n-1}\ = -[K_1k_1 + (2n-1)(X"/2)\ ]\ h_n\ (k_{n-1}/k_n)$$

and finally

$$\beta_n\ = -\,(h_n/h_{n-1})\ (k_{n-1}/k_n)\ [K_1k_1 + (2n-1)(X"/2)\ ] \qquad \text{// my result}$$

We can now compare this with Erdelyi's formula for β_n

$$\beta_n\ = -(C_n/A_n)[\ k_1K_1 + (n - 1/2)X"\ ]$$

but I know that $(C_n/A_n) = h_n/(A_{n-1}h_{n-1}) = (h_n/h_{n-1})(k_{n-1}/k_n)$, so his answer then becomes

$$\beta_n\ = -(h_n/h_{n-1})(k_{n-1}/k_n)[\ k_1K_1 + (2n-1)(X"/2)] \qquad \text{/ his}$$

which agrees with my result, so happy now about β_n.

STEP 3: (Find β_n for Jacobi polys) As an application, let's look at his Jacobi differential formula on p 170 top

$$(2n+\alpha+\beta)X\ p_n'\ = \text{first term} +\ \ 2(n+\alpha)(n+\beta)p_{n-1}$$

and our general formula said,

$$Xp_n' = [\alpha_n + (n/2)X'' x]\, p_n + \beta_n p_{n-1}$$

Let's divide his through to get

$$X\, p_n' = \text{first term}' + 2(n+\alpha)(n+\beta)/(2n+\alpha+\beta)p_{n-1}$$

so this would suggest that he uses $\beta_n = 2(n+\alpha)(n+\beta)/(2n+\alpha+\beta)$. This result is explicitly stated in his constants section, good!

Let's now look at our now-derived formula for β_n

$$\beta_n = - (h_n/h_{n-1})(k_{n-1}/k_n)[\, k_1K_1 + (2n-1)(X''/2) \,]$$

The constants on page 169 are just a huge mess! I just computed some things,

$$(h_n/h_{n-1}) = (1/n)\,\frac{(2n+\alpha+\beta-1)}{(2n+\alpha+\beta+1)}\,\frac{(n+\alpha)(n+\beta)}{(n+\alpha+\beta)}$$

$$(k_{n-1}/k_n) = 2n\,\frac{(n+\alpha+\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta-1)}$$

The product of these two things is then,

$$(h_n/h_{n-1})(k_{n-1}/k_n) = 2\,\frac{(n+\alpha)(n+\beta)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta)}$$

And we have $k_1 = (1/2)(2+\alpha+\beta)$ and $K_1 = -2$ so that $k_1K_1 = -(2+\alpha+\beta)$. Then we get

$$\beta_n = - 2\,\frac{(n+\alpha)(n+\beta)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta)}\,[\, -(2+\alpha+\beta) + (2n-1)(X''/2) \,]$$

If we now assume (for the first time ever) that $X = 1 - x^2$ so $X'' = -2$, this becomes

$$\beta_n = - 2\,\frac{(n+\alpha)(n+\beta)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta)}\,[\, -(2+\alpha+\beta) - (2n-1)]$$

$$= 2\,\frac{(n+\alpha)(n+\beta)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta)}\,[\, (2+\alpha+\beta) + (2n-1)]$$

$$= 2\,\frac{(n+\alpha)(n+\beta)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta)}\,[\, (2n+\alpha+\beta+1)]$$

$$= 2\,\frac{(n+\alpha)(n+\beta)}{(2n+\alpha+\beta)} = \text{the right answer.}$$

STEP 4: (Compute α_n). First here are three little Lemmas we will be using below:

Lemma S1 $(x^n, p_n) = (1/k_n)h_n$

Proof: $(x^n, p_n) = (p_n/k_n + o(x^{n-1}) , p_n) = (1/k_n)(p_n, p_n)$ QED.

Lemma S2 $(p_n, x^{n+1}) = -r_{n+1} (p_n, x^n) = -r_{n+1} (1/k_n)h_n$

Proof: Write $p_{n+1}/k_{n+1} = x^{n+1} + (k'_{n+1}/k_{n+1}) x^n + o(x^{n-1})$ and solve for x^{n+1}

$$(p_n, x^{n+1}) = (p_n, p_{n+1}/k_{n+1}) - (k'_{n+1}/k_{n+1})(p_n, x^n) - \text{terms like } (p_n, x^{n-1}) \text{ and lower}$$
$$= - (k'_{n+1}/k_{n+1})(p_n, x^n) = - r_{n+1}(p_n, x^n)$$

which gives the first result. Then use lemma (1)'s result to get the final result.

Lemma S3 $(p_n, xp_n) = k_n[r_n - r_{n+1}](p_n, x^n) = (r_n - r_{n+1}) h_n$

Proof: Start with

$$(p_n, xp_n) = (p_n, x [k_n x^n + k_n' x^{n-1} + o(x^{n-2})]) = k_n (p_n, x^{n+1}) + k_n' (p_n, x^n)$$

Use lemma 2 to get

$$= k_n \{-r_{n+1} (p_n, x^n)\} + k_n' (p_n, x^n) = [k_n' - k_n r_{n+1}] (p_n, x^n)$$

$$= k_n[r_n - r_{n+1}](p_n, x^n) = k_n[r_n - r_{n+1}](1/k_n)h_n = [r_n - r_{n+1}] h_n$$

where we used lemma 1 for (p_n, x^n). QED.

(Compute α_n) From the end of STEP 1 we quote this result:

$$Xp_n' = [\alpha_n + (n/2)X'' x] p_n + \beta_n p_{n-1}$$

and we have already found β_n. Let's now close with p_n to get:

$$(p_n, Xp_n') = \alpha_n (p_n, p_n) + n(X''/2) (p_n, xp_n) = \alpha_n h_n + n(X''/2) (p_n, xp_n)$$

Let's start with (p_n, Xp_n'). If we were to do a parts integration at this point, we pick up constants like k_1 and K_1 just as we did in computing β_n . This method is reviewed in Appendix A. We end up with a complex result for α_n and we cannot even tell if it is right or wrong! So instead, to get the simple form result which appears in Erdelyi, we decide here to NOT do parts integration at this step, and instead we do this:

$$(p_n, Xp_n') = (p_n, [X(0) + x X'(0) + (x^2/2) X''] [nk_n x^{n-1} + (n-1)k'_n x^{n-2} + o(x^{n-3})])$$

$$= X(0) * 0 + X'(0) * nk_n (p_n, x^n) + (X''/2)*[nk_n (p_n, x^{n+1}) + (n-1)k'_n (p_n, x^n)]$$

$$= (X''/2)* nk_n (p_n, x^{n+1}) + [X'(0) * nk_n + (n-1)k'_n (X''/2)] (p_n, x^n)$$

Now let's use Lemma S2

$$(p_n, x^{n+1}) = - (k'_{n+1}/k_{n+1})(p_n, x^n) = -r_{n+1} (p_n, x^n)$$

so we then have

$$(p_n, Xp_n') = (X''/2)* nk_n \{ -r_{n+1} (p_n, x^n) \} + [X'(0) * nk_n + (n-1)k'_n (X''/2)] (p_n, x^n)$$

$$= [X'(0) * nk_n + (n-1)k'_n (X''/2) - r_{n+1} (X''/2)* nk_n](p_n, x^n)$$

$$= k_n [X'(0) * n + (n-1)r_n (X''/2) - r_{n+1} (X''/2)* n](p_n, x^n)$$

$$= k_n [X'(0) * n + (n-1)r_n (X''/2) - r_{n+1} (X''/2)* n](p_n, x^n)$$

From Lemma S1 we have

$$(x^n, p_n) = (p_n/k_n + o(x^{n-1}) , p_n) = (1/k_n)(p_n, p_n) = (1/k_n)h_n$$

so our *non-parts* version of (p_n, Xp_n') is

$$(p_n, Xp_n') = [X'(0) * n + (n-1)r_n (X''/2) - r_{n+1} (X''/2)* n] h_n \quad \text{// non-parts}$$

Meanwhile, the RHS is this

$$(p_n, Xp_n') = \alpha_n h_n + n(X''/2) (p_n, xp_n)$$

and from Lemma S3 we have

$$(p_n, xp_n) = (r_n - r_{n+1}) h_n$$

so we have for the RHS

$$(p_n, Xp_n') = \alpha_n h_n + n(X''/2) (r_n - r_{n+1}) h_n$$

Now here then is our total equation (in the non-parts path)

$$\alpha_n h_n + n(X''/2) (r_n - r_{n+1}) h_n = [X'(0) * n + (n-1)r_n (X''/2) - r_{n+1} (X''/2)* n] h_n$$

and we can cancel the h_n at once to get

$$\alpha_n + n(X''/2) (r_n - r_{n+1}) = [X'(0) * n + (n-1)r_n (X''/2) - r_{n+1} (X''/2)* n]$$

so we then have

$$\alpha_n = - n(X''/2) (r_n - r_{n+1}) + [X'(0) * n + (n-1) r_n (X''/2) - r_{n+1} (X''/2)* n]$$

$$= r_n \{ - n(X''/2) + (n-1) (X''/2) \} + r_{n+1} \{ n(X''/2) - (X''/2)* n \} + nX'(0)$$

$$= r_n(X''/2) \{ -n +n - 1\} + 0 + nX'(0)$$

$$= nX'(0) - r_n(X''/2)$$

which is the right short-form answer, as shown in 10.7.5.

Theorem 7.5. Derive the "other" formula for $X p_n'$.

Proof: Our starting point is the result on page 167 in equations (4) and (5) which we have derived in Theorem 7.4 previous. Now we use the recurrence formula on page 158

$$p_{n+1} = (A_n x + B_n) p_n - C_n p_{n-1}$$

and we solve for p_{n-1}

$$p_{n-1} = (x A_n/C_n + B_n/C_n) p_n - p_{n+1}/C_n$$

We then jam this into (4) on page 167 to get

$$Xp_n' = (\alpha_n + n(X''/2) x) p_n + \beta_n \{ (x A_n/C_n + B_n/C_n) p_n - p_{n+1}/C_n \}$$

$$= ([\alpha_n + \beta_n(B_n/C_n)] + [n(X''/2) + \beta_n(A_n/C_n)] x) p_n - (\beta_n/C_n) p_{n+1}$$

$$= (a + bx) p_n + d p_{n+1}$$

This result does appear more complicated, but not really sure. We prefer the other formula because we are happier having p_{n-1} instead of p_{n+1} in any formula!

Theorem 7.6. Derive the two formulas for the Gaussian Weights shown in page 167 equation (7).

Proof: These are (luckily) very easy to show. I am going to leave out the arguments of the p_n and X here but they are all x_υ , the zeros of $p_n(x_\upsilon)$. We start off with

$$\lambda_{\upsilon n} = - h_n A_n/[p_n' p_{n+1}] \qquad \text{as we already derived on page 162 (8)}$$

Our recursion relation says that $p_{n+1} = -C_n p_{n-1}$. So we can then have this alternative form

$$\lambda_{\upsilon n} = + h_n(A_n/C_n) / [p_n' p_{n-1}] = h_{n-1} A_{n-1} / [p_n' p_{n-1}]$$

Next, (4) says that $X p_n' = \beta_n p_{n-1}$. This then gives us

$\lambda_{\upsilon n} = h_{n-1} A_{n-1} (X/\beta_n) / [p_{n-1}]^2$ which is perhaps the best form, all eval at $x = x_\upsilon$

The other alternative is to instead replace p_{n-1} to get

$$\lambda_{\upsilon n} = h_{n-1} A_{n-1} (\beta_n/X) / [p_n']^2$$

and this concludes proof of formulas (7)

A few comments on the Differential Equation notations.

As presented in Erdelyi, we have in general

$$X p_n'' + K_1 p_1 p_n' + \lambda_n p_n = 0$$

and for the zeros of p_n we have

$$Xp_n' = \beta_n p_{n-1} \qquad \text{where } n\beta_n = \lambda_n(C_n/A_n) \qquad // \; x = x_v \text{ only}$$

Now how does this relate to A&S? From page 781, it certainly seems that they use

$$g_2 = X$$

since we only see the three values 1, x and $1-x^2$ in this entire table. In the Rodriquez, we see that they have written $g_2 = g$ and $e_n = K_n$. Then they go on to define these things

k_n and k_n' and h_n same as in Erdelyi

$b_n = k_{n+1}/k_n = A_n$

$a_n = b_n (r_{n+1} - r_n) = A_n (r_{n+1} - r_n) = B_n$

$c_n = C_n$

$g_2 = X$

$e_n = K_n$

$g_1 = K_1 p_1$

$g_0 = \lambda_n$ for the pure orthog poly cases in Table page 781
$= \beta_n$ more or less in Table page 783

For sure, on page 783 we have $g_0/g_2 = \beta_n/X$.

Now let's write the Gaussian Weights in A&S notation. Start from above,

$$\lambda_{\upsilon n} = h_{n-1} A_{n-1} (X/\beta_n) / [p_{n-1}]^2 = h_{n-1} (k_n/k_{n-1}) (X/\beta_n) / [p_{n-1}]^2 = h_{n-1} (k_n/k_{n-1}) (g_2/g_0) / [p_{n-1}]^2$$

and this agrees with my separate derivation of this result in my 1978 or older notes.

****

Appendix A: In developing an expression for the α_n of equation 10.7.5, it seems possible to start the analysis by doing a parts integration, in a manner similar to how we evaluated β_n. That method leads to a very complex answer for α_n which is either wrong (because parts is not justified or I goofed), or is right but in a way that is hard to see. In this appendix, I just want to record that calculation since I did the whole thing. Then there are comments at the end.

(Compute α_n). From the end of STEP 1 we quote this result:

$$Xp_n' = [\alpha_n + (n/2)X'' x] p_n + \beta_n p_{n-1}$$

and we have already found β_n. Let's now close with p_n to get:

$$(p_n, Xp_n') = \alpha_n (p_n, p_n) + n(X''/2) (p_n, xp_n) = \alpha_n h_n + n(X''/2) (p_n, xp_n)$$

Let's start with (p_n, Xp_n'). **Do the usual parts integration** to find that

$$- (p_n, Xp_n') = + \int dx\, p_n D[wXp_n] = + \int dx\, p_n \{ (wX)p_n' + p_n D(wX) \}$$

The first term is the same object that is on the LHS, and for the second $D(wX) = K_1 w p_1$ as we used earlier from the Rod formula, so we have

$$- (p_n, Xp_n') = + (p_n, Xp_n') + K_1 \int dx\, w\, p_n\, p_n\, p_1 = (p_n, Xp_n') + K_1 (p_n, p_n\, p_1)$$

which tells us that $-2(p_n, Xp_n') = K_1 (p_n, p_n\, p_1)$ so that

$$(p_n, Xp_n') = -(K_1/2)(p_n, p_n\, p_1)$$

Now examine $(p_n, p_n\, p_1) = (p_n, p_n [k_1 x + k_1']) = k_1(p_n, xp_n) + k_1' h_n$. Thus we have found:

$$(p_n, Xp_n') = -(K_1/2) [k_1(p_n, xp_n) + k_1' h_n] = -(k_1 K_1/2) (p_n, xp_n) - (k_1' K_1/2) h_n$$

Inserting this result as the LHS of our starting closed equation,

$$(p_n, Xp_n') = \alpha_n h_n + n(X''/2) (p_n, xp_n)$$

we get

$$-(k_1 K_1/2) (p_n, xp_n) - (k_1' K_1/2) h_n = \alpha_n h_n + n(X''/2) (p_n, xp_n)$$

which we can rewrite as

$$- \alpha_n h_n = [\, k_1 K_1/2 + n(X''/2\,) \,] \, (p_n, xp_n) + (k_1' K_1/2)\, h_n$$

Next, we evaluate (p_n, xp_n):

$$(p_n, xp_n) = (p_n, x\, [\, k_n x^n + k_n' x^{n-1} + o(x^{n-2}) \,]\,) \quad = k_n\, (p_n, x^{n+1}) + k_n'\, (p_n, x^n)$$

Let's start with the left term on the right using $p_{n+1}/k_{n+1} = x^{n+1} + (k'_{n+1}/k_{n+1})\, x^n + o(x^{n-1})$ and solve for x^{n+1}

$$(p_n, x^{n+1}) = (p_n,\ p_{n+1}/k_{n+1}) - (k'_{n+1}/k_{n+1})(p_n, x^n) \text{ - terms like } (p_n, x^{n-1}) \text{ and lower}$$
$$= - (k'_{n+1}/k_{n+1})(p_n, x^n)$$

We can now combine this with the right term on the right to get

$$(p_n, xp_n) = - k_n\, (k'_{n+1}/k_{n+1})(p_n, x^n) + k_n'\, (p_n, x^n) = k_n\, [\, (k_n'/k_n) - (k_{n+1}'/k_{n+1}) \,]\, (p_n, x^n)$$

Then repeating a piece of work done way back, we have for the rightmost factor:

$$(x^n, p_n) = (\, p_n/k_n + o(x^{n-1}) \,, p_n) = (1/k_n)(p_n, p_n) = (1/k_n) h_n$$

and this tells us that

$$(p_n, xp_n) = [\, (k_n'/k_n) - (k_{n+1}'/k_{n+1}) \,]\, h_n = (\, r_n - r_{n+1}) \, h_n$$

and we can now complete the solution using from above

$$- \alpha_n h_n = [\, k_1 K_1/2 + n(X''/2\,) \,]\, (p_n, xp_n) + (k_1' K_1/2)\, h_n$$

$$= [\, k_1 K_1/2 + n(X''/2\,) \,](\, r_n - r_{n+1})\ h_n + (k_1' K_1/2)\, h_n$$

to find that

$$\alpha_n = - [\, k_1 K_1/2 + n(X''/2\,) \,]\, [\, r_n - r_{n+1} \,] + (k_1' K_1/2)$$

which, unfortunately, is not the correct answer! The book's answer is this:

$$\alpha_n = n X'(0) - (X''/2)\, r_n$$

which is very much simpler. There are three possibilities for this disagreement

(1) the parts integration stage is unjustified. That would be true if the following did not vanish:

$$p_n^2\, wX \Big|_b^a \neq 0 \quad ??$$

How would we check this? Consider Jacobi at x = 1, for example. $w(x) \sim (1-x)^{\alpha}$ and $X \sim 2(1-x)$ and the p_n is finite, so it would seem that the parts vanish at a for $\alpha > -1$. At the b = -1 endpoint we have $w(x) \sim (1+x)^{\beta}$ and $X \sim 2(1+x)$ so they should vanish here for $\beta > -1$. In the Hermite case we have w being a huge expo at both end points, so parts = 0 in that case. In Laguerre, our only concern would be at a = 0 and there we have $w \sim x^{\alpha}$ and $X \sim x$ so again we are fine if $\alpha > -1$.

So my conclusion is that the parts integration I did was probably OK!

(2) I made one or more algebraic errors. I did check everything once. Regardless of errors, my answer is never going to have X'(0) in it, and it IS going to have things like k_1 and K_1 in it.

(3) The results really are the same. I suspect this is where the truth lies. Once again our two forms are:

$$\alpha_n = - [k_1K_1/2 + n(X''/2)] [r_n - r_{n+1}] + (k_1'K_1/2) \qquad \text{// me}$$

$$\alpha_n = nX'(0) - (X''/2)\, r_n \qquad \text{// them}$$

Regardless of any errors I may have made, my parts method brings in things like k_1 and K_1 which don't appear in the second form of the result. Similarly, my method does not bring in X'(0). We know that k_1 and k_1' and X'(0) and K_1 are connected by the equation $p_1K_1w = D[wX]$, but the connection is non-trivial. So I don't think you can just state flatly that the above two forms do not agree. Who knows what strange relations my be between all these constants.

Appendix B: What would β_n have been had I used the non-parts approach there?

$I = (Xp_n', p_{n-1}) = (p_{n-1}, Xp_n')$, so

$$(p_{n-1}, Xp_n') = (p_{n-1}, [X(0) + x\, X'(0) + (x^2/2)\, X''] [nk_nx^{n-1} + (n-1)k'_n\, x^{n-2} + (n-2)\, k_n''x^{n-3} + o(x^{n-4})])$$

Here you see the problem. We pick up a term of this form: $(X''/2)\,(n-2)\,k_n''\,(p_{n-1}, x^{n-1})$ and then we are stuck forever with this k_n'' parameter which is not one of Erdelyi's chosen parameters. Doing the parts integration gets is mixed up with k_1 and K_1 but at least these are an "the list" of OK parameters to use.