The General Theory of Chromatic Derivatives

Matthew Cushman
Timothy Herron

March 28, 2001

Abstract

We work out the general mathematical theory of chromatic derivatives as first advanced by Aleksandar Ignjatovic in his paper [1]. In the first section we consider the general theory and associated formulas behind families of differential operators associated with the Fourier transform. The middle two sections of this paper contain a formulary in which we record the operators and associated functions for each of the most important families of differential operators that can be used by machines processing industrially-important signals. Finally, we finish by recording some numerical methods useful in obtaining functions and formula constants associated with arbitrary families of such operators.

Contents

1 Chromatic Derivatives and Expansions

1.1 Orthogonal Polynomials in the Frequency Domain

The following gives the general theory of Chromatic expansions for an arbitrary family of orthogonal polynomials. Fix a real valued weight function $\rho(\omega)$, $\omega \in \mathbb{R}$, with $\int_{-\infty}^{\infty} \rho(\omega)d\omega = 2\pi$. We define an Hermitian inner product Q_ρ with respect to this weight:

$$< f, g >_\rho \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} \rho(\omega)f(\omega)\overline{g(\omega)}d\omega$$

A word on normalization: we put a factor of $(2\pi)^{-1}$ in front of this inner product since we will be applying this formula exclusively in the frequency domain. The Fourier transform is unitary if one renormalizes the integral in the frequency domain by this factor.

Definition 1 *The space of measurable functions f on the set $S \subset \mathbb{R}$ for which $< f, f >_\rho < \infty$ is denoted $L^2_\rho(S)$.*

If we take the case $S = \text{supp}(\rho) = \overline{\{\omega \in \mathbb{R} | \rho(\omega) \neq 0\}}$, then the inner product Q_ρ is positive definite. On this set, we will also have occasion to consider ρ^{-1} as a weight, and the corresponding space $L^2_{\rho^{-1}}$ will play a role.

We now fix a family of orthonormal polynomials p_i with respect to the real-valued weight ρ. Traditionally, one thinks of these polynomials as completely determined by the orthonormality condition, a requirement that the leading coefficient be positive, plus the condition that the degree of p_i is i. We will frequently renormalize these polynomials so that the have complex valued coefficients, so that they are conjugate-symmetric around 0 (in other words, so that $p_k(\omega) = \overline{p_k(-\omega)}$). For example, many classical families of orthogonal polynomials have the property that p_k is even when k is even, and is odd when k is odd. We wil renormalize these by $i^k p_k$, which then become conjugate symmetric. This is convenient, since we think of these functions as being define in the frequency domain, and conjugate-symmetric functions have real Fourier transforms.

For such a family to be defined, one requires $\int_{-\infty}^{\infty} |\omega^n|\rho(\omega)d\omega < \infty$ for every n. Recall that ρ is chosen so that $\int_{-\infty}^{\infty} \rho = 2\pi$.

Proposition 1 *Let S be the support of ρ. The elements $e_k = p_k\sqrt{\rho}$ are an orthonormal basis of $L^2(S)$.*

Proof. Orthonormality is trivial. The key is to see that they are a basis. This means that their linear span is dense. If S is bounded (as it is in many of the cases of interest), this is trivial from the Stone-Weierstrass theorem. The proof we give here works for arbitrary S.

Let W be the linear subspace of $L^2(S)$ generated by polynomials times $\sqrt{\rho}$. It suffices to show that any continuous linear functional on $L^2(S)$ which vanishes on W is 0. The dual space of $L^2(S)$ is again $L^2(S)$, which is spanned by elements g with compact support. Let g be such an element, which also vanishes on W. Let F_g be the Fourier transform of $g\sqrt{\rho}$, which is in L^1 (since both g and $\sqrt{\rho}$ is in L^2) and compactly supported. By the Paley-Wiener Theorem, F_g may be extended to a holomorphic function. Furthermore:

$$\begin{aligned}\frac{d^n}{dt^n}F_g(t) &= \int_{-\infty}^{\infty}(i\omega)^n g(\omega)\sqrt{\rho(\omega)}d\omega \\ &= 0\end{aligned}$$

Thus, $F_g = 0$. This implies that $g\sqrt{\rho} = 0$. Since the support of g is contained in the support of ρ, this implies that g itself is 0. □

The above theorem has as a special case classical results which prove this for particular families of orthogonal polynomials, such as the Hermite and Laguerre polynomials, but the author hasn't seen it proved this generally.

Fix a function F in $L^2_{\rho^{-1}}$. The elements $e_k = p_k\sqrt{\rho}$ are an orthonormal basis of $L^2(S)$. We then have by Parseval's Equality :

$$\begin{aligned}\frac{F}{\sqrt{\rho}} &= \sum_{k=0}^{\infty} < \frac{F}{\sqrt{\rho}}, e_k > e_k \\ &= \sum_{k=0}^{\infty} < F, p_k > \; p_k\sqrt{\rho}\end{aligned}$$

with the sum converging in L^2. We now multiply both sides by $\sqrt{\rho}$, which yields the following sum.

$$\begin{aligned}F &= \sum_{k=0}^{\infty} < \frac{F}{\rho}, p_k >_\rho \; p_k\rho \\ &= \sum_{k=0}^{\infty} < F, p_k > \; p_k\rho\end{aligned}$$

This obviously converges in $L^2(\rho^{-1})$. In fact, it actually converges in $L^1(\mathbb{R})$, as the following calculation shows:

$$\begin{aligned}\int\left|f - \sum_{k=0}^{m} < f, p_k > p_k\rho\right| &= \int\sqrt{\rho}\left|\frac{f}{\sqrt{\rho}} - \sum_{k=0}^{m} < \frac{f}{\rho}, p_k\rho > p_k\sqrt{\rho}\right| \\ &\le \left(\int\rho\right)^{1/2}\left(\int\left(\frac{f}{\sqrt{\rho}} - \sum_{k=0}^{m} < \frac{f}{\rho}, p_k\rho > p_k\sqrt{\rho}\right)^2\right)^{1/2} \\ &= \left(\int\rho\right)^{1/2}\left(\sum_{k=m+1}^{\infty} a_k^2\right)^{-1/2}\end{aligned}$$

where $a_k := < f\rho^{1/2}, e_n >$. The sum $\sum_{k=0}^{\infty}|a_k|^2$ converges, since these are the expansion of the L^2 function $f\rho^{-1/2}$ in the orthonormal bases e_k. Hence, the right hand side of the last line goes to 0 as $m \to \infty$.

1.2 Derivatives and Time Domain Expansions

We apply this to the case $F = \hat{f}$ and apply the inverse Fourier transform we obtain:

Theorem 1

$$f = \sum_{k=0}^{\infty} < \hat{f}, p_k > \widehat{p_k \rho}$$

Since the Fourier transform is a continuous map from $L_1(\mathbb{R})$ into $L_\infty(\mathbb{R})$, this sum converges pointwise.

Theorem 1 inspires the following definitions.

Definition 2 *Fix f, ρ, p_k as above.*

1. *The family of base time domain functions is $B_k = \widehat{p_k \rho}$ for each $k \in \mathbb{N}$. This depends only on ρ and p_k.*
2. *The k-th chromatic derivative of f at 0 is $< \hat{f}, p_k >$. This is a reasonable name since, by virtue of the fact that p_k is a polynomial, this a differential operator acting on f at 0. We write this symbolically as $CD_k(f, 0)$. By translating in the obvious manner we obtain the k-th chromatic derivative of f, $CD_k(f, t)$ as a function of t.*

By translating the equality in Theorem 1 we obtain:

$$f(t) = \sum_{k=0}^{\infty} CD_k(f, t_0) B_k(t - t_0)$$

This expresses f as a sum of products, one term in each product is the scalar value of a differential operator applied to f, and the other is a function of t which is independent of f. It is akin to Taylor's series, except that it is tailored to signals with specific spectral content through the frequency domain weight.

1.3 Error Bounds

We have the following error bound on the truncated sum:

Theorem 2 *Let $f \in L^2_{\rho^{-1}}$, $t, t_0 \in \mathbb{R}$. Then*

$$\left| f(t) - \sum_{k=0}^{n} CD_k(f, t_0) B_k(t - t_0) \right| \leq \sqrt{\int_{-\infty}^{\infty} \rho^{-1} |\hat{f}|^2} \sqrt{1 - \sum_{k=n+1}^{\infty} |B_k(t)|^2}$$

Proof. Since the sum converges pointwise, we can just apply the Cauchy-Schwatz inequality:

$$\begin{aligned} \left| f(t) - \sum_{k=0}^{n} < \hat{f}, p_k > \widehat{p_k \rho}(t) \right| &= \left| \sum_{k=n+1}^{\infty} < \hat{f}, p_k > \widehat{p_k \rho}(t) \right| \\ &\leq \sqrt{\sum_{k=n+1}^{\infty} < \hat{f}, p_k >^2} \sqrt{\sum_{k=n+1}^{\infty} |\widehat{p_k \rho}(t)|^2} \\ &\leq \sqrt{\int_{-\infty}^{\infty} \rho^{-1} |\hat{f}|^2} \sqrt{\sum_{k=n+1}^{\infty} |B_k(t)|^2} \end{aligned}$$

□

This theorem bounds the error by a product of two terms. One depends only upon f, and is independent of the order of the expansion and the expansion time t. The other is a function of time t and the order of expansion n, but independent of f. We define the square of the left hand factor (which depends on f) to be the ρ-energy of f, written $E_\rho(f)$:

$$E_\rho(f) = \int_{-\infty}^{\infty} \rho^{-1}|\hat{f}|^2$$

This is the square of the "length" of $f \in L^2_{\rho^{-1}}$ with respect to the positive definite quadratic form Q_ρ.

Define $B = \hat{\rho}$. We can simplify this last expression by the following proposition:

Proposition 2

$$B(s-t) = \sum_{k=0}^{\infty} B_k(s)\overline{B_k(t)}$$

Proof. We begin by calculating the chromatic derivatives of B.

$$\begin{aligned} CD_k(B,t) &= < \rho(\omega)e^{i\omega t}, p_k(\omega) > \\ &= \int_{-\infty}^{\infty} \overline{\rho(\omega)p_k(\omega)}e^{i\omega t} d\omega \\ &= \overline{B_k}(-t) \end{aligned}$$

Thus,

$$\begin{aligned} \sum_{k=0}^{\infty} B_k(s)\overline{B_k(t)} &= \sum_{k=0}^{\infty} CD_k(B,-t)B_k(s) \\ &= B(s-t) \end{aligned}$$

□

Since $\int_{-\infty}^{\infty} \rho = 2\pi$, we see that $B(0) = 0$. Hence, the truncation error bound simplifies to:

$$|f - \cdots| \leq \sqrt{e_\rho(f)}\sqrt{1 - \sum_{k=0}^{n} |B_k(t)|^2}$$

This implies that the convergence is of a special nature. The bound on the right is the product of two factors. The first factor is a frequency weighted energy measure of the function f. The second is a "window" which is independent of f, and which is uniformly bounded by $B(0) = 1$. For typical families, this expansion is very local, with the window being 0 at 0, and very small in a neighborhood of 0.

1.4 Examples

A table of common values of the parameters is below. The family p_k is actually i^k times the usual orthogonal polynomial, and in the band-limited case is scaled so that the band-limit is π (for example, in the case of Bessel-Tchebysheff, $p_k = i^k T_k(\omega/\pi)$). For more detailed of each case, see the second chapter of this paper. For the theoretical calculations we normalize ρ so that it integrates to 2π, but in order to express the functions simply in terms of classical functions we relax that criterion. Thus, the polynomials might not be orthonormal, but merely orthogonal, with respect to Q_ρ. We thus include the parameter a_k, which is the norm of p_k in in $L^2(\rho)$. Thus, for the formulae above are adjusted by the additional of additional scalar factors a_k in to appropriate terms involving p_k and B_k. For a more in-depth discussion of these and other particular cases, see Chapter 2.

Name	weight	B_k	a_k
Bessel-Tchebysheff T	$\frac{1}{1-(\omega/\pi)^2}$	$J_k(\pi t)$	1 if $k=0$, otherwise $\sqrt{2}$
Double Bessel/ Tchebysheff U	$\sqrt{1-(\omega/\pi)^2}$	$J_k(\pi t) + J_{k+2}(\pi t)$	1
Spherical Bessel/ Legendre P	1	$\frac{1}{\sqrt{2t}} J_{k+1/2}(\pi t)$	$\sqrt{2k+1}$
Hermite	$e^{-\omega^2}$	$t^k e^{t^2/4}$	$2^{-k}\pi^{-1/2}/k!$

2 Iteration of chromatic derivatives

Usual derivatives have the property $(f^{(k)})^{(n)} = f^{(k+n)}$. This holds since $f^{(n)}$ is simply the result of iterating a fixed operator (namely, differentiation). Chromatic derivatives are not defined in the manner, however. The two functions $CD_k(CD_n(f))$ and $CD_{k+n}(f)$ will not generally agree. There is a relation between the two, however.

If one examines the identity $(f^{(k)})^{(n)} = f^{(k+n)}$ in the frequency domain, it can be seen as an equality between two polynomials: $(i\omega)^n(i\omega)^k = (i\omega)^{n+k}$. In order to iterate chromatic derivatives, one has to look at the transfer function of operator given by composing CD_k with CD_n, which is the product of the transfer functions $p_k(i\omega)p_n(i\omega)$. This is again a polynomial of degree $k+n$. Since the polynomials $p_0(i\omega)\dots p_m(i\omega)$ form a basis of the space of polynomials of degree $\leq m$, we may expand this polynomial as a sum:

$$p_k(i\omega)p_n(i\omega) = \sum_{j=0}^{k+n} \beta_j p_j(i\omega)$$

This yields a composition formula:

$$CD_k(CD_n(f)) = \sum_{j=0}^{k+n} \beta_j CD_j(f)$$

The coefficients β_j depend only on k, n and the choice of weight ρ (which of course determines the family of orthogonal polynomials p_i).

3 Chromatic expansions and linear operators

In Section 1 we give a general approximation formula with interesting convergence properties. Here we examine the effect of applying a linear operator to this expansion. The case of a differential operator is particularly interesting.

Fix a signal $f \in L^2(\rho^{-1})$. The chromatic expansion of f is given by:

$$f(t) = \sum_{k=0}^{\infty} CD_k(f, t_0) B_k(t - t_0)$$

Now, let H be a shift-invariant linear operator (i.e. channel) with frequency response $\hat{h}(\omega)$. Thus:

$$H(f)(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{h}(\omega) \hat{f}(\omega) e^{-i\omega t} d\omega$$

Expand $\hat{h}(\omega)$ as a sum of orthogonal polynomials:

$$\hat{h}(\omega) = \sum_{j=0}^{\infty} b_j p_j(\omega)$$

We can then calculate how H acts on the chromatic derivatives of f (i.e. the coefficients a_k). Multiplying together the expansion of $\hat{f}$ and $\hat{h}$ we get

$$(\hat{f}\hat{h})(\omega) = \sum_{i,j} a_i b_j \rho(\omega) p_i(\omega) p_j(\omega)$$

We define constants c_{ij}^k which express products of members of the family of orthogonal polynomials:

$$p_i p_j = \sum_{k=0}^{i+j} c_{ij}^k p_k$$

We then have:

$$\begin{aligned}
(\hat{f}\hat{h}) &= \sum_{i,j} a_i b_j \rho p_i p_j \\
&= \sum_{i,j} a_i b_j \left(\sum_{k=0}^{i+j} c_{ij}^k \rho p_k \right) \\
&= \sum_k \left(\sum_{i+j \geq k} c_{ij}^k a_i b_j \right) \rho p_k
\end{aligned}$$

Thus:

$$\mathrm{CD}_k(H(f)) = \sum_i \left(\sum_{j \geq k-i} c_{ij}^k b_j \right) a_i$$

Furthermore, the coefficients b_j are simply the chromatic derivatives of $H(B_0)$:

$$\begin{aligned}
b_j &= \int_{-\pi}^{\pi} \left(\hat{h} \rho T_j \right) d\omega \\
&= \mathrm{CD}_k(H(B_0))
\end{aligned}$$

Equations 1 and 1 have the following interpretation: the chromatic derivatives of the filtered signal $H(f)$ are calculated by applying a matrix to the chromatic derivatives of f. The matrix itself is given by a predetermined linear combination of the chromatic derivatives of the filter applied to the base function $H(B_0)$. Furthermore, the matrix entry corresponding to the contribution of $CD_i(f)$ to $CD_k(H(f))$ depends only on $CD_j(B_0)$ for $j \geq k - i$. In particular, since $CD_j(B_0)$ tends to zero as j increases (in fact, the sum of the squares converges, with the speed of the convergence determined by the severity of the channel H) one can get an arbitrarily good approximation to $H(f)$ by including enough terms in a finite truncation.

3.1 An Example: the Tchebysheff/Bessel Expansion

It is worth giving an example to illustrate how this works. We use the Tchebysheff T_k polynomials, since they satisfy

$$T_i T_j = \frac{T_{i+j} + T_{|i-j|}}{2}$$

which implies that most of the coefficients $c_{i,j}^k$ are zero. We don't use $T_k(\omega)$, but $i^k T_k(\frac{\omega}{\pi})$ as the family of orthogonal polynomials, so that the expansion around 0 is of the form

$$f(t) = CD_0(f,0)J_0(\pi t) + 2\sum_{k=1}^{\infty} CD_k(f,0)J_k(\pi t)$$

The support of the weight $\rho(\omega) = (1 - (\omega/\pi)^2)^{-1/2}$ in this case is $[-\pi, \pi]$. Thus, this family can express any band-limited signal of finite energy (and even some signals of infinite energy, since we only need that $\int |\hat{f}|^2 \sqrt{1 - (\omega/\pi)^2}$ be finite).

Expand the transfer function of the filter $\hat{h}$ as

$$\hat{h}(\omega) = \sum_{k=0}^{\infty} b_k (i^k) T_k \left(\frac{\omega}{\pi}\right)$$

These values appear as the leftmost column in the transformation matrix associated to the operator H acting on the chromatic derivatives. The whole matrix can be reconstructed from this column as follows (again, the CDs are not quite the same as the usual normalization for historical reasons):

$$\begin{bmatrix} b_0 & \frac{b_1}{2} & \frac{b_2}{2} & \frac{b_3}{2} & \cdots \\ b_1 & \frac{b_0 - b_2}{2} & \frac{b_1 - b_3}{2} & \frac{b_2 - b_4}{2} & \cdots \\ b_1 & \frac{b_1 + b_3}{2} & \frac{b_0 + b_4}{2} & \frac{b_1 + b_5}{2} & \cdots \\ \vdots & \vdots & & \ddots & \end{bmatrix}$$

3.2 Differential Operators

Another particularly nice example is given by the case of H being a differential operator with constant coefficients of degree d (going back to the general family ρ, p_k). In that case, $\hat{h}(\omega)$ is itself a polynomial of degree d, and expanding it as a sum or orthogonal polynomials yields:

$$\hat{h} = \sum_{j=0}^{d} b_k p_k(\omega)$$

so that $b_j(\omega) = 0$ for $j > d$. In this case, the matrix entries are of the form

$$\sum_{j \geq k-i}^{d} c_{ij}^{k} b_j$$

Thus, the contribution of $CD_i(f)$ to $\mathrm{CD}_k(H(f))$ is zero unless $k \leq i + d$. In particular, this says that the truncation of the chromatic expansion of $H(f)$ at stage k is the truncation of the chromatic expansion of $H(g)$, where g is the chromatic expansion of f truncated at stage $d + k$.

This argument can be seen more clearly from the following point of view. In the frequency domain, taking the chromatic expansion of f of order k corresponds to projecting $\hat{f}\rho^{-1/2}$ to the subspace V_k of polynomials of degree at most k times $\rho^{1/2}$. If p is any polynomial of degree m, we have that $pV_k \subset V_{k+m}$, and its dual $pV_k^{\perp} V_{k-m}^{\perp}$ (here, $\perp$ denotes orthogonal complement). Viewing multiplication by p as a map $V_k \oplus V_k^{\perp} \to V_{k-m} \oplus V_{k-m}^{\perp}$, the second summand in the domain gives no contribution to the first summand in the codomain. Thus, multiplication by p commutes with the projections to the first summands, which, restated informally, says that the expansion of differential operator applied to f is the differential operator applied to the expansion of f.

In general, applying an operator will induce a non-zero linear map $V_k^{\perp} \to V_{k-m}$. The norm of this map is a measure of the "non-locality" of the operator. By choosing k large relative to m we can force this factor to be small.

4 Chromatic Expansions of Distributions

The previous sections show how one can expand functions satisfying certain conditions in the frequency domain. These conditions are stated in terms of weighted integrals of $\hat{f}^2$ converging. Many natural functions don't satisfy these conditions, however. A primary example is one of the most important types of functions of all: sinusoids. These must be dealt with using more sophisticated techniques, since their Fourier transforms don't exist as functions, but as distributions.

We recall the following usual notations used for tempered distributions: $\mathcal{S}$ is the space of C^{∞} functions on $\mathbb{R}$ which are rapidly decreasing. This is endowed with the usual topology for the test function space. The space of tempered distributions $\mathcal{S}'$ is the topological dual of this space. The support of a distribution is well-defined; we denote by $\mathcal{E}$ the space of compactly supported distributions, and by $\mathcal{D}(S)$ the set of distributions whose support is contained in $S \subset \mathbb{R}$.

We assume that the family of orthogonal polynomials in question satisfies the following criteria:

Criterion 1 *Let ϕ be a C^{∞} function of rapid descent with compactly supported Fourier transform. Then*

1. *The sequence $< \phi, p_k \rho >$ goes to zero faster than k^{-n} for every n.*
2. *The sequence $||p_k^{(n)}||_{\infty}$ is bounded by k^m for some m, where the norm is taken over some compact set $K \subset S^{\circ}$ (C is independent of k, but depends on K and ϕ).*

We can show that this holds for many classical families of orthogonal polynomials. For instance, using Rodrigues' formula for the Jacobi and Laguerre we can prove the first assumption (see below). The second assumption holds as well. We conjecture that it holds for all families of orthogonal polynomials.

4.1 Convergence of Test Functions

In order to analyze distributions, we must examine the convergence of the series in question on test functions. The precise relationship between convergence for test function and distributions will be developed later. The result we need is the following:

Proposition 3 *Let $S = supp\,\rho$, $\phi \in \mathcal{S}$, $\hat{\phi}$ compactly supported, K compact subset of the interior S°. Then*

$$\phi = \sum_{k=0}^{\infty} < p_k\rho, \phi > p_k$$

uniformly on K. The derivatives of both sides also converge uniformly on K.

Proof. We know that the sum converges in $L^2(S)$ to ϕ restricted to S . We need to show that the m-th derivatives of the partial sums on the right converge uniformly the m-th derivatives of ϕ. This is a consequence of the criterion, which immediately shows that x^n (for a fixed n) times the derivatives of the partial sums converge uniformly on compact sets contained in S, and thus converge uniformly on compact sets to the derivatives of ϕ. □

4.2 Chromatic Expansion of Distributions

Theorem 3 *Let $\eta \in \mathcal{E}(K)$ be a distribution supported on $K \subset S$. Then*

$$\eta = \sum_{k=0}^{\infty} < \eta, p_n > p_n\rho$$

as distributions.

Proof. Apply the partial sums of the right hand side to an element $\phi \in \mathcal{S}$ with compactly supported Fourier transform:

$$\begin{aligned}\left(\sum_{k=0}^{M} < \eta, p_k > p_k\rho\right)\phi &= \sum_{k=0}^{M} < \eta, p_k >< p_k\rho, \phi > \\ &= < \eta, \sum_{k=0}^{M} < \phi, p_k\rho > p_k >\end{aligned}$$

The partial sum inside the brackets in the last line converges to ϕ uniformly in all derivatives, by our criterion. Thus, this sum converges to $< \eta, \phi >$. This implies that the partial sums converge as distributions to η. □

4.3 Examples where these formulae hold

We verify that the criterion holds for families of orthogonal polynomials satisfying some Rodrigues'-type fomula.

Proposition 4 *Let ρ be a weight, and p_n the family of orthonormal polynomials with respect to ρ. Assume that this family satisfies a formula of Rodrigues'-type*

$$p_n(\omega) = \rho^{-1}\frac{d^n}{d\omega^n}(\alpha_n(\omega))$$

where α_n is such that

$$\lim_{n\to\infty} t^n ||\alpha_n||_1 = 0$$

for every $t \in \mathbb{R}$. Then this family of orthogonal polynomials satisfies the first part of the criterion.

Proof. Use integration by parts to show

$$\begin{aligned} |\int \phi(\omega)p_n(\omega)\rho(\omega)d\omega| &= |\int \phi(\omega)\frac{d^n}{d\omega^n}\alpha_n d\omega| \\ &= |\int \phi^{(n)}(\omega)\alpha_n(\omega)d\omega| \\ &\leq ||\phi^{(n)}||_\infty ||\alpha_n||_1 \end{aligned}$$

Since ϕ has compactly supported Fourier transform, there is a t such that $t^n||\phi^{(n)}||_\infty$ is bounded for all n. By the assumption on α_n, the product must go to zero as $n \to \infty$. □

Corollary 1 *The Jacobi and Laguerre polynomials satisfy the criterion.*

Proof. Just apply the classical Rodrigues' formula and the above proposition to obtain the first part of the criterion. The second part follows from the fact that the orthonormalized polynomials p_k have coefficients which are bounded by polynomials in k, and thus the coefficients of $p_k^{(n)}$ grow at most polynomially in k. □

We can apply Theorem 3 to give expansions of functions not in $L^2(\mathbb{R})$, and hence don't have nice Fourier transforms. A principle example of this is $\sin(\omega t)$, or $e^{i\omega t}$ for a fixed ω, which has a Dirac delta function δ_ω as Fourier transform. Since applying a differential operator to $e^{i\omega t}$ is the same as evaluating the symbol of the operator at $i\omega$, we obtain the following general expansion

Proposition 5 *Assume that the family of orthogonal polynomials satisfies the criterion. Then*

$$e^{i\omega t} = \sum_{k=0}^{\infty} i^{-k} p_k(i\omega) B_k(t)$$

5 Review of Formulary Contents

For each of the families we record the following information:

1. The names that are used for the family of differential operators.
2. The *frequency-domain weight*, which controls the frequency domain of operation of the family of operators; i.e. the exact kind of filtering the differential operator will do when applied to a signal. The weight determines all of the other formulas for a given family as well as implying differential operator characteristics such as:
   (a) the energy of the signals that the operators can process; ranging from certain infinite energy signals to restricted kinds of finite energy signals.
   (b) the kinds of signals that the operators naturally predicts [1] outside of the domain of ordinary local representation using the differential operators. I.e. Signal with the same spectrum as the weight are naturally predicted using the associated differential operators.

[1] with respect to ordinary stochastic variation, of course.

(c) how real and complex signals interact (or do not) when representing them using the family of differential operators.

3. The *frequency-domain polynomials*, which are orthogonal polynomials generated by the frequency-domain weight, and they govern exactly how the operators use frequency information contained in the signals they process. Note that we do not include the imaginary term i^k as part of the polynomial in the formulary: it just shows up where needed in the formulas.

4. The explicit *time-domain functions* which are associated with the chromatic representation of each family. These are the functions which are used to interpolate a signal locally after the family's differential operators have acted upon the signal. See the appendix for a derivation of the time-domain functions from the frequency-domain polynomials in the case of the ultraspherical polynomials.

5. The *Fourier kernel decomposition* which shows how the time and frequency-domain functions are connected and which can be used to schematically derive (for a precise derivation see section 2 of this document) the chromatic representation associated with each family as follows. Given that the kernel for the Chebyshev T-Bessel J family, e.g., is

$$e^{i\omega t} = T_0(\frac{\omega}{\pi})J_0(\pi t) + 2\sum_{k=1}^{\infty} i^k T_k(\frac{\omega}{\pi})J_k(\pi t)$$

we can see that

$$\begin{aligned} f(t) &= \frac{1}{2\pi}\int_{-\pi}^{+\pi} \hat{f}(\omega)e^{i\omega t}d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{+\pi} \hat{f}(\omega)[T_0(\frac{\omega}{\pi})J_0(\pi t) + 2\sum_{k=1}^{\infty} i^k T_k(\frac{\omega}{\pi})J_k(\pi t)]d\omega \\ &= J_0(\pi t)[\frac{1}{2\pi}\int_{-\pi}^{+\pi} \hat{f}(\omega)e^{i\omega 0}d\omega] \\ &\quad +2\sum_{k=1}^{\infty} J_k(\pi t)[\frac{1}{2\pi}\int_{-\pi}^{+\pi} i^k T_k(\frac{\omega}{\pi})\hat{f}(\omega)e^{i\omega 0}d\omega] \\ &= J_0(\pi t)CD_0[f](0) + 2\sum_{k=1}^{\infty} J_k(\pi t)CD_k[f](0) \end{aligned}$$

where $CD_k[f](t) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} i^k T_k(\frac{\omega}{\pi})\hat{f}(\omega)e^{i\omega t}d\omega$ is the Chebyshev family's sequence of differential operators (note that we are not normalizing the family here).

6. The *chromatic derivative definition*, a recursive generation of the family's differential operators, which is closely related to (and straightforwardly derivable from) the recursively generated, frequency-domain polynomials.

7. The *chromatic derivative operator iteration* formula which governs how a chromatic derivative acts upon another chromatic derivative within the same family.

8. The *chromatic approximation error bound*, which gives a sharp upper bound on the error we can expect when we only use a finite number of terms of the chromatic representation of a given representable signal.

   Other pertinent formulas also appear, such as the chromatic derivatives (CDs) of important functions or special time-domain function orthogonality relations.

6 Bandlimited Signal Representation and Signal Processing

6.1 Bessel-Chebyshev T

This is one of the most important family of operators because of the simplicity of its formulas and because of the many properties that Chebyshev T polynomials and Bessel J functions possess. In particular the Chebyshev T Bessel J CD operators are useful in the following situations:

- wherever we need to approximate a signal or transfer function in the frequency domain since Chebyshev T polynomials are unsurpassed in their approximation power - they provide an optimal minimax approximation; e.g. they can be profitably used in filtering (especially adaptive filtering).
- wherever we need to or are able to solve time-domain differential equations having solutions which are Bessel functions, of which there are many in physical phenomena; e.g. in antenna design.
- when we need to process signals which locally look like they contain infinite energy; e.g. transient-like signals in audio applications.
  For this section, we let $c_0 = 1$ and $c_i = \sqrt{2}$ for positive integers ($i > 0$).

Frequency Domain Weight $\rho^T = \frac{2}{\pi}(1-(\frac{\omega}{\pi})^2)^{-\frac{1}{2}}$, Range $[-\pi,\pi]$

Frequency Domain Polynomials $p_i^T(\omega) = c_i T_i(\frac{\omega}{\pi})$, (Chebyshev Polys of the 1st kind)

Time Domain Functions $B_i^T(t) = c_i J_i(\pi t)$, (Bessel Functions of the 1st kind)

Fourier Kernel Decomposition

$$e^{i\omega t} = T_0(\frac{\omega}{\pi})J_0(\pi t) + 2\sum_{k=1}^{\infty} i^k T_k(\frac{\omega}{\pi})J_k(\pi t)$$

Chromatic Derivative Definition

$$\begin{aligned} CD_T(0,f) &= f \\ CD_T(1,f) &= \frac{\sqrt{2}}{\pi}\frac{\partial f}{\partial t} \\ CD_T(2,f) &= \frac{2}{\pi}\frac{\partial}{\partial t}(CD_T(1,f)) + \sqrt{2}CD_T(0,f), \\ CD_T(i,f) &= \frac{2}{\pi}\frac{\partial}{\partial t}(CD_T(i-1,f)) + CD_T(i-2,f), \text{ where } i \geq 3. \end{aligned}$$

$$\text{Equivalently, } CD_T(k,f)(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} i^k p_k^T(\omega)\hat{f}(\omega)e^{i\omega t}d\omega.$$

Chromatic Derivative Operator Iteration

$$T_m(x)T_k(x) = \frac{1}{2}(T_{m+k}(x) + T_{|m-k|}(x))$$

implies

$$CD_T(m, CD_T(k,f)) = \frac{1}{2}(\left(\frac{c_m c_k}{c_{m+k}}\right) CD_T(m+k,f) + (-1)^{Min(m,k)}\left(\frac{c_m c_k}{c_{|m-k|}}\right) CD_T(|\, m-k \,|, f))$$

and since

$$CD_T(m, B_0^T(t)) = (-1)^m B_m^T(t),$$

we obtain

$$CD_T(k, B_m^T(t)) = \frac{(-1)^k}{2}(\left(\frac{c_m c_k}{c_{m+k}}\right) B_{m+k}^T(t) + (-1)^{Min(m,k)}\left(\frac{c_m c_k}{c_{|m-k|}}\right) B_{|m-k|}^T(t).$$

Chromatic Approximation Error Bound

$$\left| f(t) - \sum_{k=0}^{m} CD_T(k,f)(t_0) B_k^T(t-t_0) \right| \leq A\sqrt{1 - \sum_{k=0}^{m} B_k^T(t-t_0)^2}$$

where $A = \sqrt{\int_{-\pi}^{\pi} \frac{|\hat{f}(\omega)|^2}{\rho^T} d\omega}$ is a "windowed" energy.

Orthogonality Relation

$$\int_{-\infty}^{\infty} B_m^T(t) B_n^T(t) \frac{dt}{|t|} = 0$$

if $m \neq n$. The value is ∞ for $m = n = 0$, and $2/n$ for $m = n > 0$.

CDs of Sinc

$$CD_T(m, sinc)(0) = \begin{cases} 1 & : \ m = 0. \\ 0 & : \ m \text{ odd} \\ \sum_{j=0}^{m/2} (-1)^{j+\frac{m}{2}} 2^{m-2j+1} \frac{\sqrt{2m}\binom{m-j}{j}}{(m-j)(m-2j+1)} & : \ m \geq 2 \text{ and even} \end{cases}$$

CDs of Sine and Cosine $(0 \leq a \leq \pi)$

$$CD_T(m, \sin(at)) = \begin{cases} (-1)^{(m-1)/2} p_m^T(a) & : \ m \text{ odd} \\ 0 & : \ m \text{ even} \end{cases}$$

$$CD_T(m, \cos(at)) = \begin{cases} (-1)^{m/2} p_m^T(a) & : \ m \text{ even} \\ 0 & : \ m \text{ odd} \end{cases}$$

CDs of Gaussian Noise with PSD ρ_T .

If signal f is Gaussian noise with PSD ρ_T and power ν, then $CD_T(m, f)(x)$ as x varies (varied at the bandlimit's Nyquist rate) are uncorrelated Gaussian random variables with variance ν.

6.2 Double Bessel-Chebyshev U

The double Bessel - Chebyshev U differential operators are important themselves and also indicate the possibilities available in other ultrashperical polynomials which use frequency-domain weights which emphasize lower frequencies inside the bandlimit. In particular, these operators are useful:

- whenever we wish to use transmission modulations in severely attentuating channels which force one to transmit most information in the lower frequencies.
- whenever we wish to build waveforms which dampen quickly; an inherent property of double Bessel functions.

Frequency Domain Weight $\rho^U(\omega) = \frac{4}{\pi}(1 - (\frac{\omega}{\pi})^2)^{\frac{1}{2}}$ Range $[-\pi, \pi]$

Frequency Domain Polynomials $p_i^U(\omega) = U_i(\frac{\omega}{\pi})$, (Chebyshev Polys of the 2nd kind)

Time Domain Functions $B_i^U(t) = D_i(\pi t) = J_i(\pi t) + J_{i+2}(\pi t)$ ("Double" Bessels)

Fourier Kernel Decomposition

$$e^{i\omega t} = \sum_{k=0}^{\infty} i^k U_k(\frac{\omega}{\pi}) D_k(\pi t)$$

Chromatic Derivative Definition

$$\begin{aligned} CD_U(0, f) &= f \\ CD_U(1, f) &= \frac{2}{\pi}\frac{\partial f}{\partial t} \\ CD_U(i, f) &= \frac{2}{\pi}\frac{\partial}{\partial t}(CD_U(i-1, f)) + CD_U(i-2, f), \text{ where } i \geq 2. \end{aligned}$$

$$\text{Thus, } CD_U(k, f)(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} i^k p_k^U(\omega)\hat{f}(\omega)e^{i\omega t}d\omega.$$

Chromatic Derivative Operator Iteration

$$CD_U(m, CD_U(k, f)) = \sum_{\substack{i=|k-m| \\ i \text{ even if } k+m \text{ even} \\ i \text{ odd if } k+m \text{ odd}}}^{k+m} (-1)^{\frac{(k+m-i)}{2}} CD_U(i, f)$$

and since

$$CD_U(k, B_0^U(t)) = (-1)^k B_k^U(t)$$

$$CD_U(k, B_m^U(t)) = (-1)^k \sum_{\substack{i=|k-m| \\ i \text{ even if } k+m \text{ even} \\ i \text{ odd if } k+m \text{ odd}}}^{k+m} (-1)^{\frac{(k+m-i)}{2}} B_i^U(t)$$

Chromatic Approximation Error Bound

$$\left| f(t) - \sum_{k=0}^{m} CD_U(k, f)(t_0) B_k^U(t - t_0) \right| \leq A\sqrt{1 - \sum_{k=0}^{m} B_k^U(t - t_0)^2}$$

where $A = \sqrt{\int_{-\pi}^{\pi} \frac{|\hat{f}(\omega)|^2}{\rho^U} d\omega}$ is a "windowed" energy.

Orthogonality Relation

$$\int_{-\infty}^{\infty} B_m^U(\pi t) B_n^U(\pi t)|t|dt = 0$$

if $m \neq n$. The value is $4(m+1)$ if $m = n$.

CDs of Sinc

$$CD_U(m, sinc)(0) = \begin{cases} 1, & : \ m = 0. \\ 0, & : \ m \text{ odd} \\ \sum_{j=0}^{m/2}(-1)^{j+\frac{m}{2}} 2^{m-2j} \frac{\binom{m-j}{j}}{(m-2j+1)} & : \ k \geq 2 \text{ and even} \end{cases}$$

CDs of Sine and Cosine $(0 \leq a \leq \pi)$

$$CD_U(m, \sin(at)) = \begin{cases} (-1)^{(m-1)/2} p_m^U(a) & : \; m \text{ odd} \\ 0 & : \; m \text{ even} \end{cases}$$

$$CD_U(m, \cos(at)) = \begin{cases} (-1)^{m/2} p_m^U(a) & : \; m \text{ even} \\ 0 & : \; m \text{ odd} \end{cases}$$

CDs of Gaussian Noise with PSD ρ_U **.**
If signal f is Gaussian noise with PSD ρ_U and power ν, then $CD_U(m, f)(x)$ as x varies (varied at the bandlimit's Nyquist rate) are uncorrelated Gaussian random variables with variance ν.

6.3 Spherical Bessel-Legendre

The spherical Bessel j - Legendre P family of differential operators are very useful because of the flat frequency weighting they have. Such operators are useful:

- whenever general bandlimited signals are being processed, whether for filtering, or transmission, or compression.
- whenever we wish to generalize the sinc functions which are the harmonic interpolants of signals; sinc is the zeroth time-domain function of the spherical Bessels. Thus, this family of differential operators can be used to represent locally any signal of finite energy.

Frequency Domain Weight $\rho^P(\omega) = 1$, Range $[-\pi, \pi]$

Frequency Domain Polynomials $p_n^P(\omega) = \sqrt{2n+1} P_n(\frac{\omega}{\pi})$, (Legendre (spherical) Polynomials)

Time Domain Functions $B_n^P(t) = \sqrt{2n+1} j_n(t) = \sqrt{2n+1} \frac{J_{n+\frac{1}{2}}(\pi t)}{\sqrt{2t}}$ (spherical Bessels of the first kind - normalized)

Fourier Kernel Decomposition

$$e^{i\omega t} = \sum_{k=0}^{\infty} i^k (2k+1) P_k(\frac{\omega}{\pi}) j_k(t)$$

Chromatic Derivative Definition

$$\begin{aligned} CD_P(0, f) &= f \\ CD_P(1, f) &= \frac{\sqrt{3}}{\pi} \frac{\partial f}{\partial t} \\ CD_P(n, f) &= \left(\frac{(2n-1)^{3/2}}{\pi n \sqrt{2n-3}} \right) \frac{\partial}{\partial t} (CD_P(n-1, f)) + \\ & \left(\frac{(n-1)\sqrt{2n-1}}{n\sqrt{2n-5}} \right) \cdot CD_P(n-2, f), \; n \geq 2. \end{aligned}$$

$$\text{Equivalently, } CD_P(k, f)(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} i^k p_k^P(\omega) \hat{f}(\omega) e^{i\omega t} d\omega.$$

Chromatic Derivative Operator Iteration

$$\mathrm{CD}_P(m, \mathrm{CD}_P(k, f)) = \sum_{i=0}^{\min(m,k)} (-1)^i a_{k,m,i} \frac{\sqrt{2m-1}\sqrt{2k-1}}{\sqrt{2(k+m-2i)-1}} \cdot \mathrm{CD}_P(k+m-2i, f)$$

for some constants $a_{k,m,i}$ computable from the Ferrers-Adams linearization formula:[2]

$$\begin{aligned} P_m(x)P_k(x) &= \sum_{i=0}^{Min(m,k)} \left(\frac{2m+2k+1-4i}{2m+2k+1-2i}\right) \cdot \left(\frac{\binom{m+k-i}{k}}{\binom{m+k-i}{k}_{\frac{1}{2}}}\right) \\ &\quad \left(\frac{\binom{k}{i}}{\binom{k}{i}_{\frac{1}{2}}}\right) \cdot P_{m+k-2i}(x) \\ &= \sum_{i=0}^{Min(k,m)} a_{k,m,i} P_{m+k-2i}(x) \end{aligned}$$

And since

$$CD_P(m, B_0^P(t)) = (-1)^m B_m^P(t),$$

$$CD_P(k, B_m^P(t)) = \sum_{i=0}^{\min(k,m)} (-1)^{i+k} a_{k,m,i} \frac{\sqrt{2m-1}\sqrt{2k-1}}{\sqrt{2(k+m-2i)-1}} B_{k+m-2i}^P(t).$$

Chromatic Approximation Error Bound

$$\left| f(t) - \sum_{k=0}^{m} CD_P(k, f)(t_0) B_k^P(t-t_0) \right| \leq A \sqrt{1 - \sum_{k=0}^{m} B_k^P(t-t_0)^2}$$

where $A = \sqrt{\int_{-\pi}^{\pi} \frac{|\hat{f}(\omega)|^2}{\rho^P} d\omega}$ is a "windowed" energy.

Orthogonality Relation

$$\int_{-\infty}^{\infty} B_m^P(t) B_n^P(t) dt = 0$$

if $m \neq n$. The value is $\frac{\sqrt{(2m-1)(2n-1)}\Gamma(\frac{m+1}{2})}{\sqrt{\pi}\Gamma(\frac{m+2}{2})}$ if $m = n$.

CDs of Sinc

$$CD_P(m, sinc)(0) = \begin{cases} 1, & : \quad m = 0. \\ 0, & : \quad \text{otherwise} \end{cases}$$

CDs of Sine and Cosine $(0 \leq a \leq \pi)$

$$CD_P(m, \sin(at)) = \begin{cases} (-1)^{(m-1)/2} p_m^P(a) & : \ m \text{ odd} \\ 0 & : \ m \text{ even} \end{cases}$$

$$CD_P(m, \cos(at)) = \begin{cases} (-1)^{m/2} p_m^P(a) & : \ m \text{ even} \\ 0 & : \ m \text{ odd} \end{cases}$$

CDs of Gaussian White Noise .

If signal f is Gaussian white noise with power ν, then $CD_P(m, f)(x)$ as x varies (varied at the bandlimit's Nyquist rate) are uncorrelated Gaussian random variables with variance ν.

[2]where $\binom{z}{w}_\nu = \frac{(\nu)_z}{(\nu)_w (\nu)_{z-w}}$ a "fractional z choose w" function and where $(\nu)_z = \frac{\Gamma(\nu+z)}{\Gamma(\nu)}$ is the Pochhammer symbol

6.4 Ultrabessel-Ultraspherical ($\nu > -\frac{1}{2}, \nu \neq 0$)

The ultrabessel-ultraspherical differential operators can be used for locally processing most useful, real-valued, bandlimited signals. Such operators include the previously mentioned families of bandlimited operators as special cases.

- Probably the most useful of the remaining families in the ultraspherical group of families is where ν is relatively large; such weights would be useful for processing and generating signals containing mostly low frequencies (but with some higher frequencies). Families with rather simple governing equations are those with $\nu = \frac{1}{2}n$ for n a positive integer.
- Another use of this class of families of differential operators is in signal processing: selecting the appropriate ν after figuring out the colored noise characteristics of the channel due to other interfering transmissions. In this case we would select the ultraspherical weight which would provide the closest match to what the water-filling version of Shannon's theorem would advise.

Frequency Domain Weight $\rho^{C^\nu}(\omega) = \frac{\nu\Gamma(\nu)^2}{\pi 2^{-2\nu}\Gamma(2\nu)}(1-(\frac{\omega}{\pi})^2)^{\nu-\frac{1}{2}}$, Range $[-\pi,\pi]$

Frequency Domain Polynomials $p_i^{C^\nu}(\omega) = \sqrt{\frac{(i+\nu)i!\Gamma(2\nu)}{\nu\Gamma(i+2\nu)}}C_i^\nu(\frac{\omega}{\pi})$, ($C_i^\nu$ - Ultraspherical Polynomials [3])

Time Domain Functions $B_i^{C^\nu}(t) = \left(\frac{J_{i+\nu}(\pi t)}{t^\nu}\right)\cdot\left(\frac{2^\nu(i+\nu)\Gamma(\nu)}{\pi^\nu}\right)\sqrt{\frac{\nu\Gamma(i+2\nu)}{i!(i+\nu)\Gamma(2\nu)}}$ (with $\left(\frac{J_{i+\nu}(\pi t)}{t^\nu}\right) = \mathcal{C}_i^\nu(t)$ the "Ultrabessel" functions)

Fourier Kernel Decomposition

$$e^{i\omega t} = \sum_{k=0}^{\infty} i^k p_k^{C^\nu}(\omega)B_k^{C^\nu}(t)$$

Chromatic Derivative Definition

$$\begin{aligned}
\mathrm{CD}_{C^\nu}(0,f) &= f \\
\mathrm{CD}_{C^\nu}(1,f) &= \sqrt{2(\nu+1)}\left(\frac{1}{\pi}\right)\frac{\partial f}{\partial t} \\
\mathrm{CD}_{C^\nu}(n,f) &= \sqrt{\frac{(n+\nu)(n-1+\nu)}{n(n-1+2\nu)}}\left(\frac{2}{\pi}\right)\frac{\partial}{\partial t}(\mathrm{CD}_{C^\nu}(n-1,f)) \\
&\quad + \sqrt{\frac{(n+\nu)(n-1)(n-2+2\nu)}{n(n-1+2\nu)(n-2+\nu)}}\mathrm{CD}_{C^\nu}(n-2,f),\ n\geq 2.
\end{aligned}$$

$$\text{Equivalently, } CD_{C^\nu}(k,f)(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} i^k p_k^{C^\nu}(\omega)\hat{f}(\omega)e^{i\omega t}d\omega.$$

Chromatic Derivative Operator Iteration

$$\mathrm{CD}_{C^\nu}(m,\mathrm{CD}_{C^\nu}(k,f)) = \sum_{i=0}^{\min(m,k)}(-1)^i a_{k,m,i}\cdot\mathrm{CD}_{C^\nu}(k+m-2i,f)\cdot \sqrt{\frac{(m+\nu)(k+\nu)m!k!\Gamma(2\nu)\Gamma(k+m-2i+2\nu)}{(k+m-2i+\nu)\nu(k+m-2i)!\Gamma(m+2\nu)\Gamma(k+2\nu)}}$$

[3] Also known as the Gegenbauer polynomials

for some constants $a_{k,m,i}$ computable from the Dougall-Rogers linearization formula: [4]

$$\begin{aligned} C_m^\nu(x)C_k^\nu(x) &= \sum_{i=0}^{\min(m,k)} \left(\frac{m+k+\nu-2i}{m+k+\nu-i}\right) \cdot \left(\frac{\binom{m+k-2i}{k-i}}{\binom{m+k-i}{k}_\nu}\right) \\ &\quad \left(\frac{\binom{m+k-i+2\nu-1}{i}}{\binom{k}{i}_\nu}\right) \cdot C_{m+k-2i}^\nu(x) \\ &= \sum_{i=0}^{\min(m,k)} a_{k,m,i} \cdot C_{m+k-2i}^\nu(x) \end{aligned}$$

And since

$$CD_{C^\nu}(k, B_0^{C^\nu}(t)) = (-1)^k B_k^{C^\nu}(t),$$

$$\begin{aligned} CD_{C^\nu}(k, B_m^{C^\nu}(t)) &= \sum_{i=0}^{\min(k,m)} (-1)^{i+k} a_{k,m,i}\, B_{k+m-2i}^{C^\nu}(t) \cdot \\ &\quad \sqrt{\frac{(m+\nu)(k+\nu)m!k!\Gamma(2\nu)\Gamma(k+m-2i+2\nu)}{(k+m-2i+\nu)\nu(k+m-2i)!\Gamma(m+2\nu)\Gamma(k+2\nu)}}. \end{aligned}$$

Chromatic Approximation Error Bound

$$\left|f(t) - \sum_{k=0}^{m} CD_{C^\nu}(k,f)(t_0)B_k^{C^\nu}(t-t_0)\right| \leq A\sqrt{1 - \sum_{k=0}^{m} B_k^{C^\nu}(t-t_0)^2}$$

where $A = \sqrt{\int_{-\pi}^{\pi} \frac{|\hat{f}(\omega)|^2}{\rho^{C^\nu}} d\omega}$ is a "windowed" energy.

Moving Between Ultraspherical CD Types Since it is the case that: [5]

$$\begin{aligned} C_n^\nu(x) &= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(\nu)_{n-k}(\nu-\mu)_k(n+\mu-2k)}{(\mu+1)_{n-k}k!\mu} C_{n-2k}^\mu(x) \\ &= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} b_{n,k}^{\nu,\mu} C_{n-2k}^\mu(x) \end{aligned}$$

then we see that:

$$\begin{aligned} \mathrm{CD}_{C^\nu}(n,f) &= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k b_{n,k}^{\nu,\mu}\, \mathrm{CD}_{C^\mu}(n-2k,f) \cdot \\ &\quad \sqrt{\frac{(n+\nu)\mu n!\Gamma(n-2k+2\mu)\Gamma(2\nu)}{(n-2k+\mu)\nu(n-2k)!\Gamma(n+2\nu)\Gamma(2\mu)}} \end{aligned}$$

[4]where $\binom{z}{w}_\nu = \frac{(\nu)_z}{(\nu)_w(\nu)_{z-w}}$ a "fractional z choose w" function; where $(\nu)_z = \frac{\Gamma(\nu+z)}{\Gamma(\nu)}$ is the Pochhammer symbol; and where $\binom{z}{w} = \frac{\Gamma(z+1)}{\Gamma(w+1)\Gamma(z-w+1)}$ is the ordinary generalized "z choose w" function.

[5]where $(x)_k = x \cdot (x+1) \cdots (x+k-1)$ for k a positive integer and $(x)_0 = 1$

7 Non-Bandlimited Signal Representation and Processing

7.1 Gaussian Monomials-Hermite

When considering the difficult problem of processing non-bandlimited signals, the Gaussian - Hermite H differential operators have many attractive features because of the fact that their sharply decreasing frequency-domain weight makes them somewhat bandlimited in practice. Thus, the operators are useful:

- whenever there are mainly bandlimited signals to be processed but where we need access to higher frequencies for a specific reason; e.g. for dealing with nonlinearities in transmission.
- when we are processing signals with a lot of bandlimited content but where higher frequencies must be dealt with as central to the signal; e.g. in video processing or stock price processing.
- where we need to process 2-dimensional or higher order signals, because the form of the time and frequency domain functions is very convenient for generalizations of the formulas given below.

Frequency Domain Weight $\rho^H(\omega) = 2\sqrt{\pi}e^{-\omega^2}$, Range $[-\infty, \infty]$

Frequency Domain Polynomials $p_n^H(\omega) = \frac{1}{\sqrt{2^n n!}} H_n(\omega)$, (Hermite Polynomials)

Time Domain Functions $B_n^H(t) = \frac{1}{\sqrt{2^n n!}} t^n e^{-t^2/4}$, (Gaussian Monomials)

Fourier Kernel Decomposition

$$e^{i\omega t} = \sum_{k=0}^{\infty} \left(\frac{1}{2^k k!}\right) i^k H_k(\omega) h_k(t)$$

Chromatic Derivative Definition

$$\begin{aligned} CD_H(0,f) &= f \\ CD_H(1,f) &= \sqrt{2}\frac{\partial f}{\partial t} \\ CD_H(n,f) &= \frac{1}{\sqrt{2(n-1)}}\frac{\partial}{\partial t}(CD_H(n-1,f)) + \\ &\quad \sqrt{\frac{(n-1)}{(n-2)}} \cdot CD_P(n-2,f), \text{where } n \geq 2. \end{aligned}$$

$$\text{Equivalently, } CD_H(k,f)(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} i^k p_k^H(\omega) \hat{f}(\omega) e^{i\omega t} d\omega.$$

Chromatic Derivative Operator Iteration

$$\mathrm{CD}_H(m, \mathrm{CD}_H(k,f)) = \sum_{i=0}^{\min(m,k)} (-1)^i a_{k,m,i} \cdot \sqrt{\frac{(m+k-2i)!}{m!k!4^i}} \mathrm{CD}_H(k+m-2i, f)$$

for some constants $a_{k,m,i}$ computable from the Nielsen linearization formula:

$$\begin{aligned} H_m(x)H_k(x) &= \sum_{i=0}^{Min(m,k)} \binom{m}{i}\binom{k}{i} 2^i i!\; H_{m+k-2i}(x) \\ &= \sum_{i=0}^{Min(m,k)} a_{k,m,i} H_{m+k-2i}(x) \end{aligned}$$

And since

$$CD_H(k, B_0^H(t)) = (-1)^k B_k^H(t),$$

$$CD_H(k, B_m^H(t)) = \sum_{i=0}^{\min(k,m)} (-1)^{i+k} a_{k,m,i} \sqrt{\frac{(m+k-2i)!}{m!k!4^i}}\, B_{k+m-2i}^H(t).$$

Chromatic Approximation Error Bound

$$\left| f(t) - \sum_{k=0}^{m} CD_H(k,f)(t_0) B_k^H(t-t_0) \right| \leq A \sqrt{1 - \sum_{k=0}^{m} B_k^H(t-t_0)^2}$$

where $A = \sqrt{\int_{-\pi}^{\pi} \frac{|\hat{f}(\omega)|^2}{\rho^H} d\omega}$ is a "windowed" energy.

7.2 SechTanh-Hyperbolic

The sechtanh-hyperbolic differential operator family can be quite useful when we need to process highly non-bandlimited signals, but where we do not need all of the power of wavelets and/or where we wish to retain the ability to process functions locally (using a fundamentally *local* approximation). Such operators can be used:

- to calculate rough, but wide-ranging, frequency spectrum values of unknown signals.
- to represent high frequency signals approximately using only (relatively) slow integrations to collect the interpolation coefficients.

Frequency Domain Weight $\rho^L(\omega) = \text{sech}\left(\frac{1}{2}\omega\right)$ Range $[-\infty, \infty]$

Frequency Domain Polynomials $p^L(\omega) = \mathcal{L}_i(\omega/\pi)$, where $\mathcal{L}_i$ are the hyperbolic polynomials which are orthogonal to weight $sech(\pi x/2)$, and generated by the recursion: [6]

$$\begin{aligned} \mathcal{L}_0(x) &= 1; \\ \mathcal{L}_1(x) &= -x; \text{ and} \\ \mathcal{L}_n(x) &= -\frac{1}{n} x \cdot \mathcal{L}_{n-1}(x) - \left(\frac{n-1}{n}\right) \mathcal{L}_{n-2}(x) \qquad [n \geq 2] \end{aligned}$$

Time Domain Functions $B_i^L(t) = (-1)^i \text{sech}(\pi t) \tanh^i(\pi t) = S_i(t)$ (hyperbolic secant and hyperbolic tangent)

Fourier Kernel Decomposition

$$e^{i\omega t} = \sum_{k=0}^{\infty} i^k \mathcal{L}_k(\omega/\pi) S_k(t)$$

[6] Compare to the classical Laguerre polynomials

Chromatic Derivative Definition

$$\begin{aligned}
\mathrm{CD}_{\mathcal{L}}(0,f) &= f(t)\\
\mathrm{CD}_{\mathcal{L}}(1,f) &= -\frac{1}{\pi}\frac{\partial}{\partial t}(f(t))\\
\mathrm{CD}_{\mathcal{L}}(n,f) &= -\frac{1}{\pi n}\frac{\partial}{\partial t}(\mathrm{CD}_{\mathcal{L}}(n-1,f))\\
&\quad+\left(\frac{n-1}{n}\right)\mathrm{CD}_{\mathcal{L}}(n-2,f)
\end{aligned}$$

$$\text{Equivalently, } CD_{\mathcal{L}}(k,f)(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} i^k \mathcal{L}_k(\omega/\pi)\hat{f}(\omega)e^{i\omega t}d\omega.$$

Chromatic Derivative Operator Iteration

$$\mathrm{CD}_{\mathcal{L}}(m,\mathrm{CD}_{\mathcal{L}}(k,f)) = \sum_{i=0}^{\min(m,k)}(-1)^i a_{k,m,i}\cdot \mathrm{CD}_{\mathcal{L}}(k+m-2i,f)$$

for some constants $a_{k,m,i}$ computable from the following recursion on m for the largest coefficients $k \leq k_{max}$ and $m \leq m_{max}$ we need to compute:

- Without loss of generality, we can assume that $k \leq m$ (because multiplication of polynomials is commutative); $a_{k,m,i} = a_{m,k,i}$.
- Let $a_{k,m,k} = \binom{m}{k}$ for all $0 \leq k \leq m$.
- Then, by induction on m (and on k) we define:

$$\begin{aligned}
a_{k,m,j} = a_{k-1,m,j} + a_{k-1,m-1,j-1} + a_{k,m-1,j} &\ :\ 0<j<k\\
a_{k,m,0} = a_{k-1,m,0} + a_{k,m-1,0} &\ :\ (j=0).
\end{aligned}$$

And since

$$CD_{\mathcal{L}}(k, B_0^L(t)) = (-1)^k B_k^L(t),$$

$$CD_{\mathcal{L}}(k, B_m^L(t)) = \sum_{i=0}^{\min(k,m)}(-1)^{i+k} a_{k,m,i}\, B_{k+m-2i}^L(t).$$

Chromatic Approximation Error Bound

$$\left|f(t) - \sum_{k=0}^{m}\mathrm{CD}_{\mathcal{L}}(k,f)(t_0)S_k(t-t_0)\right| \leq A\sqrt{1-\sum_{k=0}^{m}S_k^2(t-t_0)}$$

where $A = \sqrt{\int_{-\infty}^{\infty}\frac{|\hat{f}(\omega)|^2}{\rho^L(\omega)}d\omega}$ is the "exponentially windowed" energy of f.

8 The General Case Numerically

In this section we outline how to find such families of differential operators and their associated chromatic expansions when we are given only a real valued frequency domain weight to start with. Naturally, the task is not so straightforward as was the case with the classical orthogonal polynomials because there will not be closed-form formulas and definitions to refer to when building up the necessary library of tools we need to efficiently do local signal processing with chromatic derivatives.

Note that the suggestions below are not intended for real-time computation of families of chromatic derivatives; the computations are intended to be done off-line and will generally take a considerable amount of time to attain good numerical accuracy. However, given an even, real-valued weight with a bandlimit - an important case because it covers all real-valued time-domain bandlimited signals - it is possible to extract a procedure from those given below which produces good approximate frequency and time-domain functions and differential operators during a reasonably-short training period when using auxilliary polynomials to define families, as we suggest in the sequel.

8.1 Obtaining the Frequency-Domain Orthonormal Polynomials

When given a general real-valued weight, a measure on L^2 over the frequency domain, from $+\infty$ to $-\infty$ (or a restricted frequency range, if given a bandlimited weight), we follow the steps below to produce the corresponding system of orthonormal polynomials.

It is useful here to note that for any set of polynomials $p_i(x)$ orthogonal to a given real-valued weight $w(x)$ - $\int_{-\infty}^{+\infty} p_i(x)p_j(x)w(x)dx = \delta_{i,j}$ - such a system can be generated recursively from $p_0(x)$ and $p_1(x)$ by knowing the constants in the recursion equation: $p_{n+1}(x) = (a_n x + b_n)p_n(x) + c_n p_{n-1}(x)$ (where $c_n > 0$ and all constants are real) that accompanies every such sequence of orthogonal polynomials. We aim to find such a set of polynomials through integration (using the weight) using standard techniques of orthogonalization and from which we can extract the above recursion constants.

8.1.1 Basic method

1. First, normalize the weight: $\int_{-\infty}^{+\infty} w(x)dx = 1$.
2. Then compute the following integrals for all of the positive integers $n > 0$ up to $2N$ for which one cares to know differential operator up to the N-th order: $m_n = \int_{-\infty}^{+\infty} x^n w(x)dx$. Such values must exist for the weight to be a suitable one for generating differential operators, and if computed once then the following steps are all just algebraic. If the weight is even ($w(x) = w(-x)$ for all x) then we only have to integrate the weight with even monomials.
3. Define an inner product on the space $\mathbb{R}^{2N}$, which represents the space of polynomial coefficients, as follows: $< \vec{a}, \vec{b} >= \sum_{i,j=0}^{N} a_i b_j m_{i+j}$.
4. Then, in the general case in which the weight is not an even function use the modified Gram-Schmidt method with respect to the above inner product, starting with the unit vectors $< 0, 0, \ldots, 0, 1, 0, \ldots, , 0, 0 >$, to generate the orthogonal polynomial coefficients for n-order polynomials ($n > 1$). At the end we can orthonormalize the polynomials easily by again using the inner product.
5. In the case where $w(x)$ is an even function, we note that this means that we can break the modified Gram-Schmidt process down in to two parts (even and odd order polynomials), which could be useful if N is large: the modified Gram-Schmidt process takes on the order of $(N + 1)^3$ flops.

Thus, we will have determined both the orthonormal polynomials and the recursion equations which determine the chromatic derivatives; the CD recursion equation is taken from the polynomial recursion equation by just inserting some "i's" and replacing the "x"'s by partial derivatives.

8.1.2 Alternatives

There are two apparent modifications to the above approach. First, one can substitute auxiliary orthogonal polynomials for the monomials which one integrates to determine the values m_n. This option can provide superior numerical accuracy, especially in the bandlimited case, if we select the auxiliary orthogonal polynomials whose weights (to which these polynomials are orthogonal) are close to $w(x)$. Then, proceeding as above, we end up with a set of coefficients which tell us which linear mixture of the auxiliary polynomials form polynomials orthogonal with respect to $w(x)$. Finally, it is easy to transform these "polynomials in polynomials" into ordinary polynomials since we can, using the recursion generation equation of the auxiliary orthogonal polynomials, recursively build a matrix which converts linear combinations of auxiliary polynomials into ordinary polynomials. [7]

The second alternative we can employ to calculate orthogonal polynomials is if we are computing the polynomials using symbolic computation software, such as Mathematica, and the weight $w(x)$ is in relatively tractable analytic form. In this case we compute the m_n as above, symbolically, along with the associated inner product, and just use the following three-term recursion directly after defining $p_0(x) = 1$ and $p_1(x) = x - m_1$:

$$p_{n+1}(x) = \left(x - \left(\frac{< xp_n(x), p_n(x) >}{< p_n(x), p_n(x) >}\right)\right) p_n(x) - \left(\frac{< p_n(x), p_n(x) >}{< p_{n-1}(x), p_{n-1}(x) >}\right) p_{n-1}(x)$$

This will produce orthogonal polynomials which can then be orthonormalized using the inner product. [8]

8.2 Obtaining the Time-Domain Chromatic Expansion Functions

In the general case there is not much to be done in this step except to bite the bullet and compute the time-domain functions by performing a Fourier transform for each basis function:

$$B_n(t) = \int_{-\infty}^{+\infty} i^n p_n(\omega) w(\omega) e^{i\omega t} d\omega$$

This will need to be approximated by a limit, numerically, if there is no bandlimit on the weight. If there is a bandlimit on the weight we might be able to perform an FFT on the polynomials times the weight, though in that case we would have to check on the performance of the procedure if the weight will produce functions of infinite energy.

However, given that we will want to produce time-domain functions which obey the equation: $CD(n, B_0) = (-1)^n B_n$, if it is possible to compute just B_0 using a Fourier transform and then approximate this very well using functions which are repeatedly differentiable (especially if they are symbolically differentiable), [9] then we could use the recursive form of the CD definition to compute the higher order functions from the 0th-order time-domain function.

---

[7] For example, if $p_0(x) = 1$, $p_1(x) = a_0 x$, and $p_{n+1}(x) = a_n x p_n(x) - b_n p_{n-1}(x)$ then we can inductively show that if $x^n = \sum_{m=0}^{n} \gamma_m^n p_m(x)$, then $x^{n+1} = \left(\frac{\gamma_n^n}{a_n}\right) p_{n+1}(x) + \sum_{m=1}^{n} \left[\left(\frac{\gamma_{m-1}^n}{a_{m-1}}\right) + \left(\frac{b_{m+1}\gamma_{m+1}^n}{a_{m+1}}\right)\right] p_m(x) + \left(\frac{\gamma_1^n b_1}{a_1}\right)$ where each term contributes to the sum where its subscripts are > 0.

[8] E.g. see J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, 2nd Ed., Springer Verlag, 1993, section 3.6 for this form of Gram-Schmidt orthogonalization.

[9] differentiable in the ordinary sense, of course

8.2.1 Time Domain Basis Functions Using Auxilliary Orthogonal Polynomials

Another option available to us for creating the basic time-domain functions is when we have generated the orthogonal polynomials for the weight $w(x)$ by first computing the linear combinations of some auxiliary orthogonal polynomials which are orthogonal with respect to $w(x)$:

1. First, compute the linear combinations of some auxiliary orthogonal polynomials which are orthogonal with respect to $w(x)$. [10]

2. Then, the next step is to approximate $\frac{w(x)}{\rho(x)}$ using linear combinations of those same auxiliary polynomials, where ρ is the auxiliary polynomial's weight.

3. Finally, computing basis functions $B_k(t) = i^k \widehat{w(x)p_k}(x)$ associated with the frequency domain weight $w(x)$ can be done by linearly mixing together the basis functions corresponding to the auxiliary orthogonal polynomials because everything in the above formula is a linear combination of auxiliary polynomials (which can be linearized) and the auxiliary polynomials' weight.

8.3 Obtaining the Other Formulas

The only other fundamental equation for which we need to generate coefficients is the CD iteration formula, a formula very useful in adaptive filtering and modulation among other areas. All that is lacking are the coefficients given by the linearization formula associated with the orthogonal polynomials:

$$p_m(x)p_k(x) = \sum_{i=0}^{2Min(m,k)} d_{k,m,i}p_{k+m-i}(x)$$

Such linear coefficients are always guaranteed to exist: we can solve for the linearization coefficients from the polynomial coefficients by just using the equation above, multiplying out the left-hand side, and solving for the $d_{k,m,i}$ in terms of the polynomial coefficients. However, such sets of equations tend to be poorly conditioned and so a more stable solution method should be sought. Note that even for the classical orthogonal polynomials, producing closed forms of the linearization equations are non-trivial undertakings, and even then some of those forms are not optimal for numerically calculating the coefficients we need. [11]

However, there is a recursive way to generate these tables using a recursion formula invented by Richard Askey. [12]

1. We can rearrange the standard recursion formula $p_{n+1}(x) = (a_n x + b_n)p_n(x) + c_n p_{n-1}(x)$ by requiring it to be renormalized so that the leading coefficient of each polynomial is always 1. Then, we can redo the basic recursion to say (renaming all constants):

   $$p_1(x)p_n(x) = p_{n+1}(x) + a_n p_n(x) + b_n p_{n-1}(x)$$

   This has the form of a linearization formula - the easiest non-trivial case.

[10] See the above subsection.

[11] Note that in solving for the linearization formulas, we usually do not care about the positivity of the coefficients $d_{k,m,i}$, a property which is crucial for the standard use of these formulas which is associated with polynomial approximations.

[12] Richard Askey, "Linearization of the Product of Orthogonal Polynomials" in *Problems in Analysis*, ed. Robert C. Gunning, Princeton University Press, 1970

2. Then, it is fairly easy to show, by twice using the above basic, "linearized" recursion formula, that

$$p_{l+1}p_n - p_l p_{n+1} = (a_n - a_l)p_n p_l + (b_n - b_l)p_l p_{n-1} + b_l[p_l p_{n-1} - p_{l-1}p_n]$$

and to compute the $[p_l p_{n-1} - p_{l-1}p_n]$ term by using this recursion formula again. Notice how the indices of the polynomials will keep drifting downward while the form of the recursion stays constant as we keep reusing it in order to get down to the basic ("easy") recursion formula above.

3. Thus, we can use this patterned recursion to compute tables of as many linearization coefficients as needed by starting with $l = 1$ and computing the coefficients from the original polynomial recursion for many n. Then, we use the second recursion to induct (via l) our way up, all the while collecting in tables the coefficients $d_{k,m,i}$ which fall out of the recursion. Such a procedure is bound to be very numerically stable provided that our orthogonal polynomial recursion coefficients are accurate.

4. Finally, we need to transform these high-order term-normalized (to 1) orthogonal polynomial linearization constants into linearization constants for the orthonormal polynomials we created. But this is easy to do by computing the constants α_k which orthonormalize each high-order term-normalized p_k polynomials and then noticing that

$$p_m(x)p_k(x) = \sum_{i=0}^{2Min(m,k)} \frac{\alpha_{k+m-1}}{\alpha_k \alpha_m} d'_{k,m,i} p_{k+m-i}(x)$$

where p_k are the orthonormalized polynomials and $d'_{k,m,i}$ are the constants we found using the above Askey linearization recursion.

References

[1] A. Ignjatovic, Numerical Differentiation and Signal Processing, Kromos Technology Technical Report.

[2] W. Rudin, *Functional Analysis.* Boston, MA: McGraw-Hill, 1973.

[3] G. Szegö, *Orthogonal Polynomials.* Providence, RI: American Mathematical Society, 1939,1975.