ORTHOGONAL POLYNOMIALS WITH RESPECT TO SELF-SIMILAR MEASURES

STEVEN M. HEILMAN, PHILIP OWRUTSKY, AND ROBERT S. STRICHARTZ

Abstract. We study experimentally systems of orthogonal polynomials with respect to self-similar measures. When the support of the measure is a Cantor set, we observe some interesting properties of the polynomials, both on the Cantor set and in the gaps of the Cantor set. We introduce an effective method to visualize the graph of a function on a Cantor set. We suggest a new perspective, based on the theory of dynamical systems, for studying families $P_n(x)$ of orthogonal functions as functions of $n$ for fixed values of $x$.

1. Introduction

The classical theory of orthogonal polynomials [Gautschi 2004][Szegő 1975] allows you to start the Gram-Schmidt process with virtually any measure. In the families of polynomials that are usually studied, the measure is either absolutely continuous or discrete, but the general theory allows one to use a singular continuous measure. In recent years there has been an interest in the case of fractal measures. See for example [Barnsley et al. 1983]-[Barnsley et al. 1985],[Bessis and Moussa 1983],[Mantica 1996]-[Mantica 2000]. In particular, G. Mantica has developed algorithms to efficiently compute the coefficients of the 3-term recursion relation, and hence the polynomials, in the case of a self-similar measure [Mantica 1996][Mantica 2000]. In this work we use these algorithms as a tool to study the polynomials experimentally, looking for interesting patterns and conjectures. We view this as part of a general program to explore portions of classical analysis related to fractal measures.

Let $\mu$ be a measure on the line, and for simplicity assume that $\mu$ is a probability measure supported on the unit interval. The related system $\{P_n(x)\}$ of orthogonal polynomials is characterized as follows:

$P_n(x)$ is a polynomial of degree $n$ with a non-negative coefficient of $x^n$, (1)

$$\int P_n(x)P_m(x)d\mu(x) = \delta_{n,m},$$

orthonormality. (2)
The general theory implies that there is a 3-term recursion relation
\[ xP_n(x) = r_{n+1}P_{n+1}(x) + A_n P_n(x) + r_n P_{n-1}, \]
where \( r_n, A_n > 0, n \in \mathbb{Z}^+ \) are determined by the measure \( \mu \) and \( P_{-1} = 0, P_0 = 1 \).
The coefficients \( r_n, A_n \) are also called the entries of the Jacobi matrix, an infinite symmetric tri-diagonal matrix \( J \) such that \( J_{i,i} = A_i \) and \( J_{i,i+1} = J_{i+1,i} = r_{i+1} \) for \( i \in \mathbb{Z}^+ \). Note that Eq. (3) allows us to find the polynomials recursively:
\[ P_{n+1}(x) = \frac{1}{r_{n+1}}((x - A_n)P_n(x)) - \frac{r_n}{r_{n+1}}P_{n-1}(x) \tag{4} \]
We should point out that while in principle Eq. (4) allows computation of the coefficients of \( P_n \), this computation could be unstable. In the case that \( A_n \) is constant (as we see below when our measure is supported on a Cantor set), it is often easier to compute the coefficients of \( P_n \) as polynomials of \( (x - A_n) \). When the \( A_n \) vary, we alternatively use Eq. (4) to compute the values \( P_n(x) \) for specific \( x \)-values.

A measure \( \mu \) on the line is said to be self-similar if there exists an iterated function system (IFS) of contractive similarities \( \{F_i\}_{i=1}^N \) and a set of probability weights \( \{p_i\}_{i=1}^N \) such that
\[ \mu(A) = \sum_{i=1}^N p_i \mu(F_i^{-1}A), \text{ for any measurable set } A, \text{ or equivalently,} \]
\[ \int f \, d\mu = \sum_{i=1}^N p_i \int f \circ F_i \, d\mu, \text{ for any continuous function } f. \tag{6} \]
In this paper we restrict our attention to the family of IFS’s with \( N = 2 \) and such that
\[ F_1(x) = \frac{1}{R} x, \quad F_2(x) = \frac{1}{R} (x - 1) + 1 \tag{7} \]
where \( R \geq 2 \) is a parameter. When \( R = 2 \) and \( p_1 = p_2 = \frac{1}{2} \) we obtain Lebesgue measure on \([0,1]\), and the corresponding polynomials are essentially the classical Legendre polynomials. (Actually, the classical Legendre polynomials are orthogonal with respect to Lebesgue measure on \([-1,1]\) and are normalized differently, but the differences just involve rescaling the axes.) When \( R = 2 \) and \( p_1 \neq p_2 \), we refer to the corresponding polynomials as Weighted Legendre Polynomials (WLP). The measure \( \mu \) is singular but not supported on any proper closed subset of \([0,1]\). When \( R > 2 \) we will always take \( p_1 = p_2 = \frac{1}{2} \), and we call the corresponding polynomials Cantor Legendre Polynomials (CLP). The measure \( \mu \) is then supported on a Cantor set \( C_R \) characterized by
\[ C_R = F_1 C_R \cup F_2 C_R \tag{8} \]
The standard Cantor set and Cantor measure correspond to \( R = 3 \). We refer to the intervals in \([0,1] \setminus C_R \) as gaps. The largest gap is the interval \( \left( \frac{1}{R}, 1 - \frac{1}{R} \right) \), and there are \( 2^m \) gaps of length \( \frac{1}{2^{m+1}} \left( 1 - \frac{2}{R} \right) \).

The behavior of \( P_n(x) \) in the CLP case is quite different on the gaps and on the Cantor set \( C_R \). In order to visualize the graphs of \( P_n(x) \) on \( C_R \) we introduce the distorted Cantor set \( \tilde{C}_{R,\epsilon} \) (also known as the Smith-Volterra-Cantor set or fat Cantor set), obtained by reducing the size of the gaps by a factor of \( \epsilon \) (a parameter that we choose). Note that \( \tilde{C}_{R,\epsilon} \) is still a topological Cantor set, but it has positive Lebesgue measure. There is an obvious one-to-one correspondence between \( C_R \) and
that identifies regions between corresponding gaps. We use this identification to graph functions defined on $C_R$ against $\tilde{C}_{R,\epsilon}$.

In section 2 we present data for the entries of the Jacobi matrix. In the CLP case we note the different behaviors of $r_n$ for even and odd $n$, and the occurrence of small values. In section 3 we display graphs of the polynomials. In the CLP case we show graphs of the restrictions to the Cantor set and to the gaps.

We then discuss various features of the data. In section 4 we discuss the associated Dirichlet kernels. By using the Christoffel-Darboux formula we are able to relate approximate identity behavior with small values of $r_n$. In section 5 we discuss some approximate equalities relating CLP restricted to the Cantor set. In particular, $P_{2n}(x)$ is approximately equal to $P_{2n+1}(x)$ on the right half of $C_R$, for large $n$. We also define a “shuffle” map that approximately preserves $P_{2n+1}$ when $n$ is a power of 2. In section 6 we discuss the behavior of $P_n(x)$ on the gaps in the CLP case. On the central gap, for high $n$, $P_n(x)$ vaguely approximates either a Gaussian or the derivative of a Gaussian, depending on $n \mod 2$. More precisely, we find that $P_{2n}(x) = c_{2n}e^{-d_{2n}x^2}$ where $\alpha(x) = \alpha_{2n}(x) = 2$ at $x = 1/2$, and $\alpha(x) \approx 2$ for $x$ around $1/2$. On the other gaps, for large enough $n$, the behavior is roughly the same. In section 7 we discuss the behavior of $P_n(x)$ at the points $x = 0, 1/4, 1/2, 3/4, 1$ in the WLP case, and we contrast the results with the known behavior for Legendre polynomials. This study leads to the dynamical systems perspective.

Instead of thinking of $P_n(x)$ as a function of $x$ for fixed $n$, we look at $P_n(x)$ as a function of $n$, for a fixed $x$. Because of the 3-term recursion relation, it is more natural to look at vectors $(P_n(x), P_{n+1}(x))$ in the plane. Then there exist $2 \times 2$ matrices $M_n(x)$ such that

\[
\begin{pmatrix}
P_{n+1}(x) \\
P_n(x)
\end{pmatrix} = M_n(x) \begin{pmatrix}
P_n(x) \\
P_{n-1}(x)
\end{pmatrix},
\]

In fact,

\[
M_n(x) = \begin{pmatrix}
x - A_n & -r_n \\
r_{n+1} & 1
\end{pmatrix}.
\]

So, we are really looking at orbits of a time-dependent linear dynamical system. We can then define the analogy of the Mandelbrot set as all $x$ for which the orbit is bounded. It appears that this coincides, more or less, with the support of the measure. Then for each $x$ in this Mandelbrot set, we can define a Julia set $J(x)$ as the limit set of the orbit. We display some examples of these Julia sets. We then extend the investigation of section 7 to generic $x$ for WLP and CLP in section 8.

We end with a short concluding discussion in Section 9.


Acknowledgement: We are grateful to Giorgio Mantica for allowing us to use his codes.
2. Entries of the Jacobi Matrix

The coefficients $r_n$ and $A_n$ in the 3-term recursion relation (Eq. (3)) determine the polynomials in a rather subtle way. More work is needed to clarify this relationship. In this section we report data for our two classes of examples. In Figs. 1-4 we graph $A_n$ and $r_n$ versus $n$, for several choices of $p_1$ in the WLP case. For classical Legendre polynomials we have $A_n = 1/2$ for all $n$ by symmetry about $x = 1/2$, and $\lim_{n \to \infty} r_n = 1/4$. The WLP case shows a small but significant difference from this model case.

![Figure 1. Plot of $A_n$ vs. index $n$, $1 \leq n \leq 10000$ for $p_1 = .6$](image1)

![Figure 2. Plot of $A_n$ vs. index $n$, $1 \leq n \leq 10000$ for $p_1 = .7$](image2)

In the CLP case, all $A_n = 1/2$ by symmetry about $x = 1/2$. In Fig. 5 we graph $r_n$ versus $n$ for $R = 8$.

**Problem 2.1**: What is the nature of the limit set

$$\cap_{N=1}^{\infty} \text{cl} \left( \cup_{n=N}^{\infty} \{r_n\} \right)?$$

Do the measures

$$\frac{1}{m-k} \sum_{n=k+1}^{m} \delta_{r_n}$$


converge weakly to some fractal measure as $m, k$ tend to infinity in some specific manner?

A striking feature of the data is the different behavior of $r_n$ for $n$ even and $n$ odd. In Figs. 3 and 4 we show the same data as Fig. 2, separating the even and odd values of $n$. Another striking observation is that some values of $r_n$ for $n$ even are close to zero.

**Problem 2.2:** What is
\[
\liminf_{n \to \infty} r_n?
\] (13)

In particular, is the lim inf zero? What is the sequence of $n$’s along which the lim inf is attained?

As we will see later, having values of $r_n$ close to zero has interesting implications. We could also ask for the lim sup, but it is not clear what significance this has.

**Conjecture 2.3:** We always have
\[
r_n \leq \frac{1}{2}
\] (14)
Figure 5. Plot of $r_n$ vs. index $n$, $1 \leq n \leq 10000$ for CLP, $R = 8$

Figure 6. Plot of $r_n$ vs. index $n$, $1 \leq n \leq 10000$ for $n$ odd, CLP, $R = 8$

Figure 7. Plot of $r_n$ vs. index $n$, $1 \leq n \leq 1000$ for $n$ even, CLP, $R = 8$
3. Graphs of Polynomials

In Fig. 8 we show the graphs of $P_n(x)$ for $1 \leq n \leq 5$ for the WLP with $p_1 = 1/2$, so these are just (rescaled versions of) the classical Legendre Polynomials, computed using G. Mantica’s algorithm. In Figs. 9 and 10 we show the same functions for $p_1 = .6$ and $p_1 = .7$. Already we observe that the symmetry is broken in a decisive fashion, as these functions are much larger near $x = 1$ (where the measure “has less weight”) than near $x = 0$ (where the measure “has more weight”). This observation is expected since the Gram-Schmidt process with respect to $\mu$ almost immediately implies the following property: $P_n(x)$ is the unique $n^{th}$ degree polynomial, with highest degree coefficient $d_n$, which minimizes the $L^2([0,1],\mu)$ norm. (In the case $p_1 = 1/2$, we have $d_n = \prod_{i=1}^{n} \frac{1}{b_i}$ and $1/b_i = \sqrt{16 - 4/i^2}$.) Therefore, $P_n$ is expected to be the smallest where $\mu$ “has more support.” In Fig 11 we show the graphs of $P_n(x)$ for $49 \leq n \leq 52$ for $p_1 = .7$. The same graphs are shown in Fig. 12 with the $y$-axis truncated to more clearly display the structure of the functions.

Conjecture 3.1: For fixed $p_1 > 1/2$ and any given $\epsilon \geq 0$ the WLP are uniformly bounded on the interval $[0, 1 - \epsilon]$.

Next we look at the CLP case, where certain features of the $P_n(x)$ increase in complexity. In Figs. 13 and 14 we show the graphs of $P_n(x)$ for $1 \leq n \leq 5$ on the whole interval for two different choices of $R$. In Fig. 15 we show the graph of $P_{52}(x)$ on the whole interval for $R = 4$. Note that the values on the central gap are so large that no information about the graph on the complement of the central gap is discernable using linear scaling in the $y$-axis. This is typical of $P_n(x)$ for large even values of $n = 2k$ and any $R > 2$ (though the sign of $P_n(x)$ changes sign between $n = 4j + 2$ and $n = 4j$). To view these polynomials more effectively, we scale the $y$-axis logarithmically, but also multiply by the sign of $P_n(x)$. That is, we plot $\text{sign}(P_n(x)) \log(|P_n(x)|)$ vs $x$. In Fig. 16 we show the graph of $P_{51}(x)$ for $R = 4$, which is typical of $P_n(x)$ for $n$ odd and large. With the same logarithmic scaling, Fig. 17 shows $P_{52}(x)$ for $R = 4$ restricted to the gap $[1/R^2, \frac{1}{R^2}] = [1/16, 3/16]$. This behavior is typical of $P_n(x)$ for large $n$. Note that the maximum value is quite...
Figure 9. Plot of $P_n(x)$ on the unit interval with $p_1 = .6, p_2 = .4$. Recall that the 60 percent weighting is given to the map which contracts towards zero.

Figure 10. Plot of $P_n(x)$ on the unit interval with $p_1 = .7, p_2 = .3$.

Figure 11. Plot of $P_n(x)$ on the unit interval with $p_1 = .7, p_2 = .3$ for $49 \leq n \leq 52$. 
Figure 12. Plot of $P_n(x)$ on the unit interval with $p_1 = .7, p_2 = .3$ for $49 \leq n \leq 52$ with the y-axis restricted to the range $-10 \leq y \leq 10$.

large, but still it is small relative to the maximum value on the central gap. We will discuss this behavior more in Section 6.

Figure 13. Plot of first five Cantor Legendre Polynomials, $R = 3$.

Next, we graph CLP on the Cantor set $C_R$ using a distorted Cantor set $\tilde{C}_{R,\epsilon}$ for the $x$-axis. In Figs. 18 and 19 we display the same functions as in Figs. 13 and 14. Already we see that the values of $P_n(x)$ are considerably smaller on the Cantor set. In Fig. 20 we show $P_n(x)$ restricted to the Cantor set for $97 \leq n \leq 100$ and $R = 8$. We will comment in detail about some of the structure of those graphs in Section 5. Many more examples may be viewed at [Owrutsky 2005].
Figure 14. Plot of first five Cantor Legendre Polynomials, $R = 4$. As in the Cantor Legendre Case, the polynomials have lowest absolute value where the measure has the most weight (as expected).

Figure 15. Plot of $P_{52}(x)$ for CLP $R = 4$, logarithmically scaled on the entire interval. Notice the large values obtained by the polynomials off the Cantor set ($10^{28}$) compared to the behavior on the Cantor set (for $1 \leq j \leq n$ large, $|P_n| < 20$). The overall shape in the gap is Gaussian, and this occurs for all $n = 4k$, $k \in \mathbb{Z}^+$. For $n = 4k + 2$, the behavior in the gap is roughly a negative Gaussian.
Figure 16. Plot of $P_{51}(x)$ for CLP $R = 4$, logarithmically scaled on the entire interval. This is roughly the derivative of a Gaussian on the center gap. For all $n = 4k + 3$ we observe this shape, and for $n = 4k + 1$, the negative of this shape.

Figure 17. Plot of $P_{52}(x)$ for CLP $R = 4$, logarithmically scaled on the secondary gap $[1/R, (R - 1)/R] = [1/16, 3/16]$. Again, we observe a Gaussian shape.
Figure 18. Plot of first five Cantor Legendre Polynomials, $R = 3$.

Figure 19. Plot of first five Cantor Legendre Polynomials, $R = 4$

Figure 20. Plot of some larger $n$ Cantor Legendre Polynomials, $R = 8$, $97 \leq n \leq 100$. 
4. Dirichlet Kernels

For a general function $f$ in $L^2(d\mu)$, we can expand it as a series

$$
\sum_{n=0}^{\infty} a_n P_n(x),
$$

(15)

where the coefficients are given by

$$
a_n = \int f(y) P_n(y) \, d\mu(y).
$$

(16)

The partial sums

$$
\sum_{n=0}^{N} a_n P_n(x)
$$

(17)

may be represented as an integral

$$
\int D_N(x,y) f(y) \, d\mu(y)
$$

(18)

where the Dirichlet kernel is given by

$$
D_n(x,y) = \sum_{n=0}^{N} P_n(x) P_n(y)
$$

(19)

The partial sums in Eq. (17) converge to $f$ in $L^2$ norm, but to get better convergence we need to know more about the Dirichlet kernel. In view of related results in [Strichartz 1993-2, Strichartz 1994], one might hope that there exists a sequence $\{N_k\}$ along which the partial sums converge uniformly if $f$ is continuous. This would follow by standard approximate identity arguments if we could show that the quantity

$$
\int |D_{N_k}(x,y)| \, d\mu(y)
$$

(20)

is uniformly bounded, and that

$$
\lim_{k \to \infty} \int_{|y-x| \geq \epsilon} |D_{N_k}(x,y)| \, d\mu(y) = 0
$$

(21)

for all $\epsilon > 0$. While we have no insight on how to establish Eq. (20), we can say something about Eq. (21), thanks to the Christoffel-Darboux formula

$$
D_N(x,y) = \frac{r_{N+1}}{x-y} (P_{N+1}(x)P_N(y) - P_N(x)P_{N+1}(y)), \text{ for } x \neq y
$$

(22)

(It is easy to derive Eq. (22) by multiplying Eq. (19) by $(x-y)$ and using the 3-term recursion relation Eq. (3).) Assuming that the polynomials $P_n(x)$ are uniformly bounded on the support of $\mu$, if we could find a sequence $N_k$ such that $r_{N_k+1} \to 0$, then Eq. (21) follows from Eq. (22).

It appears from our data that for CLP polynomials there exist indices $n$ such that $r_n$ is close to zero, but there is no evidence for a sequence tending to zero. This means that there will be some Dirichlet kernels that seem very concentrated near the diagonal (those with $r_{N+1}$ close to zero), but it is unlikely that we can improve this behavior indefinitely.

Figs. 21 and 22 illustrate Dirichlet kernels $D_N(\cdot,y)$ for fixed $y$ that are not concentrated near $x=y$, while Figs. 23 and 24 illustrate Dirichlet kernels that are moderately well concentrated.
Figure 21. Dirichlet kernel for CLP with $n = 60$ and $R = 8$ centered at the left most point in the right half of the Cantor set. We see that the graph is reasonably small away from the center.

Figure 22. Dirichlet kernel for CLP with $n = 89$ and $R = 8$ centered at the left most point in the right half of the Cantor set. Again, the kernel is small away from its center.
Figure 23. Dirichlet kernel for CLP with $n = 50$ and $R = 8$ centered at the left most point in the right half of the Cantor set. This kernel takes much larger values away from the center than the $n = 60$ case in Fig. 21.

Figure 24. Dirichlet kernel for CLP with $n = 99$ and $R = 8$ centered at the left most point in the right half of the Cantor set.
5. Approximate Equalities for CLP Polynomials

In this section we discuss two types of approximate equalities in the CLP case. First we note that by symmetry, $P_n(x)$ is even under the reflection $x \mapsto 1-x$ when $n$ is even and odd when $n$ is odd. Nevertheless, the plots of $P_{2n}(x)$ and $P_{2n+1}(x)$ appear very similar (see $n = 98, 99$ in Fig. 20). This is especially true on the right half of the Cantor set. In Fig. 25 we graph $P_{50}(x)$ and $P_{51}(x)$ for $R = 8$. The qualitative similarity is striking, but it is difficult to quantify. Fig. 26 shows the graph of the difference, and Fig. 27 shows the graph of the ratio. We can give a rough explanation using Eq. (4), which for even values may be written

$$P_{2n+1}(x) = \frac{x - 1/2}{r_{2n+1}} P_{2n}(x) - \frac{r_{2n}}{r_{2n+1}} P_{2n-1}(x).$$

(23)

For large values of $R$, we may have $r_{2n}$ close to zero and $r_{2n+1}$ close to $1/2$. So the second term on the right side of Eq. (23) is close to zero. But on the right half of the Cantor set, $x$ is close to one, so the coefficient $(x-1/2)/r_{2n+1}$ is close to one. Therefore, Eq. (23) says that $P_{2n+1}(x) \approx P_{2n}(x)$ on the right half of $C_R$. By the odd-even behavior, we have $P_{2n+1}(x) \approx -P_{2n}(x)$ on the left half of $C_R$. Since we see a qualitative similarity of the plots on all of $C_R$, we must attribute these approximate equalities to an approximate reflectional symmetry across the $x$-axis of the graphs of all $P_n(x)$. This is roughly apparent in Fig. 20, but it does not hold up to close inspection. In particular, the graphs of $P_{2n+1}(x)$ and $P_{2n+1}(1-x)$ are not that close (of course $P_{2n+1}(1-x) = -P_{2n+1}(x)$ exactly).

The second type of approximate equality refines the above idea. If we express the points of $C_R$ as infinite binary decimals, then $x \mapsto (1-x)$ simply interchanges all digits in the binary expansion. Let $T_m(x)$ denote the map $C_R \to C_R$ that interchanges the first $m$ binary digits, leaving all other digits unchanged. In other words, $T_m$ permutes the $2^m$ Cantor subsets of level $m$ by reversing the order of the subsets. In Fig. 28 we show the plots of $P_{17}(x)$ and $P_{17}(T_3(x))$ for $R = 8$. There is clearly a strong qualitative fit, but the agreement of the two functions is not very close numerically. The same pattern persists for $P_{2m+1}(x)$ and $P_{2m+1}(T_m(x))$ for all values of $R$. At present we have no explanation for this phenomenon.
Figure 26. $P_{80}(x) - P_{81}(x)$, CLP $R = 8$

Figure 27. $P_{80}(x)/P_{81}(x)$, CLP $R = 8$

Figure 28. $P_{17}(x)$ and $P_{17}(T_3(x))$, $R = 8$
6. CLP Polynomials on Gaps

We saw in Figs. 15 and 17 that the graph of $P_{2n}(x)$ on the gaps in $C_R$ is approximately Gaussian, for $n$ large. In reality, we find that $P_{2n}|_{[1/R, (R−1)/R]}(x) = c_{2n} \exp \left(-d_{2n} x^\alpha(x)\right)$ where $\alpha(x) = \alpha_{2n}(x) = 2$ at $x = 1/2$, and $\alpha(x) \approx 2$ for $x$ around $1/2$. Fig. 29 shows $\alpha(x) = \alpha_{2n}(x)$ for $n = 50$, CLP $R = 4$. Due to symmetry, we only view $\alpha_{2n}$ on the interval $[1/R, 1/2]$. From the data, we conjecture that $\alpha_{2n}$ converges as $n \to \infty$ on any closed interval of the form $[1/R + \epsilon, 1/2]$ to a function resembling that in Fig. 29.

The data also indicate exponential growth (in $n$) of $P_{2n}(x)$ on the central gap $[1/R, (R−1)/R]$. This exponential growth results from the behavior of Eq. (4) at $x = 1/2$ (with $A_n = 1/2$ for all $n$) as follows:

$$P_{2n}(1/2) = (-1)^n \prod_{j=1}^n \left(\frac{r_{2j−1}}{r_{2j}}\right).$$

(24)

Numerically, we find that $(r_{2j−1}/r_{2j})$ is bounded above 1 with few exceptions. In the CLP $R = 2.5$ case, there are 134 instances that $(r_{2j−1}/r_{2j}) \leq 1$ for $j = 1, \ldots, 5000$, in the $R = 3$ case there is one, and for the $R = 4$ case there are none. Thus, as $R$ increases, $(r_{2j−1}/r_{2j})$ is bounded above 1 more consistently, as we have seen in the $R = 8$ case in Figs. 6 and 7. Eq. (24) therefore (generally) gives exponential growth in $n$ for $|P_{2n}(1/2)|$.

We then summarize the behavior of $P_{2n}(x)$ on the gap $[1/R, (R−1)/R]$ as follows: $c_{2n} = P_{2n}(1/2)$ grows exponentially (according to Eq. (24)), $d_{2n}$ grows linearly, and $\alpha_{2n}(x)$ appears to converge (to something resembling Fig. 29). In fact, we can say more: for $k \in \mathbb{Z}^+$, write

$$P_j(x) = \sum_{k=0}^j \alpha_{j,k}(x - 1/2)^k.$$  

(25)

Eq. (24) gives all $\alpha_{j,0}$ (observe that $\alpha_{2n,0} = c_{2n}$ and $\alpha_{2n+1,0} = 0$ for all $n$), and then we can again use Eq. (4) to compute all of the $\alpha_{j,k}$. (Here we read Eq. (4) as an equation of polynomials.) If we view the coefficients $\{\alpha_{j,k}\}$ as a (lower triangular) matrix, then we see that the left column $\{\alpha_{j,0}\}$ and the diagonal $\{\alpha_{j,j}\}$ determine...
all of the other $\alpha_{j,k}$, via Eq. 4. Thus, the behavior of the left column and diagonal (and the $r_j$ coefficients) dictates the behavior of the other $\alpha_{j,k}$.

The left column and diagonal grow log-linearly, as we see in Figs. 30 and 31 and Table 1. Note the similarity in the three error plots of Figs. 30 and 31. If we label the (top) even indexed error $\Phi_{\text{even}}$, the (bottom) odd error $\Phi_{\text{odd}}$, and the (even) error function from Fig. 30 $\Psi_{\text{even}}$, we have the following approximate equalities for $k = 0, \ldots, 50$: $\Phi_{\text{even}}(2k) \approx \Phi_{\text{odd}}(2k+1) + 1.36 \approx \Psi_{\text{even}}(2k) + .67$. Also, from Table 1 we see that $\text{sign}(\alpha_{j,k}) = -\text{sign}(\alpha_{j,k+2})$, for appropriate $j,k$. For the column and the diagonal of the matrix of $\alpha_{j,k}$, the slope of each best fit line varies with log $R$, as we see in Fig. 32.

![Figure 30](image)

**Figure 30.** Logarithmically scaled left column $\{\alpha_{j,0}\}$ of polynomial coefficients, with error, CLP $R = 8$. Here we only plot $j$ even since for odd $j$, $\alpha_{j,0} = 0$.

![Figure 31](image)

**Figure 31.** Logarithmically scaled diagonal $\{\alpha_{j,j}\}$ of polynomial coefficients, with error, CLP $R = 8$.

On the other gaps (i.e. the images of $[1/R, (R-1)/R]$ under compositions of the maps $F_1$ and $F_2$), we observe similar qualitative behavior. That is, on these smaller gaps, $P_n(x)$ resembles a Gaussian for all large and even $n$. We have seen this self-similar behavior already in Figs. 15 and 16. However, the local extremum
Figure 32. Slopes of best fit lines for $\alpha_{j,j}$ and $\alpha_{j,0}$ (as in Figs. 30 and 31), for $R \in \{2.5, 3, 4, 8, 16, 32\}$.

<table>
<thead>
<tr>
<th>$\alpha_{j,k}/10^4$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.00001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0.00022678</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>-0.0010002</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0.0048457</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>-0.067032</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0.0013607</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>-0.027586</td>
<td>0</td>
<td>0.43733</td>
<td>0</td>
<td>-2.2537</td>
<td>0</td>
<td>3.789</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0.069611</td>
<td>0</td>
<td>1.0873</td>
<td>0</td>
<td>-5.5105</td>
<td>0</td>
<td>9.1999</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>1.4436</td>
<td>0</td>
<td>-31.645</td>
<td>0</td>
<td>254.75</td>
<td>0</td>
<td>-891.62</td>
<td>0</td>
<td>1146.2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>3.0921</td>
<td>0</td>
<td>-67.772</td>
<td>0</td>
<td>545.53</td>
<td>0</td>
<td>-1909.1</td>
<td>0</td>
<td>2454.1</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>-7.2086</td>
<td>0</td>
<td>191.07</td>
<td>0</td>
<td>-1996.6</td>
<td>0</td>
<td>10284</td>
<td>0</td>
<td>-26134</td>
<td>0</td>
<td>26237</td>
</tr>
</tbody>
</table>

Table 1. CLP $R = 8$, Matrix of (rounded) $\alpha_{j,k}$ coefficients, divided by $10^4$

Figure 33. $x$-value of local extremum of $P_2(x)$ on $[1/R^2, (R - 1)/R^2] = [1/16, 3/16]$ as a function of $n$, CLP $R = 4$.

of $P_{2n}(x)$ on $[1/R^2, (R - 1)/R^2]$ is not in general the center of the given gap. We glimpse this phenomenon in Fig. 17. In Fig. 33 we plot the local extremum of $P_{2n}|_{[1/16,3/16]}$ in the CLP $R = 4$ case. The center of this gap occurs at $x = .125$, but the local extremum fluctuates with average around $.122$.  


7. WLP as a function of $n$ for fixed $x$

For the classical Legendre polynomials on $[0,1]$, we find structure in the values of $P_n(x)$ as functions of $n$ for fixed $x$. As an example, for $x = 0$, $P_n(x) = (-1)^n \sqrt{2n+1}$, and for $x = 1$, $P_n(1) = \sqrt{2n+1}$. The computations of $P_n(0)$ and of $P_n(1)$ are shown at the top of Fig. 34. These algebraic relations are a renormalization of the behavior of the classical Legendre polynomials $\tilde{P}_n$ on $[-1,1]$, where we recall that $\tilde{P}_n(1) = 1, \tilde{P}_n(-1) = (-1)^n$ and $||\tilde{P}_n||_{L^2[-1,1]} = 1/(n+1/2)$ for $n \geq 1$.\footnote{By a generic point we mean a point chosen, with respect to a uniform distribution, among a suitable set of (rational) values in floating point arithmetic. In all cases, we only treat $x$ representable in double precision floating point arithmetic.} \cite{Gautschi 2004}. We now increase the WLP weight $p_1$ at the points $x = 0,1$. As a result, we see a perturbation of the behavior of the classical Legendre polynomials (on $[0,1]$) in Fig. 34. For $p_1 \approx 7$, a transition seems to occur in the behavior of $\{P_n(0)\}_{n=1}^{\infty}$. At this transition, we appear to have a multiplicative periodic function $P_n(0)$, i.e. a function $f: \mathbb{Z}^+ \rightarrow \mathbb{R}$ where for some $a > 0$ we have $f(ax) = f(x)$ for all $x$. For $p_1 = 7$, we can logarithmically scale the $x$-axis as in Fig. 35 to see a nearly multiplicative periodic function. For other $p_1$, we evidently have a multiplicative periodic function $P_n(0)/n^\beta$ for an appropriate choice of $\beta = \beta(p_1)$.

For generic $x \in [0,1]$, if we plot the vectors $(P_n(x), P_{n+1}(x))$ in the WLP $p_1 = .5$ (i.e. classical) case, then these vectors are attracted to an ellipse as $n \rightarrow \infty$. This ellipse is centered at the origin, and its axes and orientation vary with $x$. Also, the vectors $(P_n(x), P_{n+1}(x))$ rotate around the ellipse in a periodic way. That is, for a certain $k > 1$ which depends on $x$ (where $k \in \mathbb{Z}$ is the period), $d((P_n(x), P_{n+1}(x)), (P_{n+k}(x), P_{n+k+1}(x))) < \epsilon$ for some small $\epsilon > 0$. Geometrically speaking, $(P_n(x), P_{n+1}(x), n) \in \mathbb{R}^3$ travels along a helix near the surface of an ellipsoidal cylinder. Thus, plotting $(P_n(x), P_{n+1}(x))$ projects this helix onto the plane. In Section 8 we focus on the behavior of $(P_n(x), P_{n+1}(x))$ for generic $x \in [0,1]$. For now, we merely note the contrast between the images of generic and non-generic points.

As observed for $x \in \{\{0\}, \{1\}\}$ (a set of two non-generic points), increasing $p_1$ perturbs the dynamics of the vectors $(P_n(x), P_{n+1}(x))$. Therefore, increasing $p_1$ should result in perturbed dynamics for other non-generic $x$. To this end, consider $x \in \{\{.25\}, \{.5\}, \{.75\}\}$. For the classical Legendre polynomials on $[0,1]$, these $x$ values yield finite attractors for $\{(P_n(x), P_{n+1}(x))\}_{n=1}^{\infty}$. These attractors have three, four and six points, respectively. We call the corresponding number of points $k_x$. The vectors $(P_n(x), P_{n+1}(x))$ travel in a clockwise fashion about the $k_x$ attracting points, as $n$ increases. When we increase $p_1$, the dynamics of $(P_n(x), P_{n+1}(x))$ for $x \in \{\{.25\}, \{.5\}, \{.75\}\}$ change dramatically, as we see in Figs. 36 and 37. Instead of cycling ever closer to $k_x$ attracting points, the vectors $(P_n(x), P_{n+1}(x))$ cycle through $k_x$ fractal spiral arms, as we see in Fig. 37. However, in the case that $p_1 < .5$, the attractor of these spiral arms is difficult to determine. By measuring the distance of $(P_n(x), P_{n+1}(x))$ from $(0,0)$, it seems that the spirals are bounded away from the origin as $n \rightarrow \infty$. Therefore, by the symmetry apparent from Figs. 36 and 37, $(P_n(x), P_{n+1}(x))$ should be attracted to a (fractal) ellipse, or a finite set of $k_x$ points. As a further contrast to the $p_1 = .5$ behavior, we see from Fig. 37 that $(P_n(x), P_{n+1}(x), n) \in \mathbb{R}^3$ travels along a “fractal helix” with radius decreasing in $n$.\footnote{By a generic point we mean a point chosen, with respect to a uniform distribution, among a suitable set of (rational) values in floating point arithmetic. In all cases, we only treat $x$ representable in double precision floating point arithmetic.}
Figure 34. $P_n(x)$ values vs. $n$ for $x = 0, 1$ plotted for the WLP $p_1 = .5, .6, .7, .8, .9$ families.
Figure 35. $P_n(0)$ values vs. $n$ logarithmically scaled in the $x$-axis for WLP $p_1 = .7$.

Figure 36. $P_n(x)$ values vs. $n$ for $x = .75, .5, .25$ plotted for the WLP $p_1 = .6, .7, .8$ families. Note that for $p_1 = .5$, we would get solid horizontal lines.
Figure 37. Vectors \((P_n(x), P_{n+1}(x))\) for \(x = .75, .5, .25\), plotted for the WLP \(p_1 = .6, .7, .8\) families.
8. WLP and CLP Dynamics

As promised, we now examine the dynamics of $P_n(x)$ for generic $x$. Here we examine both WLP and CLP, and we summarize the results in Figure 38. For WLP, we recall for $p_1 = .5$ and fixed, generic $x \in [0, 1]$, there exists an integer $k_x > 1$ so that

$$d((P_n(x), P_{n+1}(x)), (P_{n+k_x}(x), P_{n+k_x+1}(x))) < \epsilon$$ (26)

for some small $\epsilon > 0$. Here we mean that the sequence of points 
$$\{(P_{n+jk_x}(x), P_{n+jk_x+1}(x))\}_{j=0}^{\infty}$$ is periodic of period $k_x$, up to a small error of $\epsilon$ at each step. We therefore make $k_x$ the minimal positive integer satisfying our condition 26. If we plot $P_n(x)$ vs. $n$, we find that $P_n(x)$ is a superposition of $k_x$ sinusoidal functions (see entry (1, 1) in Fig. 38). These $k_x$ (approximately) periodic functions are given by $P_{(k_x,n+j)}(x)$ for $j \in \{1, 2, \ldots, k_x\}$.

For example, we can see that $k_{95} = 14$ by counting the number of distinct periodic functions in the plot of $P_n(x)$ versus $n$. Since we essentially have a superposition of phase shifted cosines, it follows that the distribution of values of $P_n(x)$ is the function $1/\sqrt{1-y^2}$, suitably rescaled (see entry (2, 1) in Fig. 38). Finally, if we color the iterates $(P_n(x), P_{n+1}(x))$ in a 14-periodic manner, we get entry (3, 1) in Fig. 38. In this plot, the point $(P_{n+1}(x), P_{n+2}(x))$ overlaps all points of lower index.

So, what happens when we increase $p_1$? As in Section 7 we observe a perturbation of the $p_1 = .5$ behavior. This perturbation becomes more exaggerated the larger $p_1$ becomes. In the middle column of Fig. 38, we show the same plots described in the previous paragraph, for $p_1 = .6$. We again choose $x = .95$ so $k_{95} = 14$, and we observe fourteen vaguely periodic functions. In other words, for $j \in \{1, 2, \ldots, k_x\}$, we see that $P_{(k_x,n+j)}(.95)$ is a perturbed sinusoidal function of $n$. We can see these perturbations via the distribution function (entry (2, 2)) and the plot of $(P_n(x), P_{n+1}(x))$ (entry (3, 2)) in Fig. 38. That is, the distribution of values of $P_n(x)$ is a perturbed version of the function $1/\sqrt{1-y^2}$, suitably rescaled. Also, the attractor of the vectors $(P_n(x), P_{n+1}(x))$ appears to be a thickened (fractal) ellipse.

The behavior for generic $x$ in the CLP case is displayed in Fig. 35 in the right-most column. As an example, we view the iterates of $x = 1/5 = 1/(R+1) \in C_R$ for the case $R = 4$. (Recall that $P_n(x)$ grows exponentially for $x \in [0, 1] \setminus C_R$, so we expect more information by viewing $x \in C_R$.) For generic $x$, the distribution of $P_n(x)$ values is unclear (Fig. 38, entry (2, 3)). Nevertheless, the $P_n(x)$ values do seem concentrated around the origin. Also, the vectors $(P_n(x), P_{n+1}(x))$ obey a 2-periodic behavior. Specifically, the points $\{(P_{2n}(x), P_{2n+1}(x))\}_{n=1}^{\infty}$ cluster along the line $\{(x, y) \in \mathbb{R}^2 : y = -x\}$ (to abuse notation). Also, the points $\{(P_{2n+1}(x), P_{2n+2}(x))\}_{n=1}^{\infty}$ cluster loosely around the origin. To observe this behavior, note entry (3, 3) in Fig. 38.

In our investigation, we have found only one non-generic point $x \in C_R \cap [0, 1/2]$ for the CLP case (other than the endpoint $x = 0$). (Note that the symmetry of these $P_n(x)$ across $x = 1/2$ reduces our investigation to $C_R \cap [0, 1/2]$.) Unlike the generic iterates which demonstrate only 2-periodic behavior, the non-generic point $x = 1/R$ exhibits 4-periodic behavior. The columns of Fig. 39 correspond to different $p_1$ values and the iterates of $x = 1/R$. For $x = 1/R$, the vectors $(P_{4n+j}(x), P_{4n+j+1}(x))$ form four disjoint attractors for $j \in \{1, 2, 3, 4\}$. The set $\{(P_{4n+j}(x), P_{4n+j+1}(x))\}_{n=0}^{\infty}$ for $j = 1$ resides in the lower left quadrant, for $j = 2$
Figure 38. Summary of Generic point values: \( p_n \) values, their histograms, and the vectors \((p_n(x), p_{n+1}(x))\) plotted for the Legendre, WLP \( p_1 = .6 \) and CLP, \( R = 4 \) polynomial families.

This set lives in the upper left quadrant, for \( j = 3 \) the upper right, and for \( j = 4 \) the lower right.
Figure 39. $P_n$ values, their histograms, and the vectors $(P_n(x), P_{n+1}(x))$ plotted for $x = 1/R$, CLP $R = 2.5, 3.4$ and $8$ polynomial families.
Concluding Discussion

Experimental mathematics allows us to see patterns that otherwise would never be observed. At times these patterns allow us to formulate explicit conjectures, and eventually proofs may be found to turn conjectures into theorems. But often the story is more complicated. The observed patterns may turn out, on closer inspection, to be only approximately true. No clear cut conjectures emerge that are truly supported by the experimental evidence. In such cases, what should we do?

We could simply discard the experimental results and move on. But the approximate patterns that we observe may be of great interest. The experiment might be trying to tell us something that we are not quite able to capture in the conventional format of mathematical statements. Experimental science is often messy in exactly this way, and we think it would be a shame to limit experimental mathematics to only the clean paradigm.

We see the results reported in this paper in exactly this light. We do have clean cut conjectures on the $r_n$ values in Section 2 and boundedness for WLP in Section 3, but beyond that we only have messy evidence. We see some CLP Dirichlet kernels that look more like approximate identities than most Dirichlet kernels. We are able to relate this to the small size of the $r_n$ coefficients. The experimental evidence suggests that we will not find a sequence with $r_n$ converging to zero. So we cannot offer a conjecture analogous to the results of [Strichartz 2006] about uniform convergence of certain special partial sums of the orthogonal polynomial expansion of an arbitrary continuous function. However, the results of [Strichartz 2006] suggest that perhaps such a statement might hold only for special values of $R$, and we have only examined a few choices of $R$.

The approximate equalities for CLP discussed in Section 5 appear very striking when one looks at the graphs of the polynomials. It is on closer inspection that one observes the deviation from equality. A clean conjecture might say that the deviation goes to zero in some limit, but the evidence does not really support such a conjecture. Nevertheless, we find the approximate equalities intrinsically interesting.

Similarly, a glance at the graphs of CLP on the gaps suggests they are close to being Gaussian. Of course one comes to expect Gaussian limits in many mathematical situations. But here in Section 6, on closer inspection, we see a decided deviation from the Gaussian model. In this case we believe there is a true limit law; we just have not been able to find it.

Perhaps the most interesting contribution of this paper is the dynamical perspective discussed in Sections 7 and 8. Here we may be excused from offering explicit conjectures on the grounds that the pictures offer a view of significant structures that were previously unrecognized. We hope that others will be inspired to investigate the dynamics perspective in a systematic way.
References


