BIRKHOFF POLYNOMIAL INTERPOLATION WITH APPLICATIONS TO DIFFERENTIAL EQUATIONS

By Ryan Kelly

A Thesis

Submitted in Partial Fulfillment

of the Requirements for the Degree of

Master of Science

in Mathematics

Northern Arizona University

May 2023

Approved:

John Neuberger, Ph.D., Chair

James W. Swift, Ph.D.

Nandor Sieben, Ph.D.

ABSTRACT

BIRKHOFF POLYNOMIAL INTERPOLATION WITH APPLICATIONS TO DIFFERENTIAL EQUATIONS

RYAN KELLY

An error formula for Birkhoff interpolation of functions on $\mathbb{R}^s$ is developed. Some applications of the error formula to Birkhoff quadrature methods and finite difference methods are presented. In one dimension, the formula is used to find optimal placements of interpolation nodes that maximize the local rate of convergence of the interpolating polynomial to the interpolated function. An application of Birkhoff interpolation in approximating solutions of first order initial value problems is noted.

Contents

List of Tables

Chapter 1

Introduction

In George Birkhoff's 1906 paper [2], a generalized Mean Value Theorem for configurations of derivative and function values is established for a function in one real variable. Since this original paper, there has been much development on the existence and uniqueness of large classes of Birkhoff interpolation polynomials such as in [1, 4, 5, 10, 12, 14, 13, 17], and others. A large collection of results on univariate Birkhoff interpolation may be found in [11].

This thesis is primarily concerned with an apparently novel Birkhoff polynomial interpolation error formula of functions in one real variable, presented in Chapter 2. The main interpolation error formula has applications in the error calculation of Birkhoff quadrature methods and finite difference methods. An immediate corollary of Theorem 2.2.1 is the calculation of asymptotic order of convergence of interpolation methods, along with the corresponding asymptotic error constants.

The algebraic nature of the main error theorem in one dimension allows for its extension to multivariate polynomial interpolation, presented in Chapter 3, and to multi-polynomial (pre-spline) interpolation, presented in Chapter 4.

Chapter 2

Birkhoff Polynomial Interpolation on $\mathbb{R}$

This chapter studies Birkhoff polynomial interpolation of a function on one real variable, f. A Birkhoff polynomial P is a polynomial function on one real variable which agrees with f at some finite collection of function values and finite order derivative values. A classic question of interest is on the existence and uniqueness of interpolation polynomials. This chapter is primarily concerned with the calculation of the error of interpolation at a particular derivative location, the optimal node placements that minimize this error, and the corresponding asymptotic error as the interval of interpolation tends to zero.

2.1 Preliminary Definitions and Results

In one real variable, a function is interpolated through n nodes, for some natural number n. Given a function $f\colon \mathbb{R} \to \mathbb{R}$, its k_i-th derivative will be fit at x_i for the ordered collection real values $\mathbf{x} = (x_1, \ldots, x_n)$ and non-negative integers $\mathbf{k} = (k_1, \ldots, k_n)$. To make statements about existence and uniqueness of interpolation polynomials it is natural to consider a polynomial that fits these conditions at a collection of real constants $\mathbf{c} = (c_1, \ldots, c_n)$.

Definition 2.1.1. A *real Birkhoff Polynomial through* $(\mathbf{x}, \mathbf{k}, \mathbf{c})$ is a polynomial $P\colon \mathbb{R} \to \mathbb{R}$ satisfying

$$P^{(k_i)}(x_i) = c_i, \qquad 1 \leq i \leq n.$$

If additionally $c_i = f^{(k_i)}(x_i)$ for each $i \in \{1, \ldots, n\}$, then P *interpolates* f *at* $(\mathbf{x}, \mathbf{k})$. Throughout this thesis, we utilize notation $f^{(0)} := f$.

Throughout this thesis, we may refer to a tuple (x_i, k_i) as an interpolation node, and the tuple $(\mathbf{x}, \mathbf{k})$ as a collection of n nodes. The existence of Birkhoff Polynomials for a fixed collection of nodes $(\mathbf{x}, \mathbf{k})$ is necessary to establish interpolation error.

Definition 2.1.2. A collection of n nodes $(\mathbf{x}, \mathbf{k})$ is *consistent of degree* m if for every tuple $\mathbf{c}$ of n real constants, there exists a Birkhoff Polynomial of degree at most m through $(\mathbf{x}, \mathbf{k}, \mathbf{c})$, and m is the minimum such non-negative integer. The collection of n nodes $(\mathbf{x}, \mathbf{k})$ is *consistent* if it is consistent of degree $n-1$.

Since the interpolation conditions of Definition 2.1.1 are linear, it is clear that a collection of nodes is consistent if and only if the corresponding Birkhoff polynomial is unique for each set of interpolation constants. Given a collection of nodes that are consistent of degree m, it is useful to have a basis for interpolation.

Definition 2.1.3. Suppose the collection of n nodes, $(\mathbf{x}, \mathbf{k})$, is consistent of degree m, then a *standard interpolation basis* is an ordered set of polynomials $(L_1, \ldots, L_n)$ that satisfy

$$L_i^{(k_j)}(x_j) = \delta_{ij}, \qquad 1 \leq i, j \leq n,$$

such that L_i is of degree at most m for each $i \in \{1, \ldots, n\}$.

Here, and throughout this thesis, δ_{ij} is used to denote the function

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

The existence of a standard interpolation basis is given by the definition of consistency, and the observation that each L_i is a Birkhoff Polynomial through $(\mathbf{x}, \mathbf{k}, (\delta_{ij})_{j=1}^n)$. The uniqueness of a standard interpolation basis is given only when $(\mathbf{x}, \mathbf{k})$ is consistent. When $(\mathbf{x}, \mathbf{k})$ is consistent of degree $m > n-1$, then a standard interpolation basis is not unique, as there is a family of polynomials that are of degree less than or equal to m that are zero at all of the interpolation conditions. For any sufficiently differentiable function f there exists a polynomial in the span of a standard interpolation basis that interpolates f at $(\mathbf{x}, \mathbf{k})$.

Proposition 2.1.4. *Suppose that* $(\mathbf{x}, \mathbf{k})$ *is consistent of degree* m*,* $(L_1, \ldots, L_n)$ *is a standard interpolation basis, and* $f^{(k_i)}(x_i)$ *exists for each* $i \in \{1, \ldots, n\}$*. Then*

$$P(x) = \sum_{i=1}^{n} f^{(k_i)}(x_i) L_i(x)$$

interpolates f *at* $(\mathbf{x}, \mathbf{k})$*, and is of degree at most* m*.*

Proof. Let $j \in \{1, \ldots, n\}$ and P be given as above. Then

$$P^{(k_j)}(x_j) = \sum_{i=1}^{n} f^{(k_i)}(x_i) L^{(k_j)}(x_j)$$

$$\begin{aligned} &= \sum_{i=1}^{n} f^{(k_i)}(x_i)\delta_{ij} \\ &= f^{(k_j)}(x_j). \end{aligned}$$

Additionally, by definition, L_i is of degree at most m for each $i \in \{1, \dots, n\}$. $\square$

It is well known that the existence and uniqueness of Birkhoff Interpolation Polynomials is invariant with respect to scaling and translation of the interpolation positions $\mathbf{x}$ [11]. A constructive proof is provided here.

Proposition 2.1.5. *Suppose* $(\mathbf{x}, \mathbf{k})$ *is consistent of degree* m*. Let* a *and* h *be real numbers such that* h *is nonzero. Let* $a_i = x_i - a$ *for every* $i \in \{1, \dots, n\}$ *and let* $\mathbf{x}_a(h) = (a + a_1h, \dots, a + a_nh)$*. For every* $i \in \{1, \dots, n\}$*, let* $L_{a,h,i} \colon \mathbb{R} \to \mathbb{R}$ *be defined by*

$$L_{a,h,i}(x) = h^{k_i} L_i\Big(a + \frac{x-a}{h}\Big)$$

for $x \in \mathbb{R}$*. Then* $(L_{a,h,1}, \dots, L_{a,h,n})$ *is a standard interpolation basis for* $(\mathbf{x}_a(h), \mathbf{k})$*, for every* $a \in \mathbb{R}$ *and* $h \in \mathbb{R} \setminus \{0\}$*, if and only if* $(L_1, \dots, L_n)$ *is a standard interpolation basis for* $(\mathbf{x}, \mathbf{k})$*.*

Proof. Suppose $(L_{a,h,1}, \dots, L_{a,h,n})$ is a standard interpolation basis for $(\mathbf{x}_a(h), \mathbf{k})$, for every $a \in \mathbb{R}$ and $h \in \mathbb{R} \setminus \{0\}$. Then when $a = 0$ and $h = 1$, observe that $\mathbf{x}_0(1) = \mathbf{x}$ and $L_{0,1,i} = L_i$ for every $i \in \{1, \dots, n\}$. Hence $(L_i, \dots, L_n)$ is a standard interpolation basis for $(\mathbf{x}, \mathbf{k})$.

Conversely, suppose $(L_i, \dots, L_n)$ is a standard interpolation basis for $(\mathbf{x}, \mathbf{k})$. Then direct computation yields

$$\begin{aligned} L_{a,h,i}^{(k_j)}(a + a_jh) &= \frac{d^{k_j}}{dx^{k_j}} \left(h^{k_i} L_i\Big(a + \frac{x-a}{h}\Big)\right) \Bigg|_{x=a+a_jh} \\ &= L_i^{(k_j)}\Big(a + \frac{x-a}{h}\Big)\Bigg|_{x=a+a_jh} \\ &= L_i^{(k_j)}(x_j) \\ &= \delta_{ij} \end{aligned}$$

for every $i, j \in \{1, \dots, n\}$. Hence the property of forming an interpolation basis of some specific degree is scale and translation invariant. Thus, the minimality of an interpolation basis must also be preserved under scaling and translation. Therefore $(L_{a,h,1}, \dots, L_{a,h,n})$ is a standard interpolation basis for $(\mathbf{x}_a(h), \mathbf{k})$, for every $a \in \mathbb{R}$ and $h \in \mathbb{R} \setminus \{0\}$.

$\square$

We now recall Taylor's Theorem using the Lagrange form of the remainder term; for a proof, see a standard text in analysis, such as [3] or [16]. The main error theorem of this thesis heavily relies on the use of Taylor polynomials and on their remainder. Thus, a notation for Taylor polynomials that depend on the degree, the function, and center are given.

Definition 2.1.6. Suppose f is m times differentiable at $a \in \mathbb{R}$, then the *degree m Taylor Polynomial of f centered at a* is defined by

$$T_{f,m,a}(x) = \sum_{k=0}^{m} f^{(k)}(a)\frac{(x-a)^k}{k!}.$$

Theorem 2.1.7. *(Taylor's Theorem - Lagrange Remainder) Let $x, a \in \mathbb{R}$ and suppose that f is $m+1$ times differentiable on an open interval containing x and a, then there is some $\xi(x)$ between a and x such that*

$$f(x) = T_{f,m,a}(x) + f^{(m+1)}(\xi(x))\frac{(x-a)^{m+1}}{(m+1)!}.$$

Definition 2.1.8. The *Lagrange Form of the Taylor Remainder* is defined by

$$R_{f,m,a}(x) = f(x) - T_{f,m,a}(x) = f^{(m+1)}(\xi(x))\frac{(x-a)^{m+1}}{(m+1)!}$$

for some $\xi(x)$ between a and x, as given by Theorem 2.1.7.

It will be convenient to have a corresponding Taylor's theorem of the derivative(s) of a function f.

Corollary 2.1.9. *Let $l, m \geq 0$, $a \in \mathbb{R}$ and suppose f satisfies the hypothesis of Theorem 2.1.7. Then for every $x \in \mathbb{R}$*

(i) $T_{f^{(l)},m-l,a}(x) = T^{(l)}_{f,m,a}(x)$ *and*

(ii) $f^{(l)}(x) = T^{(l)}_{f,m,a}(x) + R_{f^{(l)},m-l,a}(x)$.

Proof. Part (i) follows from a few steps of algebra:

$$T_{f^{(l)},m-l,a}(x) = \sum_{k=0}^{m-l} f^{(l+k)}(a)\frac{(x-a)^k}{k!}$$

$$
\begin{aligned}
&= \sum_{k=l}^{m} f^{(k)}(a)\frac{(x-a)^{k-l}}{(k-l)!} \\
&= \frac{d^l}{dx^l}\sum_{k=l}^{m} f^{(k)}(a)\frac{(x-a)^{k}}{k!} \\
&= \frac{d^l}{dx^l}\sum_{k=0}^{m} f^{(k)}(a)\frac{(x-a)^{k}}{k!} \\
&= T^{(l)}_{f,m,a}(x).
\end{aligned}
$$

Part (ii) follows directly from Part (i) and Theorem 2.1.7 in that

$$
\begin{aligned}
f^{(l)}(x) &= T_{f^{(l)},m-l,a}(x) + R_{f^{(l)},m-l,a}(x) \\
&= T^{(l)}_{f,m,a}(x) + R_{f^{(l)},m-l,a}(x).
\end{aligned}
$$

□

2.2 Interpolation Error

A novel error formula for the Birkhoff polynomial interpolation is stated as the main result of this thesis. Then, the formula is applied to calculate a remainder polynomial and consequently to calculate the corresponding asymptotic error.

Theorem 2.2.1. *Let* $(\mathbf{x}, \mathbf{k})$ *be consistent of degree* m*, let* $(L_1, \ldots, L_n)$ *be a standard interpolation basis of* $(\mathbf{x}, \mathbf{k})$*, let* $l \geq 0$*,* $x \in \mathbb{R}$ *such that* $(x, l) \neq (x_i, k_i)$ *for every* $i \in \{1, \ldots, n\}$ *and let* $a \in \mathbb{R}$*. Suppose* r *is the least non-negative integer such that there is a degree* $r+1$ *polynomial* R *that satisfies*

$$R^{(k_i)}(x_i) = 0 \qquad 1 \leq i \leq n,$$

but $R^{(l)}(x) = 1$*. Let* $f\colon \mathbb{R} \to \mathbb{R}$ *be* $r+1$ *times differentiable on an open interval containing* x*,* a*, and* x_i *for each* $i \in \{1, \ldots, n\}$*. If* $m < r+1$ *and*

$$P = \sum_{i=1}^{n} f^{(k_i)}(x_i) L_i,$$

then

$$f^{(l)}(x) - P^{(l)}(x) = f^{(r+1)}(\xi(x))\frac{(x-a)^{r+1-l}}{(r+1-l)!} - \sum_{i=1}^{n}\left[f^{(r+1)}(\xi_i(x_i))\frac{(x_i-a)^{r+1-k_i}}{(r+1-k_i)!}\right] L_i^{(l)}(x)$$

for some $\xi(x)$ *between* a *and* x*, and some* $\xi_i(x_i)$ *between* a *and* x_i *for each* $i \in \{1, \ldots, n\}$*.*

Proof. Let $P_T\colon \mathbb{R} \to \mathbb{R}$ be the degree at most m polynomial that interpolates $T_{f,r,a}$ defined by

$$P_T = \sum_{i=1}^{n} T^{(k_i)}_{f,r,a}(x_i) L_i.$$

Note that since $r+1 > m$, $P_T^{(l)}(x) = T^{(l)}_{f,r,a}(x)$, otherwise

$$\frac{P_T - T_{f,r,a}}{P_T^{(l)}(x) - T^{(l)}_{f,r,a}(x)}$$

would be a polynomial of degree at most r that is zero at $(\mathbf{x}, \mathbf{k})$ but whose l-th derivative is 1 at x, which contradicts the assumption that $r+1$ is the smallest such value. Thus

$$\begin{aligned}
P^{(l)}(x) &= \sum_{i=1}^{n} f^{(k_i)}(x_i) L_i^{(l)}(x) \\
&= \sum_{i=1}^{n} \left[T_{f^{(k_i)}, r-k_i, a}(x_i) + R_{f^{(k_i)}, r-k_i, a}(x_i) \right] L_i^{(l)}(x) \\
&= \sum_{i=1}^{n} T_{f^{(k_i)}, r-k_i, a}(x_i) L_i^{(l)}(x) + \sum_{i=1}^{n} R_{f^{(k_i)}, r-k_i, a}(x_i) L_i^{(l)}(x) \\
&= \sum_{i=1}^{n} T^{(k_i)}_{f,r,a}(x_i) L_i^{(l)}(x) + \sum_{i=1}^{n} R_{f^{(k_i)}, r-k_i, a}(x_i) L_i^{(l)}(x) \\
&= P_T^{(l)}(x) + \sum_{i=1}^{n} R_{f^{(k_i)}, r-k_i, a}(x_i) L_i^{(l)}(x) \\
&= T^{(l)}_{f,r,a}(x) + \sum_{i=1}^{n} R_{f^{(k_i)}, r-k_i, a}(x_i) L_i^{(l)}(x) \\
&= f^{(l)}(x) - R_{f^{(l)}, r-l, a}(x) + \sum_{i=1}^{n} R_{f^{(k_i)}, r-k_i, a}(x_i) L_i^{(l)}(x).
\end{aligned}$$

So

$$f^{(l)}(x) - P^{(l)}(x) = R_{f^{(l)}, r-l, a}(x) - \sum_{i=1}^{n} R_{f^{(k_i)}, r-k_i, a}(x_i) L_i^{(l)}(x),$$

and substituting explicit remainder terms from Definition 2.1.8 gives the desired result. ☐

Remark 2.2.2. If $(\mathbf{x}, \mathbf{k})$ is consistent, then $m = n-1 < r+1$ for any choice of x and l since by uniqueness the zero polynomial would be the only degree at most $n-1$ polynomial that is zero at $(\mathbf{x}, \mathbf{k})$.

We note several corollaries of Theorem 2.2.1. First, a few definitions are provided.

Definition 2.2.3. Suppose $(\mathbf{x}, \mathbf{k})$ is consistent of degree m. A *minimal remainder polynomial at* (x, l) is a polynomial $R_{x,l} \colon \mathbb{R} \to \mathbb{R}$ of minimum degree such that

$$R_{x,l}^{(k_i)}(x_i) = 0, \qquad 1 \leq i \leq n,$$

but

$$R_{x,l}^{(l)}(x) = 1.$$

Given a standard interpolation basis, one can actually compute a minimal remainder polynomial. Towards this computation, a basis for the collection of polynomials that are zero at $(\mathbf{x}, \mathbf{k})$ is noted, then used to find a minimal remainder polynomial.

Definition 2.2.4. Suppose $(\mathbf{x}, \mathbf{k})$ is consistent of degree m and $(L_1, \ldots, L_n)$ is a standard interpolation basis. Let p be a non negative integer and let a be a real number. Then *the degree p zero polynomial through* $(\mathbf{x}, \mathbf{k})$*, centered at* a, is the polynomial $Q_{p,a} \colon \mathbb{R} \to \mathbb{R}$ defined by

$$Q_{p,a}(x) = \frac{(x-a)^p}{p!} - \sum_{\substack{i=1 \\ p \geq k_i}}^{n} \frac{(x_i - a)^{p-k_i}}{(p-k_i)!} L_i(x).$$

Proposition 2.2.5. *If* $(\mathbf{x}, \mathbf{k})$ *is consistent of degree* m*,* $(L_1, \ldots, L_n)$ *is a standard interpolation basis, and* $Q_{p,a}$ *is a degree* p *zero polynomial through* $(\mathbf{x}, \mathbf{k})$*, centered at* a*, then*

$$Q_{p,a}^{(k_j)}(x_j) = 0,$$

for every $j \in \{1, \ldots, n\}$.

Proof. Let $\sigma \colon \mathbb{R} \to \mathbb{R}$ such that $\sigma(x) = 1$ if $x \geq 0$ and $\sigma(x) = 0$ if $x < 0$. Then

$$\begin{aligned} Q_{p,a}^{(k_j)}(x_j) &= \sigma(p-k_j)\frac{(x_j - a)^{p-k_j}}{(p-k_j)!} - \sum_{\substack{i=1 \\ p \geq k_i}}^{n} \frac{(x_i - a)^{p-k_i}}{(p-k_i)!} L_i^{(k_j)}(x_j) \\ &= \sigma(p-k_j)\frac{(x_j - a)^{p-k_j}}{(p-k_j)!} - \sum_{i=1}^{n} \sigma(p-k_j)\frac{(x_i - a)^{p-k_i}}{(p-k_i)!} \delta_{ij} \\ &= 0. \end{aligned}$$

□

Corollary 2.2.6. *Suppose* $(\mathbf{x}, \mathbf{k})$ *is consistent of degree* m*. Let* $a \in \mathbb{R}$ *and* $(x, l) \in \mathbb{R} \times \mathbb{Z}_{\geq 0}$ *such that* $(x, l) \neq (x_i, k_i)$ *for every* $i \in \{1, \ldots, n\}$*. For every* $p \geq 0$*, let* $Q_{p,a}$ *be a degree* p *zero polynomial through* $(\mathbf{x}, \mathbf{k})$*, centered at* a*. Let* $r+1$ *be the degree of a minimal remainder polynomial at* (x, l)*. If* $r + 1 > m$ *then*

(i) $r + 1$ *is the smallest non-negative integer* p*, where* $p \geq l$*, for which the quantity*

$$Q_{p,a}^{(l)}(x) = \frac{(x-a)^{p-l}}{(p-l)!} - \sum_{\substack{i=1 \\ p \geq k_i}}^{n} \frac{(x_i - a)^{p-k_i}}{(p-k_i)!} L_i^{(l)}(x)$$

is non zero,

(ii) if $r + 1 = p$*, then*

$$R_{x,l} := \frac{Q_{p,a}}{Q_{p,a}^{(l)}(x)}$$

is a minimal remainder polynomial at (x, l)*, and for any* $y \in \mathbb{R}$*,*

$$\frac{1}{R_{x,l}^{(p)}(y)} = Q_{p,a}^{(l)}(x).$$

Proof. Suppose R is any polynomial of degree p such that

$$R^{(k_i)}(x_i) = 0, \qquad 1 \leq i \leq n,$$

and

$$R^{(l)}(x) = 1.$$

First, it is clear that $p \geq l$, since $R^{(l)}$ is not the zero polynomial. Moreover, if $r + 1$ is the minimum degree of such a remainder polynomial at $(\mathbf{x}, \mathbf{k})$, then $p \geq r+1$. Hence $Q_{p,a}^{(l)}(x) = 0$ for every $p < r + 1$, otherwise by Proposition 2.2.5,

$$\frac{Q_{p,a}}{Q_{p,a}^{(l)}(x)}$$

would be a remainder polynomial of smaller degree.

Now suppose $p = r + 1$ and apply Theorem 2.2.1 to R. Let

$$P = \sum_{i=1}^{n} R^{(k_i)}(x_i) L_i.$$

By assumption $r+1 > m$, and P is degree m, so $P^{(l)}(x) = 0$. Hence, applying Theorem 2.2.1 gives

$$\begin{aligned}
1 &= R^{(l)}(x) - P^{(l)}(x) \\
&= R^{(p)}(y)\left[\frac{(x-a)^{p-l}}{(p-l)!} - \sum_{i=1}^{n} \frac{(x_i-a)^{p-k_i}}{(p-k_i)} L_i^{(l)}(x)\right] \\
&= R^{(p)}(y) Q_{p,a}^{(l)}(x).
\end{aligned}$$

Therefore $Q_{p,a}^{(l)}(x) \neq 0$, and

$$\frac{1}{R^{(p)}(y)} = Q_{p,a}^{(l)}(x)$$

completing the proof.

$\square$

Now, the associated asymptotic interpolation error can be calculated by utilizing the degree p zero polynomial through $(\mathbf{x}, \mathbf{k})$.

Corollary 2.2.7. *Let $a \in \mathbb{R}$, and suppose $(\mathbf{x}, \mathbf{k})$ is consistent of degree m. Suppose $R_{x,l}$ is a minimal remainder polynomial of $(\mathbf{x}, \mathbf{k})$ at (x, l), with degree p and $p > m$. Let $Q_{p,a}$ be the degree p zero polynomial through $(\mathbf{x}, \mathbf{k})$, centered at a. Let $f\colon \mathbb{R} \to \mathbb{R}$ such that $f^{(p)}$ is continuous. Let $(L_1, \ldots, L_n)$ be a standard interpolation basis for $(\mathbf{x}, \mathbf{k})$. For every $h \in \mathbb{R}$ such that $h \neq 0$, let $P_h\colon \mathbb{R} \to \mathbb{R}$ be defined as*

$$P_h(y) = \sum_{i=1}^{n} f^{(k_i)}(a + h(x_i - a)) \cdot h^{k_i} L_i\Big(a + \frac{y-a}{h}\Big).$$

Then

$$\lim_{h \to 0}\left[\frac{f^{(l)}(a + h(x-a)) - P_h^{(l)}(a + h(x-a))}{h^{p-l}}\right] = f^{(p)}(a)\frac{Q_{p,a}^{(l)}(x)}{(x-a)^{p-l}}.$$

Thus the approximation $P_h^{(l)}(a+h)$ of $f^{(l)}(a+h)$ is at most $\mathcal{O}(h^{p-l})$.

Proof. By Proposition 2.1.5, the Birkhoff interpolation problem is scale and translation invariant. Hence if $(\mathbf{x}, \mathbf{k})$ is consistent of degree m, then $(\mathbf{x}_a(h), \mathbf{k})$ is consistent of degree m, and a minimal remainder polynomial of $(\mathbf{x}(h), \mathbf{k})$ at $(a+h, l)$, $R_{a+h,l}$, is also of degree p. Let $(L_{a,h,1}, \ldots, L_{a,h,n})$ be the standard interpolation basis for $(\mathbf{x}_a(h), \mathbf{k})$ defined by

$$L_{a,h,i}(y) = h^{k_i} L_i\left(a + \frac{y-a}{h}\right)$$

for every $i \in \{1, \dots, n\}$. Then, by Theorem 2.2.1

$$\begin{aligned}
&f^{(l)}(a + h(x-a)) - P_h^{(l)}(a + h(x-a)) \\
&= f^{(p)}(\xi(a + h(x-a)))\frac{h^{p-l}}{(p-l)!} - \sum_{i=1}^{n} \left[f^{(p)}(\xi_i(a + a_i h))\frac{(a_i h)^{p-k_i}}{(p-k_i)!} \right] L_{a,h,i}^{(l)}(a+h)
\end{aligned}$$

for some $\xi(a + h(x-a))$ between a and $a + h(x-a)$, and some $\xi_i(a + a_i h)$ between a and $a + a_i h$ for each $i \in \{1, \dots, n\}$. Hence, since $f^{(p)}$ is continuous,

$$\lim_{h \to 0} \left[\frac{f^{(l)}(a+h) - P_h^{(l)}(a+h)}{h^{p-l}} \right] = f^{(p)}(a) \lim_{h \to 0} \frac{\frac{h^{p-l}}{(p-l)!} - \sum_{i=1}^{n} \left[\frac{(a_i h)^{p-k_i}}{(p-k_i)!} \right] L_{a,h,i}^{(l)}(a + h(x-a))}{h^{p-l}}.$$

However, for every $i \in \{1, \dots, n\}$,

$$\begin{aligned}
L_{a,h,i}^{(l)}(a + h(x-a)) &= \frac{d^l}{dy^l} \left(h^{k_i} L_i\big(a + \frac{y-a}{h}\big) \right) \Bigg|_{y = a + h(x-a)} \\
&= h^{k_i - l} L_i^{(l)}\big(x\big),
\end{aligned}$$

which gives

$$\begin{aligned}
&f^{(p)}(a) \lim_{h \to 0} \frac{\frac{h^{p-l}}{(p-l)!} - \sum_{i=1}^{n} \left[\frac{(a_i h)^{p-k_i}}{(p-k_i)!} \right] L_{a,h,i}^{(l)}(a + h(x-a))}{h^{p-l}} \\
&= f^{(p)}(a) \lim_{h \to 0} \frac{\frac{h^{p-l}}{(p-l)!} - \sum_{i=1}^{n} \left[\frac{(a_i h)^{p-k_i}}{(p-k_i)!} \right] h^{k_i - l} L_i^{(l)}(x))}{h^{p-l}} \\
&= f^{(p)}(a) \lim_{h \to 0} \left[\frac{1}{(p-l)!} - \sum_{i=1}^{n} \left[\frac{(a_i)^{p-k_i}}{(p-k_i)!} \right] L_i^{(l)}\big(x)\big) \right] \\
&= f^{(p)}(a) \frac{Q_{p,a}^{(l)}(x)}{(x-a)^{p-l}},
\end{aligned}$$

which is a constant value. □

Remark 2.2.8. Corollary 2.2.7 shows that the local error of approximation is at most $\mathcal{O}(h^{p-l})$, with an asymptotic error constant of $f^{(p)}(a)\frac{Q_{p,a}^{(l)}(x)}{(x-a)^{p-l}}$. By Corollary 2.2.6, one could rewrite this asymptotic error constant as $\frac{f^{(p)}(a)}{(x-a)^{p-l} R_{x,l}^{(p)}(y)}$.

2.3 Application to First Order Initial Value Problems

In this section, a numerical method to approximate solutions of a first order initial value problem is derived from a Birkhoff quadrature method. Consider the initial value problem defined by

$$\mathbf{y}'(t) = \mathbf{f}(t, \mathbf{y}(t)) \qquad \mathbf{y}(t_0) = \mathbf{y}_0, \tag{2.1}$$

$\mathbf{y}\colon \mathbb{R} \to \mathbb{R}^d$ and $\mathbf{f}\colon \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$. A natural generalization of the well known collocation method, as defined in [7], is given here.

Definition 2.3.1. Let $(\mathbf{x}, \mathbf{k})$ be a set of nodes of length n and consistent of degree m. Let $h > 0$ and let $t_l = t_0 + lh$ for each $l \in \{1, \ldots, N\}$. The *Birkhoff collocation method* to approximate the solution of Equation (2.1) is defined by iteratively calculating approximate function values $\mathbf{y}_l$, $l \in \{0, \ldots, N\}$. First, find a degree at most m polynomial $\mathbf{u}_l$ that satisfies $\mathbf{u}_l(t_l) = \mathbf{y}_l$ and

$$\mathbf{u}_l^{(k_i+1)}(t_l + x_i h) = \mathbf{f}^{(k_i)}(t_l + x_i h, \mathbf{u}_l(t_l + x_i h)), \qquad 1 \leq i \leq n.$$

Then let $\mathbf{y}_{l+1} := \mathbf{u}_l(t_{l+1})$, for each $l \in \{1, \ldots, N-1\}$. Here $\mathbf{f}^{(k_i)}$ denotes the k_i-th total derivative of $\mathbf{f}$ with respect to t.

Remark 2.3.2. To use a Birkhoff collocation method one typically needs to utilize an implicit equation solver, such as Newton's method, to find the coefficients of each polynomial $\mathbf{u}_l$.

As is typical, one can associate such a scheme with a quadrature method. In the case of Definition 2.3.1 it is natural to consider the Birkhoff quadrature method defined by approximating the integral of a function g as

$$\int_{x_0}^{x} g(t)dt \approx \int_{x_0}^{x} P(t)dt, \tag{2.2}$$

where P interpolates g through $(\mathbf{x}, \mathbf{k})$. For the quadrature method given by Equation (2.2) there is an associated integrated interpolation method.

Proposition 2.3.3. *Let* $\mathbf{x} = (x_1, \ldots, x_n)$*,* $\mathbf{k} = (k_1, \ldots, k_n)$ *and let* P, Q *be polynomials on* $\mathbb{R}$ *with degree* $n-1$ *and* n *respectively. Suppose* $(\mathbf{x}, \mathbf{k})$ *is consistent and* P *interpolates* g' *at* $(\mathbf{x}, \mathbf{k})$*. Let* $x_0 \in \mathbb{R}$ *and define the associated integrated nodes as* $\mathbf{x}^* = (x_0, x_1, \ldots, x_n)$*,* $\mathbf{k}^* = (0, k_1+1, \ldots k_n+1)$*. Then* $(\mathbf{x}^*, \mathbf{k}^*)$ *is consistent and if and* Q *interpolates* g *at* $(\mathbf{x}^*, \mathbf{k}^*)$*, then*

$$g(x_0) + \int_{x_0}^{x} P(t)dt = Q(x).$$

Proof. By consistency, P is unique. Let $H\colon \mathbb{R} \to \mathbb{R}$ be defined by

$$H(x) = g(x_0) + \int_{x_0}^{x} P(t)dt.$$

Note that H is a degree at most n polynomial, $H(0) = g(x_0)$ and for any $i \in \{1, \dots, n\}$

$$H^{(k_i+1)}(x_i) = P^{(k_i)}(x_i) = g^{(k_i+1)}(x_i).$$

Hence, H is an interpolation polynomial of g through $(\mathbf{x}^*, \mathbf{k}^*)$ for any sufficiently differentiable g, and $(\mathbf{x}, \mathbf{k})$ is consistent, so it follows that $(\mathbf{x}^*, \mathbf{k}^*)$ is consistent. Therefore $H = Q$ since Q and H are the unique interpolation polynomials of g at $(\mathbf{x}^*, \mathbf{k}^*)$. ☐

Now, we apply Theorem 2.2.1 to obtain an error for the integrated method. This error, in conjunction with the Alekseev-Gröbner Lemma, potentially yields a method to calculate local truncation error of the method given in Definition 2.3.1.

Corollary 2.3.4. *Let $x \in \mathbb{R}$, suppose $(\mathbf{x}, \mathbf{k})$ is consistent and let $(\mathbf{x}^*, \mathbf{k}^*)$ be the associated integrated scheme. Let P be a degree $n-1$ polynomial that interpolates g' at $(\mathbf{x}, \mathbf{k})$ and Q be a degree n polynomial that interpolates g at $(\mathbf{x}^*, \mathbf{k}^*)$. Suppose $R_{x,0}$ has degree p. Then the error of the Birkhoff quadrature method defined by $(\mathbf{x}, \mathbf{k})$ at x,*

$$g(x) - g(x_0) - \int_{x_0}^{x} P(t), dt$$

is given by the error of the interpolation scheme $(\mathbf{x}^*, \mathbf{k}^*)$ of g (as in Theorem 2.2.1 with $a = x_0$ and $l = 0$),

$$g(x) - Q(x) = g^{(p)}(\xi(x))\frac{(x-x_0)^p}{(p)!} - \sum_{i=0}^{n}\left[g^{(p)}(\xi_i(x_i))\frac{(x_i - x_0)^{p-k_i-1}}{(p-k_i-1)!}\right] L_i(x)$$

where $(L_0, L_1, \dots L_n)$ is a standard interpolation basis of $(\mathbf{x}^*, \mathbf{k}^*)$.

Finally, one could potentially utilize the local quadrature error given in Corollary 2.2.7 to establish a global error bound on the Birkhoff collocation method. A global error is established for standard collocation methods, without higher order derivative conditions, in [7] and [8]. In general if a standard collocation method is derived from a quadrature method that is locally $\mathcal{O}(h^p)$, then the global collocation method for first order initial value problems is $\mathcal{O}(h^{p-1})$. The global error of the standard collocation methods is calculated by utilizing the Alekseev-Gröbner Lemma. This method suggests a potential method to calculate the global error of Birkhoff collocation methods, that is to utilize repeated applications of the Alekseev-Gröbner Lemma. However, it is not clear at this point what the global rate of convergence is for a general Birkhoff collocation method.

2.4 Optimal Node Placements for Quadrature Methods

In this section a few examples that illustrate applications of the results from previous sections are given. The positions of nodes that result in the highest order error for one and two point quadrature methods are established analytically. The process of finding optimal nodes amounts to finding the minimum degree of a polynomial that is zero at each if the interpolation conditions, but is one at the position that the error is to be calculated. This is accomplished by analyzing the collection of degree p zero polynomials, as in Corollary 2.2.6. A two point Birkhoff Collocation method that does not originate from any standard collocation method is analyzed. Finally, optimal node placements for every three point quadrature method are established via numerical approximation.

Every statement in this section is about the local convergence order of the associated quadrature method given in Proposition 2.3.3, a result that is known due to Corollary 2.2.7.

Example 2.4.1. (Euler's Method(s) – one point collocation) Let $h \in \mathbb{R}$ where $h \neq 0$. Consider the one point interpolation problem defined by $\mathbf{x} = (\alpha h)$ and $\mathbf{k} = (0)$, where $\alpha \in \mathbb{R}$. The associated quadrature method at the origin is then given by $(\mathbf{x}^*) = (0, \alpha h)$ and $(\mathbf{k}^*) = (0, 1)$ with basis $(L_1, L_2) = (1, x)$; clearly the system $(\mathbf{x}^*, \mathbf{k}^*)$ is consistent. The interpolation polynomial P of f at $(\mathbf{x}^*, \mathbf{k}^*)$ is given by

$$P(x) = f(0) + f'(\alpha h)x.$$

Suppose we wish to estimate f at h using P, then by Corollary 2.2.7 the method is locally $\mathcal{O}(h^p)$ where p is the smallest degree of a polynomial that is zero at $(\mathbf{x}^*, \mathbf{k}^*)$, but is one at $(h, 0)$. By Corollary 2.2.6, p can be calculated by finding the smallest non-negative integer p for which the expression

$$\frac{h^p}{p!} - \sum_{i=1}^{2} \frac{x_i^{p-k_i}}{(p-k_i)!} L_i(h) = \frac{h^p}{p!} - \left[\frac{0^{p-0}}{(p-0)!}1 + \frac{(\alpha h)^{p-1}}{(p-1)!}h\right] = h^p \frac{1 - \alpha^{p-1}p}{p!} \tag{2.3}$$

is nonzero; here $a = 0$ and $(x, l) = (h, 0)$. Since $(\mathbf{x}^*, \mathbf{k}^*)$ is consistent, p is at least 2. When $p = 2$, Equation (2.3) is zero if and only $\alpha = \frac{1}{2}$. When $p = 3$ Equation (2.3) is not zero if $\alpha = \frac{1}{2}$. Thus the method defined by $(\mathbf{x}^*, \mathbf{k}^*)$ is locally $\mathcal{O}(h^2)$ when $\alpha \neq \frac{1}{2}$ and $\mathcal{O}(h^3)$ when $\alpha = \frac{1}{2}$. Additionally, by Corollary 2.2.6, when $p = 2$ and $\alpha \neq \frac{1}{2}$, the minimal remainder polynomial at $(h, 0)$ is given by $Q/Q(h)$, where

$$Q(x) = \frac{x^2}{2!} - \sum_{i=1}^{2} \frac{x_i^{2-k_i}}{(2-k_i)!} L_i(x) = \frac{x^2}{2} - \alpha h x.$$

Hence

$$R(x) = \frac{\frac{x^2}{2} - 2\alpha hx}{\frac{h^2}{2} - \alpha h^2} = \frac{x^2 - \alpha hx}{h^2(1-2\alpha)}$$

is the minimal remainder polynomial; observe that $R(0) = 0$, $R'(\alpha h) = 0$ and $R(h) = 1$.

When used for solving a first order ODE, as in Definition 2.3.1, if $\alpha = 0$ this method is exactly (Forward) Euler's method and if $\alpha = 1$, this method is Backwards Euler's method. Both methods are know to be locally $\mathcal{O}(h^2)$ and globally $\mathcal{O}(h)$. If $\alpha = \frac{1}{2}$ the method gives midpoint Euler method, which is known to be locally $\mathcal{O}(h^3)$ and globally $\mathcal{O}(h^2)$ (see [7]).

The next example is concerned with the approximation error of two point collocation methods.

Example 2.4.2. (Two point collocation) Let $h, \alpha, \beta \in \mathbb{R}$ such that $h \neq 0$ and $\alpha \neq \beta$. Consider the interpolation problem defined by $\mathbf{x} = (\alpha h, \beta h)$ and $\mathbf{k} = (0,0)$ with associated integrated nodes $\mathbf{x}^* = (0, \alpha h, \beta h)$ and $\mathbf{k}^* = (0,1,1)$. The nodes $(\mathbf{x}^*, \mathbf{k}^*)$ have the standard interpolation basis given by

$$(L_1(x), L_2(x), L_3(x)) = \left(1, \frac{x^2 - 2\beta hx}{2h(\alpha-\beta)}, \frac{x^2 - 2\alpha hx}{2h(\beta-\alpha)}\right)$$

which may be used to find optimal placements for α and β by Corollary 2.2.6. Again, we wish to find values of α and β for which

$$\frac{h^p}{p!} - \sum_{i=1}^{3} \frac{x_i^{p-k_i}}{(p-k_i)!} L_i(h) = \frac{h^p}{p!} - \left[\frac{(\alpha h)^{p-1}}{(p-1)!} \cdot \frac{h(1-2\beta)}{2(\alpha-\beta)} + \frac{(\beta h)^{p-1}}{(p-1)!} \cdot \frac{h(1-2\alpha)}{2(\beta-\alpha)}\right], \tag{2.4}$$

which can be re-written as

$$\frac{h^p}{2p!}\left[2 - p[(\alpha^{p-2} + \alpha^{p-3}\beta + \cdots + \alpha\beta^{p-3} + \beta^{p-2}) - 2\alpha\beta(\alpha^{p-3} + \alpha^{p-4}\beta + \cdots + \alpha\beta^{p-4} + \beta^{p-3})]\right], \tag{2.5}$$

is nonzero. Since $(\mathbf{x}^*, \mathbf{k}^*)$ is consistent, p is at least 3. When $p = 3$ setting Equation (2.5) to zero yields

$$0 = 2 - 3(\alpha + \beta - 2\alpha\beta)$$

with solution for any $\beta \neq \frac{1}{2}$

$$\alpha = \frac{3\beta - 2}{6\beta - 3}.$$

α	β	Local Error	Global Error
$\alpha \neq \frac{3\beta-2}{6\beta-3}$	β	$\mathcal{O}(h^3)$	$\mathcal{O}(h^2)$
$\frac{3\beta-2}{6\beta-3}$	$\beta \neq \frac{1}{2} \pm \frac{1}{2\sqrt{3}}$	$\mathcal{O}(h^4)$	$\mathcal{O}(h^3)$
$\frac{1}{2} \mp \frac{1}{2\sqrt{3}}$	$\frac{1}{2} \pm \frac{1}{2\sqrt{3}}$	$\mathcal{O}(h^5)$	$\mathcal{O}(h^4)$

Table 2.1: Theoretical order of local error and of global error for the Birkhoff collocation method with $\mathbf{k} = (0,0)$ and $\mathbf{x} = (\alpha, \beta)$.

Hence the method is $\mathcal{O}(h^3)$ locally when $\alpha \neq \frac{3\beta-2}{6\beta-3}$. If $\alpha = \frac{3\beta-2}{6\beta-3}$, then p is at least 4. When $p = 4$ we set Equation (2.5) to zero to obtain

$$0 = 2 - 4(\alpha^2 + \alpha\beta + \beta^2 - 2\alpha\beta(\alpha + \beta)),$$

substituting $\alpha = \frac{3\beta-2}{6\beta-3}$ gives

$$0 = \frac{2(6\beta^2 - 6\beta + 1)}{9(1 - 2\beta))}$$

with the familiar solutions given by the shifted and scaled zeros of the degree two Legendre polynomial

$$\beta = \frac{1}{2} \pm \frac{1}{2\sqrt{3}}.$$

Hence the local error for different placements of α and β of the collocation method defined by $(\mathbf{x}^*, \mathbf{k}^*)$ can be summarized by Table 2.1. Of course one could derive local asymptotic error constants by using Corollary 2.2.7. The global error column of Table 2.1 is given by [7].

Finally, the only other well-defined Birkhoff Collocation method defined by a consistent two point Birkhoff Quadrature method is considered.

Example 2.4.3. (A two point Birkhoff-Collocation method) Let $h, \alpha, \beta \in \mathbb{R}$ such that $h \neq 0$ and $\alpha \neq \beta$. Consider the interpolation problem defined by $\mathbf{x} = (\alpha h, \beta h)$ and $\mathbf{k} = (0, 1)$ with associated integrated scheme $\mathbf{x}^* = (0, \alpha h, \beta h)$ and $\mathbf{k}^* = (0, 1, 2)$. The nodes $(\mathbf{x}^*, \mathbf{k}^*)$ have the standard interpolation basis given by

$$(L_1(x), L_2(x), L_3(x)) = \left(1, x, \frac{x(x - 2\alpha h)}{2}\right)$$

which may be used to find optimal placements for α and β by Corollary 2.2.6. Again we wish to solve for where the quantity

$$\begin{aligned}\frac{h^p}{p!} - \sum_{i=1}^{3} \frac{x_i^{p-k_i}}{(p-k_i)!} L_i(h) &= \frac{h^p}{p!} - \left[\frac{(\alpha h)^{p-1}}{(p-1)!} h + \frac{(\beta h)^{p-2}}{(p-2)!} \cdot \frac{h^2(1-2\alpha)}{2} \right] \\ &= \frac{h^p}{2p!} \left[2 - 2p\alpha^{p-1} - p(p-1)\beta^{p-2}(1-2\alpha) \right]\end{aligned} \tag{2.6}$$

is nonzero. Since $(\mathbf{x}^*, \mathbf{k}^*)$ is consistent, p is at least 3. When $p = 3$ Equation (2.6) becomes

$$0 = 2 - 6\alpha^2 - 6\beta(1-2\alpha),$$

with solution for any β given by

$$\alpha = \beta \pm \frac{\sqrt{3(3\beta^2 - 3\beta + 1)}}{3}.$$

Hence the method is $\mathcal{O}(h^3)$ locally when $\alpha \neq \beta \pm \frac{\sqrt{3(3\beta^2-3\beta+1)}}{3}$. If $\alpha = \beta \pm \frac{\sqrt{3(3\beta^2-3\beta+1)}}{3}$, then p is at least 4. When $p = 4$, setting Equation (2.6) to zero gives

$$0 = 2 - 8\alpha^3 - 12\beta^2(1-2\alpha),$$

and substituting $\alpha = \beta \pm \frac{\sqrt{3(3\beta^2-3\beta+1)}}{3}$ gives the equation

$$0 = -8\beta^3 + 12\beta^2 - 8\beta + 2 \pm \frac{8\sqrt{3}(3\beta^2 - 3\beta + 1)^{\frac{3}{2}}}{9}$$

with solutions

$$\beta = \frac{1}{2} \pm \frac{\sqrt{2\sqrt{3} - 3}}{6}.$$

Substituting back into the expression for α, yields

$$\alpha = \frac{1}{2} \pm \frac{\sqrt{2}\left(\sqrt[4]{27} - 3\sqrt[4]{3}\right)}{12}.$$

Hence the order of error for different placements of α and β of the Birkhoff collocation method defined by $(\mathbf{x}^*, \mathbf{k}^*)$ can be summarized by Table 2.2.

While it is quite nice to find optimal node placements using the above method for one and two point methods, it is much more challenging to implement the same technique and find exact solutions for three point methods. However, utilizing a numerical root finding

α	β	Local Error
$\alpha \neq \beta \pm \frac{\sqrt{3(3\beta^2-3\beta+1)}}{3}$	β	$\mathcal{O}(h^3)$
$\beta \pm \frac{\sqrt{3(3\beta^2-3\beta+1)}}{3}$	$\beta \neq \frac{1}{2} \pm \frac{\sqrt{2\sqrt{3}-3}}{6}$	$\mathcal{O}(h^4)$
$\frac{1}{2} \pm \frac{\sqrt{2}\left(\sqrt[4]{27}-3\sqrt[4]{3}\right)}{12}$	$\frac{1}{2} \pm \frac{\sqrt{2\sqrt{3}-3}}{6}$	$\mathcal{O}(h^5)$

Table 2.2: Local error of the quadrature method derived from $\mathbf{k} = (0, 1)$ and $\mathbf{x} = (\alpha, \beta)$.

α	β	γ	Local Error	Global Error
0.1127016654	0.5000000000	0.8872983346	$\mathcal{O}(h^7)$	$\mathcal{O}(h^6)$

Table 2.3: Local error and of global error of the Birkhoff collocation method with $\mathbf{k} = (0, 0, 0)$ and $\mathbf{x} = (\alpha, \beta, \gamma)$.

algorithm one can solve for the optimal node placements with relative ease. The existence of nodes that yield double precision for quadrature methods involving function values and derivatives is given in [9].

The values α, β, and γ in Tables 2.3 to 2.7 are generated numerically using a similar process as in Examples 2.4.1 to 2.4.3. These values approximate node placements for which the degree of a minimal remainder polynomial is maximized. Any symmetric values of the constants α, β and γ, obtained by a permutation of values of the same derivative order, are not listed throughout Tables 2.3 to 2.7.

α	β	γ	Local Error
0.8283672870	0.1240841012	0.6560585536	$\mathcal{O}(h^7)$
0.1219979157	0.5760571948	0.7954805730	$\mathcal{O}(h^7)$
0.8780020843	0.4239428052	0.2045194270	$\mathcal{O}(h^7)$
0.1716327130	0.8759158988	0.3439414463	$\mathcal{O}(h^7)$

Table 2.4: Local error of the quadrature method derived from $\mathbf{k} = (0, 0, 1)$ and $\mathbf{x} = (\alpha, \beta, \gamma)$.

α	β	γ	Local Error
0.6723831273	0.1272548190	0.7249365994	$\mathcal{O}(h^7)$
0.3276168726	0.8727451810	0.2750634005	$\mathcal{O}(h^7)$
0.8343701525	0.1656298475	0.5000000000	$\mathcal{O}(h^7)$

Table 2.5: Local error of the quadrature method derived from $\mathbf{k} = (0, 0, 2)$ and $\mathbf{x} = (\alpha, \beta, \gamma)$.

α	β	γ	Local Error
0.5000000000	0.2261387212	0.7738612788	$\mathcal{O}(h^7)$
0.7944855600	0.6125225784	0.2353185980	$\mathcal{O}(h^7)$
0.2055144400	0.7646814020	0.3874774216	$\mathcal{O}(h^7)$

Table 2.6: Local error of the quadrature method derived from $\mathbf{k} = (0, 1, 1)$ and $\mathbf{x} = (\alpha, \beta, \gamma)$.

α	β	γ	Local Error
0.2175750650	0.4379532711	0.6662970045	$\mathcal{O}(h^7)$
0.2235631093	0.6853255785	0.5256720915	$\mathcal{O}(h^7)$
0.4125224680	0.7593177542	0.3132970091	$\mathcal{O}(h^7)$
0.5874775320	0.2406822458	0.6867029909	$\mathcal{O}(h^7)$
0.7764368906	0.3146744214	0.4743279084	$\mathcal{O}(h^7)$
0.7824249350	0.5620467288	0.3337029955	$\mathcal{O}(h^7)$

Table 2.7: Local error of the quadrature method derived from $\mathbf{k} = (0, 1, 2)$ and $\mathbf{x} = (\alpha, \beta, \gamma)$.

2.5 Finite Differences Approximation Convergence Order

An immediate application of Corollary 2.2.7 is the calculation of the convergence order of finite difference approximations. In fact, one can proceed in a similar manner as Section 2.4 to find the optimal interpolation positions for calculating finite difference approximations. As a first example, consider the family of two-point first difference methods.

Example 2.5.1. Given the function $f\colon \mathbb{R} \to \mathbb{R}$ one may obtain an approximation of $f'(0)$ by calculating the Lagrange polynomial through the two points αh and βh, for the real numbers $\alpha \neq \beta$ and $h \neq 0$. This Lagrange polynomial P, is defined by

$$P(x) = f(\alpha h)\frac{x - \beta h}{\alpha h - \beta h} + f(\beta h)\frac{x - \alpha h}{\beta h - \alpha h},$$

yielding the approximation

$$f'(0) \approx P'(0) = \frac{f(\alpha h) - f(\beta h)}{\alpha h - \beta h}. \tag{2.7}$$

To find the order of approximation, first find the minimum degree of a polynomial R such that $R(\alpha h) = 0$, $R(\beta h) = 0$ and $R'(0) = 1$, then apply Corollary 2.2.7. By Corollary 2.2.6, this polynomial my be found in terms of minimal zero polynomials. In this case,

$$\left(\frac{x - \beta h}{\alpha h - \beta h}, \frac{x - \alpha h}{\beta h - \alpha h}\right)$$

forms an interpolation basis. Hence the degree p zero polynomial centered at 0 is the polynomial $Q_{p,0}\colon \mathbb{R} \to \mathbb{R}$ defined by

$$Q_{p,0}(x) = \frac{x^p}{p!} - \frac{(\alpha h)^p}{p!}\frac{x - \beta h}{\alpha h - \beta h} - \frac{(\beta h)^p}{p!}\frac{x - \alpha h}{\beta h - \alpha h}.$$

Thus, the degree of a minimal remainder polynomial is the smallest non-negative integer $p \geq 1$ for which $Q'_{p,0}(0)$ is non zero. However, $Q'_{1,0}$ is the zero polynomial, so p must be larger than 1. When $p > 1$,

$$Q'_{p,0}(0) = \frac{1}{p!}\frac{(\beta h)^p - (\alpha h)^p}{\alpha h - \beta h},$$

so

$$Q'_{2,0}(0) = \frac{1}{2}\frac{\beta^2 - \alpha^2}{\alpha - \beta}$$

is nonzero whenever $\beta \neq -\alpha$, and is zero when $\beta = -\alpha$. Moreover,

$$Q'_{3,0}(0) = \frac{1}{6}\frac{\beta^3 - \alpha^3}{\alpha - \beta}$$

is nonzero when $\beta = -\alpha$. Therefore, by Corollary 2.2.7, Equation (2.7) converges at most $\mathcal{O}(h)$ when $\beta \neq -\alpha$ and converges at most $\mathcal{O}(h^2)$ when $\beta = -\alpha$.

While the process outlined in Example 2.5.1 feels somewhat formal and may be slightly cumbersome, the essential character of this example is summarized by the statement

"Any quadratic R for which $R(-\alpha) = 0$ and $R(\alpha) = 0$ also satisfies $R'(0) = 0$."

Thus a minimal remainder polynomial R for which $R(-\alpha) = 0 = R(\alpha)$ and $R'(0) = 1$ must be at least a cubic polynomial. Then, one may immediately write down such a remainder polynomial without performing any calculations,

$$R(x) = \frac{-x(x-\alpha)(x+\alpha)}{\alpha^2} + c(x-\alpha)(x+\alpha),$$

for any constant $c \in \mathbb{R}$. Of course, in the case of a two point first difference method which is not central, one has no problems conjuring up a quadratic remainder polynomial.

The benefits of the formal process outlined in the previous example is that it is entirely algorithmic. One can quite easily utilize a symbolic computational software to perform these calculations and find the degree of minimal remainder polynomials. To illustrate this particular process once more, consider the family of three point first and second difference methods given in the next example.

Example 2.5.2. Suppose α, β, γ are all distinct fixed constants, and h is a nonzero real number. Then, the Lagrange interpolation polynomial through these three points, scaled by h, is characterized by nodes $\mathbf{x} = (\alpha h, \beta h, \gamma h)$ and $\mathbf{k} = (0,0,0)$. The nodes $(\mathbf{x}, \mathbf{k})$ are consistent, and yield the standard interpolation basis

$$\left(\frac{x-\beta h}{\alpha h-\beta h}\cdot\frac{x-\gamma h}{\alpha h-\gamma h}, \frac{x-\alpha h}{\beta h-\alpha h}\cdot\frac{x-\gamma h}{\beta h-\gamma h}, \frac{x-\alpha h}{\gamma h-\alpha h}\cdot\frac{x-\beta h}{\gamma h-\beta h}\right)$$

and hence the interpolation polynomial

$$P(x) = f(\alpha h)\frac{x-\beta h}{\alpha h-\beta h}\cdot\frac{x-\gamma h}{\alpha h-\gamma h} + f(\beta h)\frac{x-\alpha h}{\beta h-\alpha h}\cdot\frac{x-\gamma h}{\beta h-\gamma h} + f(\gamma h)\frac{x-\alpha h}{\gamma h-\alpha h}\cdot\frac{x-\beta h}{\gamma h-\beta h}$$

of f at $(\mathbf{x}, \mathbf{k})$. This polynomial yields the simultaneous approximations

$$f'(0) \approx P'(0) = \frac{-f(\alpha h)(\beta+\gamma)}{h(\alpha-\beta)(\alpha-\gamma)} + \frac{-f(\beta h)(\alpha+\gamma)}{h(\beta-\alpha)(\beta-\gamma)} + \frac{-f(\gamma h)(\alpha+\beta)}{h(\gamma-\alpha)(\gamma-\beta)} \tag{2.8}$$

and

$$f''(0) \approx P''(0) = \frac{2f(\alpha h)}{h^2(\alpha-\beta)(\alpha-\gamma)} + \frac{2f(\beta h)}{h^2(\beta-\alpha)(\beta-\gamma)} + \frac{2f(\gamma h)}{h^2(\gamma-\alpha)(\gamma-\beta)}. \tag{2.9}$$

Note that if $(\alpha, \beta, \gamma) = (-1, 0, 1)$ then the familiar second difference method

$$f''(0) \approx \frac{f(-h) - 2f(0) + f(h)}{h^2}$$

is recovered from Equation (2.9).

To calculate the order of convergence of the approximations given by Equation (2.8) and Equation (2.9), the same degree p zero polynomial through $(\mathbf{x}, \mathbf{k})$ centered at 0, $Q_{p,0}(x)$, may be studied. This polynomial is defined, as in Definition 2.2.4, by

$$Q_{p,0}(x) = \frac{x^p}{p!} - \frac{(\alpha h)^p}{p!} \frac{x - \beta h}{\alpha h - \beta h} \cdot \frac{x - \gamma h}{\alpha h - \gamma h} - \frac{(\beta h)^p}{p!} \frac{x - \alpha h}{\beta h - \alpha h} \cdot \frac{x - \gamma h}{\beta h - \gamma h} - \frac{(\gamma h)^p}{p!} \frac{x - \alpha h}{\gamma h - \alpha h} \cdot \frac{x - \beta h}{\gamma h - \beta h}.$$

Observe that when $p > 1$,

$$\begin{aligned} Q'_{p,0}(0) &= \frac{(\alpha h)^p(\beta + \gamma)}{p!h(\alpha - \beta)(\alpha - \gamma)} + \frac{(\beta h)^p(\alpha + \gamma)}{p!h(\beta - \alpha)(\beta - \gamma)} + \frac{(\gamma h)^p(\alpha + \beta)}{p!h(\gamma - \alpha)(\gamma - \beta)} \\ &= \frac{h^{p-1}}{p!} \cdot \frac{\alpha^p(\beta^2 - \gamma^2) + \beta^p(\gamma^2 - \alpha^2) + \gamma^p(\alpha^2 - \beta^2)}{(\alpha - \beta)(\alpha - \gamma)(\beta - \gamma)} \end{aligned}$$

and when $p > 2$,

$$\begin{aligned} Q''_{p,0}(0) &= \frac{-2(\alpha h)^p}{p!h^2(\alpha - \beta)(\alpha - \gamma)} + \frac{-2(\beta h)^p}{p!h^2(\beta - \alpha)(\beta - \gamma)} + \frac{-2(\gamma h)^p}{p!h^2(\gamma - \alpha)(\gamma - \beta)} \\ &= \frac{-2h^{p-2}}{p!} \cdot \frac{\alpha^p(\beta - \gamma) + \beta^p(\gamma - \alpha) + \gamma^p(\alpha - \beta)}{(\alpha - \beta)(\alpha - \gamma)(\beta - \gamma)}. \end{aligned}$$

Since $(\mathbf{x}, \mathbf{k})$ is consistent, in either case, a minimal remainder polynomial will have degree at least 3. When $p = 3$, $Q'_{3,0}(0) = 0$ has the solution

$$\gamma = -\frac{\alpha\beta}{\alpha + \beta}$$

when $\alpha + \beta \neq 0$ and $Q''_{3,0}(0) = 0$ has the solution

$$-\gamma = \alpha + \beta,$$

for any distinct α and β. If $p = 4$ and $\gamma = -\alpha\beta/(\alpha = \beta)$, then $Q'_{4,0}(0) \neq 0$. Similarly, if $-\gamma = \alpha + \beta$, then $Q''_{4,0}(0) \neq 0$. Hence the order of convergence for three point first and second difference methods can be summarized by Table 2.8 and Table 2.9 respectively.

α	β	γ	Convergence Rate of Equation (2.8)
α	β	$\gamma \neq \frac{-\alpha\beta}{\alpha+\beta}$	$\mathcal{O}(h^2)$
α	β	$\gamma = \frac{-\alpha\beta}{\alpha+\beta}, \gamma \neq \alpha + \beta$	$\mathcal{O}(h^3)$

Table 2.8: Convergence order of three point first difference approximation at zero (Equation (2.8)) with $\mathbf{k} = (0, 0, 0)$ and $\mathbf{x} = (\alpha, \beta, \gamma)$. This convergence rate is given by Corollary 2.2.7.

α	β	γ	Convergence Rate of Equation (2.9)
α	β	$-\gamma \neq \alpha + \beta$	$\mathcal{O}(h^1)$
α	β	$-\gamma = \alpha + \beta$	$\mathcal{O}(h^2)$

Table 2.9: Convergence order of three point second difference approximation at zero (Equation (2.9)) with $\mathbf{k} = (0, 0, 0)$ and $\mathbf{x} = (\alpha, \beta, \gamma)$. This convergence rate is given by Corollary 2.2.7.

Chapter 3

Birkhoff Polynomial Interpolation on $\mathbb{R}^s$

The purpose of this chapter is to extend all of the results from Chapter 2 to the interpolation of a function $f\colon \mathbb{R}^s \to \mathbb{R}$, for some fixed $s \geq 1$. The algebraic nature of the previous results allow for an extension to this larger context. The approach of this chapter is nearly identical to that of Chapter 2.

3.1 Preliminary Definitions

In multiple variables, Birkhoff interpolation is concerned with matching a function with a polynomial at a finite collection of finite order derivative values. The notation associated with this process becomes much more cumbersome, and less clear. In general, a finite order derivative value of a function $f\colon \mathbb{R}^s \to \mathbb{R}$ is the evaluation of a sequence of directional derivatives of f. In this section, the notation is established to be similar to the notation presented in Chapter 2.

Definition 3.1.1. *A k-th order derivative* will be denoted by D^k and is determined by fixing a k-tuple of unit vectors $(v_1, \ldots, v_k)$. The k-th rder derivative D^k then acts on functions $f\colon \mathbb{R}^s \to \mathbb{R}$ and outputs the composite directional derivative of f with respect to each direction vector v_i, $i \in \{1, \ldots, k\}$. That is,

$$D^k f = \partial_{v_k} \cdots \partial_{v_1} f,$$

where ∂_v denotes the directional derivative $v \cdot \nabla$. If $k = 0$ then $D^0 f := f$.

Additionally, the assumption throughout this chapter is that all functions of interest will be sufficiently continuously differentiable so that directional derivatives commute. Hence, two

k-th order derivatives $D_1^k = \partial_{v_k} \cdots \partial_{v_1}$ and $D_2^k = \partial_{2_k} \cdots \partial_{2_1}$ are equivalent, $D_1^k = D_2^k$, if and only if there exists a permutation σ on $\{1, \ldots, k\}$ such that $(v_1, \ldots, v_k) = (w_{\sigma(1)}, \ldots, w_{\sigma(k)})$.

Utilizing this notation, a definition of Birkhoff interpolation in s variables is given. Let $\mathbf{D} = (D_1^{k_1}, \ldots, D_n^{k_n})$ be an ordered collection of n finite order derivatives. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be a finite collection of vectors in $\mathbb{R}^s$, $x_i \in \mathbb{R}^s$ for every $i \in \{1, \ldots, n\}$. Let $\mathbf{c} = (c_1, \ldots, c_n)$ denote an ordered set of n real constants, $c_i \in \mathbb{R}$ for each $i\{1, \ldots, n\}$.

Definition 3.1.2. A *Birkhoff polynomial through* $(\mathbf{x}, \mathbf{D}, \mathbf{c})$ is a polynomial in s variables, $P \colon \mathbb{R}^s \to \mathbb{R}$ that satisfies

$$D_i^{k_i} P(x_i) = c_i, \qquad 1 \leq i \leq n.$$

If $c_i = D_i^{k_i} f(x_i)$ for each $i \in \{1, \ldots, n\}$, then P *interpolates* f *at* $(\mathbf{x}, \mathbf{D})$.

Definition 3.1.3. The collection $(\mathbf{x}, \mathbf{D})$ is *consistent of degree* m if for every tuple of n constants $\mathbf{c}$ there exists a Birkhoff polynomial of degree at most m through $(\mathbf{x}, \mathbf{D}, \mathbf{c})$, and m is the least such value. The collection $(\mathbf{x}, \mathbf{D})$ is *consistent* if $n = \sum_{i=0}^{m} \binom{s+i-1}{i}$.

In other words, $(\mathbf{x}, \mathbf{D})$ is consistent if it is consistent of degree m and the number of given interpolation conditions is the number of coefficients in an arbitrary degree at most m polynomial in s variables.

Definition 3.1.4. Suppose $(\mathbf{x}, \mathbf{D})$ is consistent of degree m. A *standard interpolation basis* is a tuple of polynomials $(L_1, \ldots, L_n)$, $L_i \colon \mathbb{R}^s \to \mathbb{R}$ for each $i \in \{1, \ldots, n\}$, that satisfy

$$D_j^{k_j} L_i(x_j) = \delta_{ij}, \qquad 1 \leq i, j \leq n,$$

such that L_i is of degree at most m for each $i \in \{1, \ldots, n\}$.

In an effort to keep with the notation of Chapter 2, given an element x of $\mathbb{R}^s$, there will not be any reference to the components of x throughout this section, and the next section. This is so that a subscript i on x may be reserved to denote the i-th vector in a list of vectors, rather than the i-th component of a vector. To accomplish this, every theorem in this section and the next section will be stated in terms of multi-indicex notation. This multi-index notation is established in [6] or [15]. In this manner, the components of a vector may remain implicit. Utilizing this multi index notation, a definition of a Taylor Polynomial is provided.

Definition 3.1.5. Let $f\colon \mathbb{R}^s \to \mathbb{R}$ be m times continuously differentiable at $a \in \mathbb{R}^s$. Then the *degree m Taylor Polynomial of f centered at a* is the polynomial function defined for each $x \in \mathbb{R}^s$ by

$$T_{f,m,a}(x) = \sum_{|\alpha| \le m} D^\alpha f(a) \frac{(x-a)^\alpha}{\alpha!}.$$

Proposition 3.1.6. *Let D^k be a k-th order derivative and let $f\colon \mathbb{R}^s \to \mathbb{R}$ be m times continuously differentiable at a. If $k \le m$, then*

$$D^k f(a) = D^k T_{f,m,a}(a).$$

Proof. Let β be a multi-index such that $0 \le |\beta| \le k$. It is sufficient to verify that

$$D^\beta f(a) = D^\beta T_{f,m,a}(a),$$

then by linearity of the derivative, the result extends to an arbitrary k-th order derivative. Evidently,

$$\begin{aligned} D^\beta T_{f,m,a}(a) &= \sum_{|\alpha| \le m} D^\alpha f(a) D^\beta \frac{(x-a)^\alpha}{\alpha!}\bigg|_a \\ &= D^\beta f(a) \end{aligned}$$

since $D^\beta \frac{(x-a)^\alpha}{\alpha!}\big|_a$ is 1 if $\alpha = \beta$ and 0 otherwise.

□

Now a version of Taylor's theorem in s real variables is presented, which may also be found in [18].

Theorem 3.1.7. *(Taylor's Theorem) Let $x, a \in \mathbb{R}^s$ and suppose that $f\colon \mathbb{R}^s \to \mathbb{R}$ is $m+1$ times differentiable on an open set containing the line segment from a to x. Let $v = (x-a)/|x-a|$. Then there is some $\xi(x) \in \mathbb{R}^s$ on the open line segment from x to a such that*

$$f(x) = T_{f,m,a}(x) + \partial_v^{m+1} f(\xi(x)) \frac{|x-a|^{m+1}}{(m+1)!}.$$

Throughout this chapter, $R_{f,m,a}(x)$ will denote the remainder term

$$R_{f,m,a}(x) = \partial_v^{m+1} f(\xi(x)) \frac{|x-a|^{m+1}}{(m+1)!}.$$

The idea of the proof is provided.

Proof. Let $g, T^*_{f,m,a} \colon \mathbb{R} \to \mathbb{R}$ be defined for every $t \in \mathbb{R}$ by $g(t) = f(a+tv)$ and $T^*_{f,m,a}(t) = T_{f,m,a}(a+tv)$. Then, by Taylor's Theorem in one real variable, there is some c between 0 and $|x-a|$ such that

$$g(|x-a|) = T_{g,m,a}(|x-a|) + g^{(m+1)}(c)\frac{|x-a|^{m+1}}{(m+1)!}.$$

But $g(|x-a|) = f(x)$ and $g^{(m+1)}(c) = \partial_v^{m+1} f(\xi(x))$ where $\xi(x) = a + cv$. Moreover, by Proposition 3.1.6, $T_{g,m,a}(|x-a|) = T^*_{f,m,a}(|x-a|) = T_{f,m,a}(x)$. ☐

Proposition 3.1.8. *Suppose D^k is a k-th order derivative and $f \colon \mathbb{R}^s \to \mathbb{R}$ is m times continuously differentiable on $\mathbb{R}^s$. If $k \leq m$, then*

$$D^k T_{f,m,a}(x) = T_{D^k f, m-k, a}(x).$$

Proof. Let β be a multi index such that $0 \leq |\beta| = l \leq k$. The statement will only be shown for D^β and by linearity of derivation and linearity of Taylor interpolation, the result will extend to an arbitrary k-th order derivative. Observe that

$$\begin{aligned}
D^\beta T_{f,m,a}(x) &= \sum_{|\alpha| \leq m} D^\alpha f(a) D^\beta \frac{(x-a)^\alpha}{\alpha!} \\
&= \sum_{\substack{|\alpha| \leq m \\ \alpha \geq \beta}} D^\alpha f(a) \frac{(x-a)^{\alpha-\beta}}{(\alpha-\beta)!} \\
&= \sum_{\alpha \leq m-l} D^{\alpha+\beta} f(a) \frac{(x-a)^\alpha}{\alpha!} \\
&= \sum_{\alpha \leq m-l} D^\alpha (D^\beta f)(a) \frac{(x-a)^\alpha}{\alpha!} \\
&= T_{D^\beta f, m-l, a}.
\end{aligned}$$

☐

Corollary 3.1.9. *If f satisfies the hypothesis of Theorem 3.1.7 and D^k is a k-th order derivative, then*

$$D^k f = D^k T_{f,m,a} + R_{D^k f, m-l, a}.$$

Proof. Apply Theorem 3.1.7 and Proposition 3.1.8. ☐

3.2 Interpolation Error

It is striking how readily Theorem 2.2.1 generalizes to larger contexts. This is largely due to its algebraic nature, in that the main idea of the proof is that there is some degree of polynomials that are exactly interpolated at a particular derivative value, by some finite set of other derivative values. Then, a Taylor polynomial of this degree is also exactly interpolated by the collection of interpolation conditions. This exact interpolation of a Taylor polynomial then gives a method to calculate the error.

Theorem 3.2.1. *Let* $\mathbf{D}$ *be an ordered set of* n *finite order derivatives and* $\mathbf{x}$ *be an ordered set of* n *vectors in* $\mathbb{R}^s$*, such that* $(\mathbf{x}, \mathbf{D})$ *is consistent of degree* m*. Let* $(L_1, \ldots, L_n)$ *be a standard interpolation basis of* $(\mathbf{x}, \mathbf{D})$*. Let* D^l *be an* l*-th order derivative and let* $x \in \mathbb{R}^s$ *such that* $(x, D^l) \neq (x_i, D_i^{k_i})$ *for every* $i \in \{1, \ldots, n\}$*. Let* $a \in \mathbb{R}^s$ *be fixed. Suppose* r *is the least non-negative integer such that there is a degree* $r+1$ *polynomial* R *that satisfies*

$$D_i^{k_i} R(x_i) = 0, \qquad 1 \leq i \leq n,$$

but $D^l R(x) = 1$*. Let* $f\colon \mathbb{R}^s \to \mathbb{R}$ *be* $r+1$ *times differentiable on an open ball centered at* a *of radius large enough to contain* x_i *for each* $i \in \{1, \ldots, n\}$*. Let*

$$P = \sum_{i=1}^{n} D_i^{k_i} f \cdot L_i,$$

let $v = (x-a)/|x-a|$ *and* $v_i = |x_i - a|/|x_i - a|$ *for each* $i \in \{1, \ldots, n\}$*. If* $m < r+1$*, then*

$$D^l f(x) - D^l P(x) = \partial_v^{r+1} f(\xi(x)) \frac{|x-a|^{r+1-l}}{(r+1-l)!} - \sum_{i=1}^{n} \left[\partial_{v_i}^{r+1} f(\xi_i(x_i)) \frac{|x_i - a|^{r+1-k_i}}{(r+1-k_i)!} \right] D^l L_i(x)$$

for some $\xi(x)$ *on the line between* a *and* x*, and some* $\xi_i(x_i)$ *on the line between* a *and* x_i*, for each* $i \in \{1, \ldots, n\}$*.*

Proof. The proof is nearly identical to that of Theorem 2.2.1. Let $P_T\colon \mathbb{R}^s \to \mathbb{R}$ be the interpolation polynomial of $T_{f,m,a}$ given by

$$P_T = \sum_{i=1}^{n} D_i^{k_i} T_{f,m,a} \cdot L_i.$$

By assumption, $m < r+1$ so

$$D^l P_T(x) = D^l T_{f,m,a}(x).$$

Thus,

$$D^l P(x) = \sum_{i=1}^{n} D_i^{k_i} f(x_i) \cdot D^l L_i(x).$$

However, by Theorem 3.1.7, $D_i^{k_i} f(x_i) = T_{D_i^{k_i} f,m-k_i,a}(x_i) + R_{D_i^{k_i} f,m-k_i,a}(x_i)$. Additionally, Proposition 3.1.8 gives $T_{D_i^{k_i} f,m-k_i,a}(x_i) = D_i^{k_i} T_{f,m-k_i,a}(x_i)$. Thus

$$\begin{aligned} D^l P(x) &= \sum_{i=1}^{n} D_i^{k_i} T_{f,m,a}(x_i) D^l L_i(x) + \sum_{i=1}^{n} R_{D_i^{k_i} f,m-k_i,a}(x_i) D^l L_i(x) \\ &= D^l P_T(x) + \sum_{i=1}^{n} R_{D_i^{k_i} f,m-k_i,a}(x_i) D^l L_i(x) \\ &= D^l f(x) + R_{D^l f,m,a}(x) + \sum_{i=1}^{n} R_{D_i^{k_i} f,m-k_i,a}(x_i) D^l L_i(x). \end{aligned}$$

Substituting remainders from Theorem 3.1.7 and rearranging terms gives the sought out result. ☐

While the previous result is the natural immediate extension of Theorem 2.2.1 established earlier, often the error of interest is that of a differential operator which may not be expressed as a sequence of directional derivatives; such as the Laplacian in s variables when $s > 1$. Fortunately, such differential operators are expressible as a linear combination of finite order derivative values, and the result may be extended easily.

Corollary 3.2.2. *Let $x \in \mathbb{R}^s$ be fixed. Let $D_j^{l_j}$ be an l_j-th order derivative such that $(x, D_j^{l_j}) \neq (x_i, D_i^{k_i})$, for each $j \in \{1, \ldots, t\}$, for some $t \in \mathbb{N}$, and every $i \in \{1, \ldots, n\}$. Let $D = \sum_{j=1}^{t} c_j D_j^{l_j}$ where $c_1, \ldots, c_t \in \mathbb{R}$. Then Theorem 3.2.1 holds with each D^l replaced by D.*

Proof. The proof is identical to that of Theorem 3.2.1, where every D^l is replaced by D. ☐

As in Chapter 2, the definition of a remainder polynomial and of a zero polynomial will be provided so that the order of interpolation can be established.

Definition 3.2.3. Suppose $(\mathbf{x}, \mathbf{D})$ is consistent of degree m. Let $x \in \mathbb{R}^s$ and let $D = \sum_{j=1}^{t} D_j^{l_j}$ be a differential operator satisfying the hypothesis of Corollary 3.2.2. A *minimal remainder polynomial at* (x, D) is a polynomial $R_{x,D} \colon \mathbb{R}^s \to \mathbb{R}$ of minimum degree such that

$$D_i^{k_i} R_{x,D}(x_i) = 0, \qquad 1 \leq i \leq n,$$

but

$$D R_{x,D}(x) = 1.$$

In the multivariate case, a zero polynomial through $(\mathbf{x}, \mathbf{D})$ is not uniquely determined by its degree.

Definition 3.2.4. Suppose $(\mathbf{x}, \mathbf{D})$ is consistent of degree m and $(L_1, \ldots, L_n)$ is a standard interpolation basis. Let p be a non-negative integer, let β be a multi-index such that $|\beta| = p$ and let $a \in \mathbb{R}^s$. Then the *β-zero polynomial through $(\mathbf{x}, \mathbf{D})$, centered at a,* is the polynomial $Q_{\beta,a} \colon \mathbb{R}^s \to \mathbb{R}$ defined for every $x \in \mathbb{R}^s$ by

$$Q_{\beta,a}(x) = \frac{(x-a)^\beta}{\beta!} - \sum_{i=1}^{n} \left(D_i^{k_i} \frac{(x-a)^\beta}{\beta!}\bigg|_{x_i} \right) L_i(x).$$

Proposition 3.2.5. *If $Q_{\beta,a}$ is a β-zero polynomial through $(\mathbf{x}, \mathbf{D})$, then*

$$D_i^{k_j} Q_{\beta,a}(x_j) = 0$$

for every $j \in \{1, \ldots, n\}$.

Proof. Observe that

$$\begin{aligned} D_i^{k_j} Q_{\beta,a}(x_j) &= D_i^{k_j} \frac{(x-a)^\beta}{\beta!}\bigg|_{x_i} - \sum_{i=1}^{n} \left(D_i^{k_i} \frac{(x-a)^\beta}{\beta!}\bigg|_{x_i} \right) D_i^{k_j} L_i(x_j) \\ &= D_i^{k_j} \frac{(x-a)^\beta}{\beta!}\bigg|_{x_i} - \sum_{i=1}^{n} \left(D_i^{k_i} \frac{(x-a)^\beta}{\beta!}\bigg|_{x_i} \right) \delta_{ij} \\ &= 0. \end{aligned}$$

□

Conjecture 3.2.6. *Suppose $(\mathbf{x}, \mathbf{D})$ is consistent of degree m. Let $a \in \mathbb{R}^s$, $x \in \mathbb{R}^s$ and D be a linear combination of directional derivatives satisfying the hypothesis of Corollary 3.2.2. For any multi index β, let $Q_{\beta,a}$ be the β-zero polynomial through $(\mathbf{x}, \mathbf{D})$, centered at a. Let $r+1$ be the degree of a minimal remainder polynomial at (x, D). If $r+1 > m$ then*

- *(i) $r+1$ is the smallest non-negative integer p such that there is a multi index β, with $|\beta| = p$, for which the quantity $DQ_{\beta,a}(x)$ is non zero,*
- *(ii) if $r+1 = p$, then*

  $$R_{x,D} := \frac{Q_{\beta,a}}{DQ_{\beta,a}(x)}$$

  is a minimal remainder polynomial at (x, D) and for any $y \in \mathbb{R}^s$,

  $$\frac{1}{D^\beta R_{x,D}(y)} = DQ_{\beta,a}(x),$$

  for any β with $|\beta| = p$ and $D^\beta Q_{\beta,a}(x) \neq 0$.

Conjecture 3.2.6 is the natural extension of Corollary 2.2.6 to multivariate interpolation, however it is not as clear to proceed with the proof, as one cannot as simply apply Theorem 3.2.1 to conclude that a remainder polynomial is a linear combination of β-zero polynomials. One possibility is to show that the set of all β-zero polynomials forms a spanning set for the collection of polynomials which are zero at the interpolation conditions $(\mathbf{x}, \mathbf{D})$, in which case Conjecture 3.2.6 follows immediately.

Chapter 4

Birkhoff Pre-Spline Interpolation on $\mathbb{R}$

This chapter is concerned with the extension of Theorem 2.2.1 to Birkhoff Spline interpolation. In general one may study splines with any combination of interpolation conditions, and any combination of continuity conditions, without enforcing any global degree of continuity. The extension of Theorem 2.2.1 to pre-splines is motivated by the application of splines to approximations of solutions of differential equations.

4.1 Preliminary Definitions

Typically, one associates a spline with a function, however, the important properties of splines are often derived from the underlying polynomials that form the spline. This motivates the definition of a pre-spline.

Definition 4.1.1. A (one dimensional) *real Birkhoff Pre-Spline* is a sequence $\mathbf{P} = (P_1, \ldots, P_N)$ of polynomials on $\mathbb{R}$ that satisfy the interpolation conditions

$$P_{\alpha_i}^{(k_i)}(x_i) = c_i \qquad 1 \leq i \leq n,$$

and additionally satisfy the continuity conditions

$$P_{\beta_i}^{(k_i)}(x_i) = P_{\gamma_i}^{(k_i)}(x_i) \qquad n+1 \leq i \leq n+t,$$

$\alpha_i, \beta_i, \gamma_i \in \{1, \ldots, N\}$, $x_i \in \mathbb{R}$, $k_i \in \mathbb{Z}_{\geq 0}$, $c_i \in \mathbb{R}$. If all these conditions are satisfied, then $\mathbf{P}$ is a real Birkhoff Pre-Spline *through* $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{x}, \mathbf{k}, \mathbf{c})$; if additionally $c_i = f^{(k_i)}(x_i)$ for each $i \in \{1, \ldots, n\}$, then $\mathbf{P}$ *interpolates* f at $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{x}, \mathbf{k})$.

Within this definition, bold face is used to represent the vectors of interpolation values

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \ \boldsymbol{\beta} = \begin{bmatrix} \beta_{n+1} \\ \vdots \\ \beta_{n+t} \end{bmatrix}, \ \boldsymbol{\gamma} = \begin{bmatrix} \gamma_{n+1} \\ \vdots \\ \gamma_{n+t} \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_{n+t} \end{bmatrix}, \ \mathbf{k} = \begin{bmatrix} k_1 \\ \vdots \\ k_{n+t} \end{bmatrix} \text{ and } \mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

The majority of results in this chapter are established in terms of pre-splines; however, for the sake of clarity the association with splines is provided. While a spline often requires some notion of continuity, or at least local continuity, the following definition does not have any such requirements.

Definition 4.1.2. A *real Birkhoff Spline* is a function on $\mathbb{R}$ defined for each $x \in \mathbb{R}$ by

$$x \mapsto P_{\mathcal{I}(x)}(x),$$

where $(P_1, \ldots, P_N)$ is a real Birkhoff Pre-Spline and $\mathcal{I} \colon \mathbb{R} \to \{1, \ldots, N\}$ is an arbitrary function.

Hence, any results developed for pre-splines may extend to splines. To properly establish the theory of pre-splines presented in this section, some properties of sequences of polynomials needs to be established.

Definition 4.1.3. The set of all sequences of length N, of polynomial functions on $\mathbb{R}$, will be denoted by $\boldsymbol{\mathcal{P}}_N$; that is, let

$$\boldsymbol{\mathcal{P}}_N = \{(P_1, \ldots, P_N) \mid P_i \text{ is a polynomial on } \mathbb{R}\}.$$

The space $\boldsymbol{\mathcal{P}}_N$ is endowed with

a) the binary operation $+$ of component-wise addition,

b) the binary operation $\cdot$ of component-wise multiplication,

c) scalar multiplication $\cdot$, $c \cdot (P_1, \ldots, P_n) \mapsto (c \cdot P_1, \ldots, c \cdot P_n)$,

d) the deg (degree) operation for which $\deg(P_1, \ldots, P_n) := (\deg(P_1), \ldots, \deg(P_N))$ and the corresponding partial order, $\leq$, of degrees defined by

$$\deg(P_1, \ldots, P_N) \leq \deg(Q_1, \ldots, Q_N) \iff \deg(P_i) \leq \deg(Q_i), \ \forall i \in \{1, \ldots, N\}.$$

The space $\boldsymbol{\mathcal{P}}_N$ is algebraically the graded $\mathbb{R}$ algebra $(\mathbb{R}[x])^N$ with a partial order on the degree of polynomial sequences. Within these operations, the pre-spline interpolation problem can be framed.

Definition 4.1.4. The collection of nodes $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{x}, \mathbf{k})$ is *consistent of degree* $\mathbf{m} = (m_1, \ldots, m_N)$ if for any set of n constants $\mathbf{c}$ there exists a real Birkhoff pre-spline through $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{x}, \mathbf{k}, \mathbf{c})$ of degree $\mathbf{m}$, and $\mathbf{m}$ is a minimal such degree. If $\sum_{i=0}^{N} m_i = n + t - N$, then the collection $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{x}, \mathbf{k})$ is *consistent.*

Again, in the case that the interpolation problem is sufficiently consistent, a natural interpolation basis presents itself.

Definition 4.1.5. If $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{x}, \mathbf{k})$ is consistent of degree $\mathbf{m} = (m_1, \ldots, m_N)$ then a *standard interpolation basis* is an ordered set of degree at most $\mathbf{m}$ real Birkhoff Pre-Splines $(\mathbf{L}_1, \ldots, \mathbf{L}_n)$–where for each $i \in \{1, \ldots, n\}$, $\mathbf{L}_i = (L_{i,1}, \ldots, L_{i,N})$–that satisfy

$$L_{i,\alpha_i}^{(k_j)}(x_j) = \delta_{ij} \qquad 1 \leq i, j \leq n$$

and additionally satisfy

$$L_{j,\beta_i}^{(k_i)}(x_i) = L_{j,\gamma_i}^{(k_i)}(x_i) \qquad 1 \leq j \leq n,\ n+1 \leq i \leq n+t.$$

Existence of such a collection of pre-splines once again follows from the observation that $\mathbf{L}_i$ is a degree at most $\mathbf{m}$ Pre-Spline through $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{x}, \mathbf{k}, (\delta_{ij})_{j=1}^{n})$, as in the single polynomial case.

Proposition 4.1.6. *If* $(\mathbf{L}_1, \ldots, \mathbf{L}_n)$ *is a standard interpolation basis of the system* $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{x}, \mathbf{k})$*, then*

$$\mathbf{P} = \sum_{i=1}^{n} f^{(k_i)}(x_i) \cdot \mathbf{L}_i$$

interpolates f *at* $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{x}, \mathbf{k})$*.*

Proof. Let $j \in \{1, \ldots, n\}$, then

$$\begin{aligned} P_{\alpha_j}^{(k_j)}(x_j) &= \sum_{i=1}^{n} f^{(k_i)}(x_i) L_{i,\alpha_j}^{(k_j)}(x_j) \\ &= \sum_{i=1}^{n} f^{(k_i)}(x_i) \delta_{ij} \\ &= f^{(k_j)}(x_j). \end{aligned}$$

Now let $j \in \{n+1, \ldots, n+t\}$, then

$$P_{\beta_j}^{(k_j)}(x_j) = \sum_{i=1}^{n} f^{(k_j)}(x_i) L_{i,\beta_j}^{(k_i)}(x_j)$$

$$= \sum_{i=1}^{n} f^{(k_j)}(x_i) L_{i,\gamma_j}^{(k_i)}(x_j)$$
$$= P_{\gamma_j}^{(k_j)}(x_j).$$

□

4.2 Birkhoff Pre-Spline Interpolation Error

Now, we extend Theorem 2.2.1 to Birkhoff Pre-Splines.

Theorem 4.2.1. *Let $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{x}, \mathbf{k})$ be consistent of degree $\mathbf{m}$, let $(\mathbf{L}_1, \ldots, \mathbf{L}_n)$ be a standard interpolation basis, let $l \geq 0$, $x \in \mathbb{R}$ such that $(x, l) \neq (x_i, k_i)$ for every $i \in \{1, \ldots, n\}$, let $a \in \mathbb{R}$ and let $j \in \{1, \ldots, N\}$. Suppose r is the least non-negative integer such that there is a degree $r+1$ polynomial R that is interpolated by a degree at most $\mathbf{m}$ pre-spline $\mathbf{Q} = (Q_1, \ldots, Q_N)$ at $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{x}, \mathbf{k})$ but for which*

$$Q_j^{(l)}(x) \neq R^{(l)}(x).$$

Let $f \colon \mathbb{R} \to \mathbb{R}$ be $r+1$ times differentiable on an open interval containing x, a, and x_i for each $x_i \in \{1, \ldots, n\}$. If $m_j < r+1$ and $\mathbf{P} = (P_1, \ldots, P_N)$ is given as in Proposition 4.1.6, then

$$f^{(l)}(x) - P_j^{(l)}(x) = f^{(r+1)}(\xi(x)) \frac{(x-a)^{r+1-l}}{(r+1-l)!} - \sum_{i=1}^{n} \left[f^{(r+1)}(\xi_i(x_i)) \frac{(x_i - a)^{r+1-k_i}}{(r+1-k_i)!} \right] L_{i,j}^{(l)}(x)$$

for some $\xi(x)$ between x and a, some $\xi_i(x_i)$ between x_i and a for each $i \in \{1, \ldots, n\}$.

Proof. The proof is nearly identical to that of Theorem 2.2.1. Let $\mathbf{P}_T$ be the degree at most $\mathbf{m}$ pre-spline that interpolates $T_{f,r,a}$ at $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{x}, \mathbf{k})$ that is given by the expression from Proposition 4.1.6. By assumption,

$$P_{T,j}^{(l)}(x) = T_{f,r,a}^{l}(x),$$

since the degree of $T_{f,r,a}$ is less than $r+1$. Therefore,

$$P_j^{(l)}(x) = \sum_{i=1}^{n} f^{(k_i)}(x_i) L_{j,i}^{(l)}(x)$$
$$= \sum_{i=1}^{n} [T_{f^{(k_i)},r-k_i,a}(x_i) + R_{f^{(k_i)},r-k_i,a}(x_i)] L_{j,i}^{(l)}(x)$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} T_{f,r,a}^{(k_i)}(x_i) L_{j,i}^{(l)}(x) + \sum_{i=1}^{n} R_{f^{(k_i)},r-k_i,a}(x_i) L_{j,i}^{(l)}(x) \\
&= P_{T,j}^{(l)}(x) + \sum_{i=1}^{n} [R_{f^{(k_i)},r-k_i,a}(x_i)] L_{j,i}^{(l)}(x) \\
&= T_{f,r,a}^{(l)}(x) + \sum_{i=1}^{n} [R_{f^{(k_i)},r-k_i,a}(x_i)] L_{j,i}^{(l)}(x) \\
&= f^{(l)}(x) - R_{f^{(l)},r-l,a}(x) + \sum_{i=1}^{n} R_{f^{(k_i)},r-k_i,a}(x_i) L_{j,i}^{(l)}(x).
\end{aligned}
$$

Substitution of explicit remainders from Definition 2.1.8 yields the desired result. $\square$

Chapter 5

Future Research

There are two main routes to continue the research presented within this thesis. The first is to apply the main error theorems Theorem 2.2.1, Theorem 3.2.1, and Theorem 4.2.1 to a broader range of Birkhoff interpolation methods, and to differential equations. The second is to generalize Theorem 2.2.1 to as broad of a context as possible.

5.1 Additional Applications of The Main Error Theorem

There are a plethora of potential further applications of Theorem 2.2.1. One could study finite difference methods in one dimension which arise from interpolation polynomials that are not Lagrange polynomials, i.e. finite difference methods which involve derivative values in addition to function values. Moreover, Birkhoff quadrature methods could be extended to higher order integrals, and in-turn lend themselves to approximation methods of solutions of higher order initial value problems. Similarly, one can approximate solutions of differential equations of any order, with mixed type boundary conditions by utilizing the notion of a Pre-Spline that fits the differential equation at a finite collection of nodes, rather than known function values.

Additionally, all of these concepts could be extended to multivariate contexts if the details of Chapter 3 are clarified. An immediate consequence of a proof for Conjecture 3.2.6 would be the calculation of multivariate asymptotic error. If such calculations follow in a similar manner as the uni-variate case, then one could perform a search for $\mathcal{O}(h^2)$ second difference methods in higher dimensions that fit derivative values at the boundary of some region, similar to the calculations of Section 2.4 and Section 2.5.

Finally the notion of a pre-spline is readily extended to the multivariate context. Of course, there are other known methods to find and optimize the error of Birkhoff interpolation methods, and associated splines such as those noted in [11].

5.2 Possible Extension of Main Error Theorem

It is worth noting that the algebraic nature of Theorem 2.2.1 suggests an extension to a larger context than multivariate interpolation. In particular, given a smooth and connected sub-manifold M of $\mathbb{R}^s$, one can study Birkhoff interpolation of functions $f\colon M \to \mathbb{R}$. There is a potentially natural way to approach this problem as an extension Theorem 2.2.1 and Theorem 3.2.1. One can take the collection of polynomials in s variables modulo the collection of polynomials that are identically zero on M as a basis for interpolation. Then identify minimal remainder polynomials, and realize an extension of Taylor's theorem by utilizing the exponential map on M. In this manner, it might be possible to generalize the result from Chapter 2 to interpolation on M.

Moreover, if Theorem 2.2.1 extends to interpolation on smooth manifolds, then a version of the pre-spline interpolation error Theorem 4.2.1 would also extend to smooth manifolds. Finally, one could easily rephrase the problem of Birkhoff Interpolation, which fits a polynomial to function and derivative values, to the problem of fitting a polynomial to linear differential conditions. In such a case it is straightforward to extend Theorem 2.2.1 and Theorem 3.2.1.

Bibliography

[1] K. Atkinson and A. Sharma. A partial characterization of poised hermite-birkhoff interpolation problems. *SIAM Journal on Numerical Analysis*, 6(2):230–235, 1969.

[2] George David Birkhoff. General mean value and remainder theorems with applications to mechanical differentiation and quadrature. *Transactions of the American Mathematical Society*, 7(1):107–136, 1906.

[3] Richard L Burden, J Douglas Faires, and Annette M Burden. *Numerical analysis*. Cengage learning, 2015.

[4] R. Devore, A. Meir, and A. Sharma. Strongly and weakly non-poised h-b interpolation problems. *Canadian Journal of Mathematics*, 25(5):1040–1050, 1973.

[5] David Ferguson. The question of uniqueness for g. d. birkhoff interpolation problems. *Journal of Approximation Theory*, 2(1):1–28, 1969.

[6] Gerald B Folland. Higher-order derivatives and taylor's formula in several variables. *Preprint*, pages 1–4, 2005.

[7] Arieh Iserles. *A First Course in the Numerical Analysis of Differential Equations*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2 edition, 2008.

[8] Arieh Iserles and Gustaf Söderlind. Global bounds on numerical error for ordinary differential equations. *Journal of Complexity*, 9(1):97–112, 1993.

[9] Kurt Jetter. A new class of gaussian quadrature formulas based on birkhoff type data. *SIAM Journal on Numerical Analysis*, 19(5):1081–1089, 1982.

[10] Samuel Karlin and John M Karon. On hermite-birkhoff interpolation. *Journal of Approximation Theory*, 6(1):90–115, 1972.

[11] G. G. Lorentz, S. D. Riemenschneider, and K. Jetter. *Birkhoff Interpolation*. Number v. 19 in Section, Interpolation and Approximation. Cambridge University Press, 1984.

[12] George G. Lorentz and Karl Zeller. Monotone approximation by algebraic polynomials. *Transactions of the American Mathematical Society*, 149:1–18, 1970.

[13] G.G Lorentz. Birkhoff interpolation and the problem of free matrices. *Journal of Approximation Theory*, 6(3):283–290, 1972.

[14] R.A. Lorentz. Uniqueness of best approximation by monotone polynomials. *Journal of Approximation Theory*, 4(4):401–418, 1971.

[15] Encyclopedia of Mathematics. Multi-index notation. https://encyclopediaofmath.org/wiki/Multi-index_notation. Accessed: 2023-04-11.

[16] Walter Rudin et al. *Principles of mathematical analysis*, volume 3. McGraw-hill New York, 1976.

[17] A. Sharma. Some poised and nonpoised problems of interpolation. *SIAM Review*, 14(1):129–151, 1972.

[18] Michael Spivak. *Calculus on manifolds: a modern approach to classical theorems of advanced calculus*. CRC press, 2018.