Exploring the Rate of Convergence of Approximations to the Riemann Integral

Lukas Owens

May 17, 2014

Contents

Abstract

There are many well known ways of approximating the value of the Riemann integral of a real-valued function. The endpoint rules, the midpoint rule, the trapezoid rule and Simpson's rule each produce sequences that converge to the value of the integral. Some methods are better than others and this paper seeks to quantify how quickly the error of each approximation converges to zero. After studying these classical results, we extend them in a variety of ways. Sums and products of functions with different tags and improper integrals are considered, as well as higher order expansion of the error term.

1 Introduction

1.1 Preliminaries

We will begin by recording some definitions and notation that will be used frequently throughout this paper.

1.1.1 The Riemann Integral

The Riemann integral arises as a method of determining the area under a curve and is a central tool in the analysis of functions of a real variable. We will use the definitions in [3] and the reader is referred to that source for the proofs of important theorems.

Suppose that we have a function $f\colon [a,b] \to \mathbb{R}$. We want to approximate the area under the curve. We are interested in the "signed area", that is, area above the x-axis and below the curve is positive area and area under the x-axis and above the curve will be negative area. The following definitions will make this idea precise.

- Fix some large number n and divide the interval $[a,b]$ into n subintervals $[x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n]$ so that $x_0 = a$ and $x_n = b$. We call such a collection of subintervals P a **partition** of $[a,b]$. These subintervals can have any length, but the definition makes it clear that the sum of all the lengths must be equal to $b-a$. We call the length of the largest subinterval the **norm** of the partition, written $\|P\| = \max\{x_k - x_{k-1} : 1 \leq k \leq n\}$

- For every k satisfying $1 \leq k \leq n$, chose a number $t_k \in [x_{k-1}, x_k]$. We call t_k a **tag** and the collection ${}^tP$ of all pairs of subintervals and their tags a **tagged partition**.

- Now we observe that, for each k with $1 \leq k \leq n$, a rectangle with base $[x_{k-1}, x_k]$ and height $f(x_k)$ has area roughly equal to the area under the curve $y = f(x)$ from $x = x_{k-1}$ to $x = x_k$. The area of this rectangle is given by $f(t_k)(x_k - x_{k-1})$. When we sum these areas over all the subintervals in the partition ${}^tP$, we get an approximation to the area under the curve $y = f(x)$ from $x = a$ to $x = b$ which we will call the **Riemann sum** of f with respect to ${}^tP$, given by

$$S(f, {}^tP) = \sum_{k=1}^{n} f(t_k)(x_k - x_{k-1}).$$

As the norm of the tagged partition gets smaller, we expect that the Riemann sum of f associated with that tagged partition converges to the area under the curve. This motivates the following definition

Definition 1.1. A function $f\colon [a,b] \to \mathbb{R}$ is **Riemann integrable** on $[a,b]$ if there exists a number L such that for every $\epsilon > 0$ there exists a $\delta > 0$ such that $|S(f, {}^tP) - L| < \epsilon$ for every tagged partition ${}^tP$ of $[a,b]$ with norm less that δ. If such an L exists, we call it the **Riemann integral** of f from a to b and we denote it by $\int_a^b f(x)\,dx$ or simply $\int_a^b f$.

The above way of defining the integral is only one approach. An alternate, and equally useful, way of defining the integral, called the Darboux integral, is presented as follows:

- Suppose f is a bounded function on $[a,b]$ and let $P = \{[x_0, x_1] \ldots [x_{n-1}, x_n]\}$ be a partition of that interval. For each subinterval $[x_{k-1}, x_k]$ in P, consider the two rectangles with shared base $[x_{k-1}, x_k]$ and heights

$$M_k = \sup_{t \in [x_{k-1}, x_k]} f(t) \text{ and } m_k = \inf_{t \in [x_{k-1}, x_k]} f(t)$$

  The rectangle with height M_k is the smallest rectangle with base $[x_{k-1}, x_k]$ and containing the curve $y = f(x)$ from $x = x_{k-1}$ to $x = x_k$. The rectangle with height m_k is the largest rectangle with that base and lying between that curve and the x-axis. The areas of those two rectangles are $M_k(x_k - x_{k-1})$ and $m_k(x_k - x_{k-1})$, respectively.

- If we add the areas of these two sets of rectangles, we get two approximations for the integral of the function. We will call these the **upper** and **lower sums** of f with respect to P, written

$$U(f,P) = \sum_{k=1}^{n} M_k(x_k - x_{k-1}) \text{ and } L(f,P) = \sum_{k=1}^{n} m_k(x_k - x_{k-1})$$

It is immediate from the definition that

$$L(f,P) \leq U(f,P).$$

for any partition P. As the partition gets finer, we would expect the upper and lower sums to converge to the same number. This motivates the following definition.

Definition 1.2. A bounded function $f : [a,b] \to \mathbb{R}$ is **Darboux integrable** if

$$\sup L(f,P) = \inf U(f,P)$$

where both the infimum and the supremum are taken over all partitions P of $[a,b]$. The common value is called the Darboux integral of f from a to b.

Proof of the fact that a function is Riemann integrable if and only if it is Darboux integrable and the two integrals have the same value to [3]. In the remainder we will drop the modifiers 'Riemann' and 'Darboux' and simply refer to the 'integral' of a function.

1.1.2 Asymptotic Notation

In this paper, we will be exploring the long term behavior of various sequences. For a sequence that converges to a limit, either finite or infinite, we would like to be able to quantify how quickly it converges to that value and compare its convergence to other sequences. For example, consider the sequences whose nth terms are

$$\ln n,\ n^2,\ n^2 + 2n,\ n!$$

All four of these sequence diverge to infinity as $n \to \infty$. However, our intuition tells us that some go to infinity faster than the others. The sequence $\{\ln n\}$ grows very slowly when compared to the sequence $\{n^2\}$ and $\{n!\}$ grows much faster. The sequences $\{n^2\}$ and $\{n^2 + 2n\}$ both grow at about the same rate, as the n^2 term begins to 'dominate' as n gets large.

These ideas about the relative rates of growth of these sequences are vague and need to be made precise. The notation introduced in this section allows us to do just this.

Definition 1.3. If f and g are real valued functions and c is a real number, we say that $f(x)$ is **of order** $g(x)$, written $f(x) = O(g(x))$ as $x \to c$, if there exist constants δ and M so that $|f(x)| \leq Mg(x)$ whenever $|x - c| \leq \delta$. We say $f(x) = O(g(x))$ as $x \to \infty$ if there are constants x_0 and M so that $|f(x)| \leq Mg(x)$ whenever $x > x_0$.

When the limit is clear from context, we will omit saying 'as $x \to c$'. The equals sign in $f(x) = O(g(x))$ is not a true statement of equality but rather a convenient notation. This notation allows us compactly make statement about the limiting behavior of a function or sequence. For example, consider the third degree Taylor polynomial of the sine function. We know from Taylor's theorem that

$$\sin(x) = x - \frac{x^3}{3!} + O(x^5) \text{ as } x \to 0.$$

This is a much simpler way of saying

$$\sin(x) = x - \frac{x^3}{3!} + g(x) \text{ where } g \text{ is a function satisfying } |g(x)| \leq Mx^5 \text{ for values of } x \text{ sufficiently close to } 0.$$

This convention will be used throughout the paper and its properties will be further explained when necessary. The statement $f(x) = O(g(x))$ means that the order of f is, in a sense, 'less than or equal to' the order of g in the limit. The next two definitions clarify the cases when the order of f is 'stricly smaller than' or 'equal to' the order of g in the limit.

Definition 1.4. Let f and g be functions of a real variable and let c be an extended real number. We say that $f(x)$ is of smaller order than $g(x)$ as $x \to c$, written $f(x) = o(g(x))$, if $\lim_{x\to c} f(x)/g(x) = 0$.

Definition 1.5. If f and g are functions of a real variable and c is an extended real number, we say that $f(x)$ is **asymptotically equivalent** to $g(x)$, written $f(x) \sim g(x)$, as $x \to c$ if $\lim_{x\to c} f(x)/g(x) = 1$.

Returning to our earlier example, we can rephrase our claims about the order of our sequences using the new notation as follows:

$$\ln n = o(n^2),\ n^2 \sim (n^2 + 2n),\ \text{and}\ n^2 + 2n = o(n!)$$

1.2 Motivation

It follows from our definition of the integral that if $f : [a, b] \to \mathbb{R}$ is Riemann integrable and $\{{}^tP_n\}$ is a sequence of tagged partitions with $\|{}^tP_n\| \to 0$ as $n \to \infty$, then the sequence of Rieman sums $\{S(f,{}^t P_n)\}$ converges to $\int_a^b f$.

For example, consider the function $f(x) = e^x$ on the interval $[0, 1]$. For any positive integer n, let ${}^tP_n$ be the partition whose subintervals are $[(k-1)/n, k/n]$ with tag k/n for $1 \le k \le n$. This gives us the sequence of Riemann sums

$$S(f,{}^t P_n) = \sum_{k=1}^{n} \frac{1}{n} f\left(\frac{k}{n}\right) = \frac{1}{n}\left(e^{1/n} + e^{2/n} + \cdots + e^{n/n}\right)$$

This is known as the nth 'right endpoint approximation' of the integral of f. To discover its limiting behavior, we start by observing that the term inside the parentheses is a geometric sum with common ratio $e^{1/n}$, which allows us to write

$$S(f,{}^t P_n) = \frac{e^{1/n} - e^{(n+1)/n}}{n(1 - e^{1/n})} = (e-1)\frac{e^{1/n}}{n(e^{1/n} - 1)} = (e-1)\frac{1}{n(1 - e^{-1/n})}$$

By using the Taylor expansion $e^x = 1 + x + x^2/2 + O(x^3)$ as $x \to 0$, we have

$$S(f,{}^t P_n) = \frac{e-1}{n\left(\frac{1}{n} - \frac{1}{2n^2} + O\left(\frac{1}{n^3}\right)\right)} = (e-1)\frac{1}{1 - \frac{1}{2n} + O\left(\frac{1}{n^2}\right)}$$

The geometric series gives us the expansion

$$\frac{1}{1-x} = 1 + x + x^2 + O(x^3) \ \text{ as } \ x \to 0$$

which simplifies the above expression to

$$\begin{aligned} S(f,{}^t P) &= (e-1)\left[1 + \frac{1}{2n} + O\left(\frac{1}{n^2}\right) + O\left(\left(\frac{1}{2n} + O\left(\frac{1}{n^2}\right)\right)^2\right)\right] \\ &= (e-1) + \frac{e-1}{2n} + O\left(\frac{1}{n^2}\right) \end{aligned}$$

This example shows that the integral of f over $[0, 1]$ is equal to $e - 1$, a fact that we could have deduced from the fundamental theorem of calculus. Moreover, we now have a sense of 'how quickly' the Riemann sums converge to that value. The error term

$$\int_0^1 f - S(f,{}^t P) = \frac{1-e}{2n} + O\left(\frac{1}{n^2}\right)$$

converges to 0 asymptotically with $(1 - e)/2n$, that is

$$\int_0^1 f - S(f,{}^t P) \sim \frac{1-e}{2n}$$

This example typifies the general situation. For many methods of estimating an integral, we are able to quantify how quickly those methods converges to the value of the integral. This is the goal of Section 2 of this paper.

2 Common Integral Approximations and the Order of Convergence

Suppose that $f : [c, d] \to \mathbb{R}$ is Riemann integrable. There are many ways to estimate the value of $\int_c^d f$. Generally, we want to approximate the function f by a simpler function whose integral is easily calculated. The endpoint and midpoint rules approximate f by a rectangle; the trapezoid rule uses, unsurprisingly, trapezoids; and Simpson's rule fits a quadratic intersecting f at three points. The smaller the length of the interval $[c, d]$ is, the smaller the error in this estimate. Suppose that f is now defined on some interval $[a, b]$. Fix a large value of n. We will divide $[a, b]$ into n subintervals of equal length and approximate the value of the integral over each subinterval using one of our estimates. In the rest of this section we will refer to this interval division by

$$I_n^k = \left[a + \frac{b-a}{n}(k-1), a + \frac{b-a}{n}k\right] \quad , \quad 1 \leq k \leq n$$

and we will denote the endpoints of these intervals by

$$x_n^k = a + \frac{b-a}{n}k \quad , \quad 0 \leq k \leq n$$

Suppose that A_n^k is an approximation of the integral of f over the interval I_n^k, using one of the methods we mentioned above. Then $A_n = \sum_{k=1}^n A_n^k$ is our approximation for $\int_a^b f$. We are concerned with the quantity

$$\Delta_n = \int_a^b f(x)\,dx - A_n$$

which represents the error in this approximation. Clearly, we should have $\Delta_n \to 0$ as $n \to \infty$ so that A_n converges to the value of the integral, but we also want to consider how quickly this number converges to 0. We will examine this question in the following sections.

2.1 The Endpoint Rules

The simplest way to approximate the value of an integral is with the right- and left-endpoint rules. These are respectively given by

$$R_n = \frac{b-a}{n}\sum_{k=1}^n f(x_n^k), \quad L_n = \frac{b-a}{n}\sum_{k=1}^n f(x_n^{k-1})$$

Each of these numbers is simply a Riemann sum of f using either the right- or left-endpoints as tags. Then we know already that $L_n, R_n \to \int_a^b f$ as $n \to \infty$. The following theorem characterizes how quick this approximation converges.

Theorem 2.1. Suppose that the function f has bounded and integrable derivative on $[a, b]$. Let L_n be the left-endpoint approximation to $\int_a^b f$ and let $\Delta_n = \int_a^b f - L_n$. Then we have

$$\Delta_n \sim \frac{f(b) - f(a)}{2n}(b-a)$$

Proof. First, observe that, by splitting up the integral, Δ_n can be written as

$$\begin{aligned}\Delta_n &= \sum_{k=0}^{n-1}\int_{x_n^k}^{x_n^{k+1}} f(x)\,dx - \frac{b-a}{n}\sum_{k=0}^{n-1} f(x_n^k)\\ &= \sum_{k=0}^{n-1}\int_{x_n^k}^{x_n^{k+1}} (f(x) - f(x_n^k))\,dx\end{aligned}$$

Recall integrable functions are always bounded. This allows us to define

$$s_n^k = \inf_{x \in I_n^k} f'(x) \text{ and } S_n^k = \sup_{x \in I_n^k} f'(x)$$

for each interval I_n^k. The mean value theorem implies that

$$s_n^k(x - x_n^k) \leq f(x) - f(x_n^k) \leq S_n^k(x - x_n^k)$$

Applying this to our above formula for Δ_n, we get

$$\sum_{k=0}^{n-1} \int_{x_n^k}^{x_n^{k+1}} s_n^k(x - x_n^k)\,dx \leq \Delta_n \leq \sum_{k=0}^{n-1} \int_{x_n^k}^{x_n^{k+1}} S_n^k(x - x_n^k)\,dx.$$

A simple computation shows that

$$\int_{x_n^k}^{x_n^{k+1}} (x - x_n^k)\,dx = \frac{(b-a)^2}{2n^2}$$

which gives us

$$\frac{b-a}{2n}\left(\frac{b-a}{n}\right)\sum_{k=0}^{n-1} s_n^k \leq \Delta_n \leq \frac{b-a}{2n}\left(\frac{b-a}{n}\right)\sum_{k=0}^{n-1} S_n^k.$$

But we know that

$$\lim_{n\to\infty} \frac{b-a}{n}\sum_{k=0}^{n-1} s_n^k = \lim_{n\to\infty} \frac{b-a}{n}\sum_{k=0}^{n-1} S_n^k = \int_a^b f'(x)\,dx = f(b) - f(a)$$

By the squeeze theorem, it follows that

$$\lim_{n\to\infty} n\Delta_n = \frac{b-a}{2n}(f(b) - f(a))$$

which establishes the result

□

A similar result holds for the right endpoint approximation.

Theorem 2.2. Suppose that f satisfies the same hypotheses as the previous theorem. Let R_n be the right endpoint approximation of f and let $\Delta_n = \int_a^b f - R_n$. Then we have

$$\Delta_n \sim \frac{f(a) - f(b)}{2n}(b-a)$$

Proof. Observe that

$$R_n = L_n + \frac{f(b) - f(a)}{n}(b-a)$$

Then, using the previous result, we have

$$\begin{aligned}
\Delta_n = \int_a^b f - R_n &= \left(\int_a^b f - L_n\right) - \frac{f(b) - f(a)}{n}(b-a) \\
&\sim \frac{f(b) - f(a)}{2n}(b-a) - \frac{f(b) - f(a)}{n}(b-a) \\
&= \frac{f(a) - f(b)}{2n}(b-a)
\end{aligned}$$

□

The requirement that f is differentiable and that its derivative is integrable is a strong condition to require of our function. Recall that a function $f : [a,b] \to \mathbb{R}$ is said to be of **bounded variation** if

$$V = \sup \sum_{k=1}^{n} |f(x_k) - f(x_{k-1}| < \infty$$

where the supremum is taken over all partitions $[x_0, x_1], \ldots, [x_{n-1}, x_n]$ of $[a,b]$. The number V is called the **total variation** of f on $[a,b]$. The following theorem shows that if f is of bounded variation, we can still establish an upper bound on the size of the error of the endpoint approximations.

Theorem 2.3. Suppose that the function f is of bounded variation on the interval $[a,b]$ and let V be the total variation of f on $[a,b]$. Let Δ_n be the error $\int_a^b f - R_n$ of the right endpoint approximation. Then we have

$$|\Delta_n| \le V\frac{b-a}{n}$$

Proof. If we split up the integral as we did before and perform a change of variables, we obtain

$$\begin{aligned}
\Delta_n &= \sum_{k=1}^{n} \int_{x_n^{k-1}}^{x_n^k} (f(x) - f(x_n^k))\,dx \\
&= \sum_{k=1}^{n} \int_0^{\frac{b-a}{n}} (f(x_n^{k-1}+t) - f(x_n^k))\,dt \\
&= \int_0^{\frac{b-a}{n}} \sum_{k=1}^{n} (f(x_n^{k-1}+t) - f(x_n^k))\,dt
\end{aligned}$$

Then by the triangle inequality, we have

$$\begin{aligned}
|\Delta_n| &\le \int_0^{\frac{b-a}{n}} \sum_{k=1}^{n} |f(x_n^{k-1}+t) - f(x_n^k)|\,dt \\
&\le \int_0^{\frac{b-a}{n}} V\,dt = V\frac{b-a}{n}
\end{aligned}$$

□

2.2 The Midpoint Rule

With the left- and right-endpoint approximations, we chose the endpoints x_n^{k-1} and x_n^k as the tags for a Riemann sum. The next choice of tag to consider is the midpoint of each interval, $(x_n^{k-1}+x_n^k)/2$. Specifically, we have for each positive integer n, the midpoint approximation

$$\begin{aligned}
M_n &= \frac{b-a}{n}\sum_{k=1}^{n} f\left(\frac{x_{k-1}+x_k}{2}\right) \\
&= \frac{b-a}{n}\sum_{k=1}^{n} f\left(a + (2k-1)\frac{b-a}{2n}\right)
\end{aligned}$$

As the following theorem shows, this turns out to be a significantly better approximation of the integral, provided that we make some stronger assumptions about the differentiability of the function.

Theorem 2.4. Assume that f is twice differentiable and that f'' is integrable on $[a,b]$. Then the error of the midpoint rule

$$\Delta_n = \int_a^b f(x)\,dx - M_n$$

satisfies

$$\Delta_n \sim \frac{f'(b)-f'(a)}{24n^2}(b-a)^2$$

Proof. First, set $m_k^n = (x_n^k + x_n^{k-1})/2$ for $1 \le k \le n$. Then the quantity we are interested in becomes

$$\Delta_n = \int_a^b f(x)\,dx - \frac{b-a}{n}\sum_{k=1}^{n} f(m_n^k) = \sum_{k=1}^{n}\int_{x_n^{k-1}}^{x_n^k} \left[f(x) - f(m_n^k)\right]\,dx$$

Suppose $x \in [x_n^{k-1}, x_k)$. Because f is twice-differentiable, we expand $f(x)$ around $f(m_n^k)$ in a second order Taylor polynomial as

$$f(x) = f(m_n^k) + f'(m_n^k)(x-m_n^k) + \frac{1}{2}f''(\zeta_x)(x-m_n^k)^2$$

where ζ_x is some number between x and m_n^k that depends on x, n and k. If we let

$$s_n^k = \inf_{x \in I_n^k} f''(x), \quad S_n^k = \sup_{x \in I_n^k} f''(x)$$

then we have

$$f'(m_n^k)(x - m_n^k) + \frac{1}{2} s_k (x - m_n^k)^2 \le f(x) - f(m_n^k) \le f'(m_n^k)(x - m_n^k) + \frac{1}{2} S_k (x - m_n^k)^2$$

This establishes

$$\sum_{k=1}^{n} \int_{x_n^{k-1}}^{x_n^k} \left(f'(m_n^k)(x - m_n^k) + \frac{1}{2} s_k (x - m_n^k)^2 \right) dx \le \Delta_n \le \sum_{k=1}^{n} \int_{x_n^{k-1}}^{x_n^k} \left(f'(m_n^k)(x - m_n^k) + \frac{1}{2} S_k (x - m_n^k)^2 \right) dx$$

Since $\int_{x_n^{k-1}}^{x_k} (x - m_n^k)\, dx = 0$ and

$$\int_{x_n^{k-1}}^{x_k} (x - m_n^k)^2\, dx = \frac{1}{3} \Big[(x - m_n^k)^3 \Big]_{x_n^{k-1}}^{x_n^k} = \frac{(b-a)^3}{12n^3}$$

the above inequalities reduce to

$$\frac{(b-a)^3}{24n^3} \sum_{j=1}^{n} s_j \le \Delta_n' \le \frac{(b-a)^3}{24n^3} \sum_{j=1}^{n} S_j.$$

But since

$$\lim_{n\to\infty} \frac{b-a}{n} \sum_{j=1}^{n} s_k = \lim_{n\to\infty} \frac{b-a}{n} \sum_{j=1}^{n} S_k = \int_a^b f''(x)\, dx = f'(b) - f'(a)$$

by the squeeze theorem, we have

$$\lim_{n\to\infty} n^2 \Delta_n = \frac{1}{24} (b-a)^2 (f'(b) - f'(a))$$

which establishes the desired result. $\square$

2.3 The Trapezoid Rule

In the previous two sections, we estimated the area under a curve by a rectangle with a specified height. Alternatively, we can approximate the area by a trapezoid whose vertices are

$$(x_{k-1}, 0), \quad (x_k, 0), \quad (x_{k-1}, f(x_{k-1})), \quad (x_k, f(x_k)).$$

The area of such a trapezoid is given by

$$\frac{b-a}{n} \cdot \frac{f(x_k) + f(x_{k-1})}{2}$$

From this, we have the nth trapezoid rule approximation of the integral of a function f given by

$$T_n = \frac{b-a}{n} \sum_{k=1}^{n} \frac{f(x_k) + f(x_{k-1})}{2}$$

Observe that the trapezoid rule is the average of the left- and right-endpoint rules since $T_n = (R_n + L_n)/2$. It is not surprising, then, that the trapezoid rule is a closer approximation to the value of the integral than either of those two rules on their own. The following theorem makes this claim precise.

Theorem 2.5. If $f : [a,b] \to \mathbb{R}$ is twice differentiable and T_n is the nth trapezoid approximation to the integral of f on $[a,b]$, then the error term $\Delta_n = \int_a^b f - T_n$ satisfies

$$\Delta_n \sim \frac{(b-a)^2}{12n^2}(f'(a) - f'(b))$$

Proof. To reduce cluttering, we will use the notation that $x_k = a + k(b-a)/n$ for $0 \leq k \leq n$ and $h = (b-a)/n$ is the norm of the partition. Though h and x_k both depend on n, the value being referred to will be identified by context.

We begin by fixing a value of k with $0 \leq k \leq n-1$ and estimating the value of the integral $\int_{x_k}^{x_{k+1}} f$. We wish to express this integral in terms of its values at its two endpoints. First we estimate its value in terms of its behavior at the left endpoint. For any x, we have

$$f(x) = f(x_k) + (x - x_k)f'(x_k) + \frac{1}{2}(x - x_k)^2 f''(x_k) + O\left((x - x_k)^3\right) \quad \text{as } x \to x_k$$

When we make the change of variables $x - x_k = hs$, we get

$$f(x) + f(x_k) + hsf'(x_k) + \frac{1}{2}h^2 s^2 f''(x_k) + O\left(h^3 s^3\right)$$

and $dx = h\,ds$. This gives us

$$\begin{aligned}
\int_{x_k}^{x_{k+1}} f &= \int_0^1 \left(f(x_k) + hsf'(x_k) + \frac{1}{2}h^2 s^2 f''(x_k) + O\left(h^3 s^3\right)\right) h\,ds \\
&= hf(x_k) + \frac{h^2}{2}f'(x_k) + \frac{h^3}{6}f''(x_k) + O\left(h^4\right)
\end{aligned}$$

as $h \to 0$. Now we use another Taylor expansion, this time of the value of $f(x_{k+1})$, that is,

$$f(x_{k+1}) = f(x_k) + hf'(x_k) + \frac{h^2}{2}f''(x_k) + O\left(h^3\right)$$

as $h \to 0$. By solving this equation for $f'(x_k)$, we get

$$f'(x_k) = \frac{f(x_{k+1} - f(x_k)}{h} - \frac{h}{2}f''(x_k) + O(h^2)$$

Substituting this into the above expression for the integral, we get

$$\begin{aligned}
\int_{x_k}^{x_{k+1}} f &= hf(x_k) + \frac{h^2}{2}\left(\frac{f(x_{k+1}) - f(x_k)}{h} - \frac{h}{2}f''(x_k) + O(h^2)\right) + \frac{h^3}{6}f''(x_k) + O\left(h^4\right) \\
&= \frac{h}{2}\Big(f(x_k) + f(x_{k+1})\Big) - \frac{h^3}{12}f''(x_k) + O\left(h^4\right)
\end{aligned}$$

When we sum both sides of this equation from $k = 0$ to $n-1$, we get

$$\int_a^b f = T_n - \frac{(b-a)^2}{n^2}\left(\frac{b-a}{n}\right)\sum_{k=0}^{n-1} f''(x_k) + O\left(\frac{1}{n^3}\right)$$

as $n \to \infty$. The term containing the sum is simply the nth left-endpoint approximation to the integral of f'', so

$$\lim_{n\to\infty} n^2\left(\int_a^b f - T_n\right) = \frac{(b-a)^2}{12}(f'(a) - f'(b))$$

which is equivalent to what we wanted to show. $\square$

2.4 Simpson's Rule

So far, we have approximated f with constant and linear functions, that is, degree 0 and degree 1 polynomials. We now will approximate f using a quadratic, that is, a degree two polynomial. A generic quadratic polynomial $Q(x)$ has three degrees of freedom, so we may choose Q so that it intersect the curve $y = f(x)$ at three points. We will choose these three points to be the left- and right-endpoints and the midpoint. If $Q(x)$ is a quadratic function, then we have

$$\int_c^d Q(x)\,dx = \frac{c-d}{3}\left[Q(c) + 4Q\left(\frac{c+d}{2}\right) + Q(d)\right]$$

for any interval $[c,d]$. This allows us to define the nth Simpson's rule estimate of the integral of f.

Definition 2.1. Let f be Riemann integrable on $[a,b]$ and let n be an even integer. For each i with $0 \le i \le n$, let $x_i = a + i(b-a)/n$. Then the nth Simpson's rule approximation to f is

$$\begin{aligned} S_n = \frac{b-a}{n}&(f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4)\cdots \\ &+ 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)) \end{aligned}$$

In the same sense that the trapezoid rule was the average of the left- and right-endpoint rules, Simpsons rule is a weighted average of the left-, right- and midpoint rules. Specifically,

$$S_n = \frac{L_n + 4M_n + R_n}{6}$$

The following lemmas and theorem shows that Simpson's rule is the fastest converging of all of the approximations we have considered so far.

Lemma 2.1. If g has fourth derivative on $[-r,r]$ for some positive number r, then there exists a z in $(-r,r)$ such that

$$\int_{-r}^{r} g - \frac{r}{3}(g(r) + 4g(0) + g(-r)) = -\frac{r^5}{90}g''''(z)$$

Proof. Let k be the number

$$k = \frac{1}{r^5}\left[\int_{-r}^{r} g - \frac{r}{3}(g(r) + 4g(0) + g(-r))\right]$$

We want to show that there exists a $z \in (-r,r)$ such that

$$k = -\frac{g''''(z)}{90}$$

Define a function G on $[0,r]$ by

$$G(x) = \int_{-x}^{x} g - \frac{x}{3}(g(x) + 4g(0) + g(-x)) - kx^5$$

Then we may compute the first three derivatives of G to be

$$\begin{aligned} G'(x) &= g(x) + g(-x) - \frac{1}{3}(g(x) + 4g(0) + g(-x)) - \frac{x}{3}(g'(x) - g'(-x)) - 5kx^4 \\ &= \frac{2}{3}(g(x) + g(-x)) - \frac{4}{3}g(0) - \frac{x}{3}(g'(x) - g'(-x)) - 5kx^4 \\ G''(x) &= \frac{2}{3}(g'(x) - g'(-x)) - \frac{1}{3}(g'(x) - g'(x)) - \frac{x}{3}(g''(x) + g''(-x)) - 20kx^3 \\ &= \frac{1}{3}(g'(x) - g'(-x)) - \frac{x}{3}(g''(x) + g''(-x)) - 20kx^3 \\ G'''(x) &= \frac{1}{3}(g''(x) - g''(-x)) - \frac{1}{3}(g''(x) - g''(-x)) - \frac{x}{3}(g'''(x) - g'''(-x)) - 60kx^2 \\ &= -\frac{x}{3}(g'''(x) - g'''(-x) + 180kx) \end{aligned}$$

It is easy to see that $G(0) = G'(0) = G''(0) = 0$ and also, by our choice of k, $G'(r) = 0$. Because $G(0) = G(r)$, by Rolle's theorem there exists a number $a \in (0, r)$ such that $G'(a) = 0$. But then since $G'(0)$ is also equal to 0, there is a number $b \in (0, a)$ such that $G''(b) = 0$, again by Rolle's theorem. Iterating this one more time gives us $c \in (0, b)$ such that $G'''(c) = 0$. Substituting this into our formula for G''' gives us

$$-\frac{c}{3}(g'''(c) - g'''(-c) + 180kc) = 0$$

By solving for k and applying the MVT, we find that there exists a $z \in (-c, c) \subset (-r, r)$ with

$$k = -\frac{g'''(c) - g'''(-c)}{90} = -\frac{1}{90}\frac{g'''(c) - g'''(-c)}{2c} = -\frac{1}{90}g''''(z)$$

This completes the proof.

□

Lemma 2.2. If f has a fourth derivative on an interval $[c, d]$, then there exists a point v in the interval (c, d) such that

$$\int_c^d f - \frac{d-c}{6}(f(c) + 4f(m) + f(d)) = -\frac{(d-c)^5}{32 \cdot 90} f^{(4)}(v)$$

where $m = (d + c)/2$ is the midpoint of the interval $[c, d]$.

Proof. Let $r = (d-c)/2$ and let $g : [-r, r] \to \mathbb{R}$ be given by $g(x) = f(m+x)$. Then g satisfies the hypotheses of Lemma 2 and therefore there exists a number $z \in (-r, r)$ such that

$$\int_{-r}^r - \frac{r}{3}(g(-r) + 4g(0) + g(r)) = -\frac{r^5}{90}g^{(4)}(z)$$

Observe that

$$\int_{-r}^r g = \int_c^d f, \quad g(-r) = f(c), \quad g(r) = f(d), \quad r^5 = \frac{(d-c)^5}{32}$$

which gives us

$$\int_c^d f - \frac{d-c}{6}(f(c) + 4f(m) + f(d)) = -\frac{(d-c)^5}{32 \cdot 90} f^{(4)}(v)$$

where $v = m + z$. □

Theorem 2.6. If f has a fourth derivative on $[a, b]$ and n is an even positive integer then there exists a point $v \in (a, b)$ such that

$$\int_a^b f - S_n = -\frac{(b-a)^5}{180n^4} f^{(4)}(v)$$

Proof. Let $x_i = a + i(b-a)/n$ for $0 \le i \le n$. Observe that we may write

$$S_n = \frac{b-a}{3n}\sum_{i=1}^{n/2}(f(x_{2i-2}) + 4f(x_{2i-1}) + f(x_{2i}))$$

Then we have

$$\int_a^b f - S_n = \sum_{i=1}^{n/1}\left(\int_{x_{2i-2}}^{x_{2i}} f - \frac{2(b-a)/n}{6}(f(x_{2i-2}) + 4f(x_{2i-1}) + f(x_{2i}))\right)$$

Applying Lemma 3 for each integer i with $1 \le i \le n/2$, we get a number $v_i \in (x_{2i-2}, x_{2i})$ such that

$$\int_{x_{2i-2}}^{x_{2i}} f - \frac{2(b-a)/n}{6}(f(x_{2i-2}) + 4f(x_{2i-1}) + f(x_{2i})) = -\frac{(d-c)^5}{90n^5} f^{(4)}(v_i)$$

Then

$$\int_a^b f - S_n = -\frac{(d-c)^5}{90n^5}\sum_{i=1}^{n/2} f^{(4)}(v_i) = -\frac{(d-c)^5}{180n^4}\,\frac{2}{n}\sum_{i=1}^{n/2} f^{(4)}(v_i)$$

Recall that derivatives satisfy the intermediate value property. If we interpret

$$\frac{f^{(4)}(v_1) + f^{(4)}(v_2) + \cdots + f^{(4)}(v_{n/2})}{n/2}$$

as an average, then it must lie between $\min\{f^{(4)}(v_i) : 1 \le i \le n/2\}$ and $\max\{f^{(4)}(v_i) : 1 \le i \le n/2\}$. Since $f^{(4)}$ has the IVP on $[v_1, v_{n/2}]$, there is a $v \in (v_1, v_{n/1})$ such that

$$\int_a^b f - S_n = -\frac{(b-a)^5}{180n^4} f^{(4)}(v)$$

which is what we wanted. □

3 Extensions

3.1 Uneven tags

We will begin by extending the right- and left-endpoint rules. These rules choose fixed tags in a uniform partition of the interval $[a,b]$. The right endpoint rule chooses the point x_k in the interval $[x_{k-1}, x_k]$ as the tag, while the left-endpoint rule chooses the point x_{k-1}. What if instead we choose a point a fixed distance from one of the endpoints in each subinterval? The next theorem answers this question. For the remainder of this section, we will consider only functions on the interval $[0,1]$. Results for generic intervals $[a,b]$ exist and are easy to deduce from the following theorems by considering the transformation $x \mapsto (x-a)/(b-a)$ from $[a,b]$ to $[0,1]$.

Theorem 3.1. Suppose the function $f : [0,1] \to \mathbb{R}$ is differentiable and its derivative is integrable. For any $\alpha \in [0,1]$, form the Riemann sum

$$U_n = \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k-\alpha}{n}\right)$$

Then the error term $\Delta_n = \int_0^1 f - U_n$ satisfies

$$\Delta_n \sim (\alpha - 1/2)\frac{f(1) - f(0)}{n}$$

Proof. The result follows by observing that

$$\begin{aligned}
\Delta_n &= \int_0^1 f - \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k-\alpha}{n}\right) \\
&= \left(\int_0^1 f - \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right)\right) + \frac{1}{n}\sum_{k=1}^{n}\left(f\left(\frac{k}{n}\right) - f\left(\frac{k-\alpha}{n}\right)\right)
\end{aligned}$$

By the mean value theorem, for each k with $1 \le k \le n$, there exists a value $c_k \in [(k-1)/n, k/n]$ such that

$$f\left(\frac{k}{n}\right) - f\left(\frac{k-\alpha}{n}\right) = \frac{\alpha}{n} f'(c_k).$$

Then we have

$$\begin{aligned}
\Delta_n &= \left(\int_0^1 f - \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right)\right) + \frac{\alpha}{n}\frac{1}{n}\sum_{k=1}^{n} f'(c_k) \\
&\sim \frac{f(0) - f(1)}{2n} + \frac{\alpha}{n}\int_0^1 f',
\end{aligned}$$

here we have used Theorem 2.2 and the fact that

$$\frac{1}{n}\sum_{k=1}^{n} f'(c_k)$$

is a Riemann sum of f' over $[0,1]$. Applying the fundamental theorem of calculus gives us

$$\Delta_n \sim \frac{1}{2n}(f(0)-f(1)) + \frac{\alpha}{n}(f(1)-f(0)) = \frac{(\alpha-1/2)(f(1)-f(0))}{n}$$

□

We can observe that when we choose $\alpha = 0$ and $\alpha = 1$, we get the right- and left-endpoint approximations to the integral and the theorem gives us the expected asymptotic behavior from Thereoms 2.1 and 2.2. One might think that we could have just proven this theorem instead of going to the trouble of proving Theorems 2.1 and 2.2, but remember that we use Theorem 2.2 in the proof of this theorem.

It is also worth observing that if we left $\alpha = 1/2$, we get the midpoint approximation of the integral of f. Theorem 3.1 tells us, correctly, that $\int_0^1 f - M_n = o(1/n)$, that is, the order of convergence of the error term is strictly smaller than $1/n$. The midpoint rule (Theorem 2.4) clarifies that it is in fact asymptotic with C/n^2, for a constant C.

We now further generalize by asking the question: What happens if we form a Riemann sum approximating the sum or product of two functions, say $f+g$ or fg, but we evaluate f and g at different tags? More precisely, suppose that for every integer k with $1 \le k \le n$, we pick tags $x_n^k, y_n^k \in [(k-1)/n, k/n]$ and form the sums

$$S_n = \frac{1}{n}\sum_{k=1}^{n}(f(x_n^k) + g(y_n^k)) \text{ and } P_n = \frac{1}{n}\sum_{k=1}^{n} f(x_n^k)g(y_n^k)$$

It is clear by linearity that $S_n \to \int_0^1 f$ as $n \to \infty$. The following theorem shows that P_n also converges to what we would expect. This result is commonly known as Bliss' theorem.

Theorem 3.2. Let $f, g : [a,b] \to \mathbb{R}$ be Riemann integrable on $[a,b]$ and suppose that $x_n^k, y_n^k \in [(k-1)/n, k/n]$ for $1 \le k \le n$. Let

$$P_n = \frac{1}{n}\sum_{k=1}^{n} f(x_n^k)g(y_n^k).$$

Then we have $\lim_{n\to\infty} S_n = \int_0^1 f$.

Proof. Fix any $\epsilon > 0$. Since f and g are integrable, their product fg is also integrable. Then, there exists an integer N so that

$$\left|\frac{1}{n}\sum_{k=1}^{n} f(y_n^k)g(y_n^k) - \int_0^1 f(x)g(x)\,dx\right| < \epsilon$$

whenever $n > N$. Further, by the Cauchy criterion, we may choose a possibly larger N so that

$$\left|\frac{1}{n}\sum_{k=1}^{n}\left(g(x_n^k) - g(y_n^k)\right)\right| < \epsilon$$

as well whenever $n > N$. Since f is Riemann integrable, it is bounded. Let $M = \sup_{x\in[a,b]} |f|$. Then whenever $n > N$, we have

$$\begin{aligned}
\left| P_n - \int_0^1 f(x)\,dx \right| &= \left| \frac{1}{n}\sum_{k=1}^n f(x_n^k)g(y_n^k) - \int_0^1 f(x)\,dx \right| \\
&\le \left| \frac{1}{n}\sum_{k=1}^n f(x_n^k)g(y_n^k) - \frac{1}{n}\sum_{k=1}^n f(x_n^k)g(x_n^k) \right| \\
&\quad + \left| \frac{1}{n}\sum_{k=1}^n f(x_n^k)g(x_n^k) - \int_0^1 f(x)\,dx \right| \\
&\le \left| \frac{1}{n}\sum_{k=1}^n f(x_n^k)\Big[g(x_n^k) - g(y_n^k)\Big] \right| + \epsilon \\
&\le M\frac{1}{n}\sum_{k=1}^n \left|g(x_n^k) - g(y_n^k)\right| + \epsilon \\
&< (1+M)\epsilon
\end{aligned}$$

This shows that $|P_n - \int fg| \to 0$ as $n \to \infty$, which is equivalent to what we wanted to show. $\square$

What can we say about the asymptotic behavior of the error of each of these approximations? If the distance from each endpoint is fixed, as it was in Theorem 3.1, we are able to quantify this behavior. For the sum, the answer is provided to us easily by Theorem 3.1. Suppose that α and β are fixed numbers in the interval $[0,1]$. Then

$$\begin{aligned}
\int_0^1 (f+g) - \frac{1}{n}\sum_{k=1}^n \left(f\left(\frac{k-\alpha}{n}\right) + g\left(\frac{k-\beta}{n}\right)\right) &= \left(\int_0^1 f - \frac{1}{n}\sum_{k=1}^n f\left(\frac{k-\alpha}{n}\right)\right) + \left(\int_0^1 g - \frac{1}{n}\sum_{k=1}^n g\left(\frac{k-\beta}{n}\right)\right) \\
&\sim \frac{1}{n}\left((\alpha - 1/2)\left(f(1) - f(0)\right) + (\beta - 1/2)(g(1) - g(0))\right)
\end{aligned}$$

The case for products is slightly more complicated and was considered in the paper [1]. We reproduce their result for this situation, with a slightly modified proof.

Theorem 3.3. Let $f, g : [0,1] \to \mathbb{R}$ be two differentiable functions whose derivatives are Riemann integrable. Let $\alpha, \beta \in [0,1]$ and let

$$\Delta_n = \int_0^1 f(x)g(x)\,dx - \frac{1}{n}\sum_{k=1}^n f\left(\frac{k-\alpha}{n}\right) g\left(\frac{k-\beta}{n}\right)$$

for every positive integer n. Then we have

$$\lim_{n\to\infty} n\Delta_n = \frac{f(0)g(0) - f(1)g(1)}{2} + \alpha\int_0^1 f'(x)g(x)\,dx + \beta\int_0^1 f(x)g'(x)\,dx$$

Proof. Let

$$D_n = \int_0^1 f(x)g(x)\,dx - \frac{1}{n}\sum_{k=1}^n f\left(\frac{k}{n}\right) g\left(\frac{k}{n}\right).$$

Since fg is continuous and therefore integrable, we have

$$\lim_{n\to\infty} nD_n = \frac{f(0)g(0) - f(1)g(1)}{2}$$

by Theorem 2.1. Because f is differentiable, the mean value theorem implies that for all integers n and k with $1 \le k \le n$, there are numbers c_{nk}, d_{nk} with

$$\frac{k-\alpha}{n} < c_{nk} < \frac{k}{n}, \quad \frac{k-\beta}{n} < d_{nk} < \frac{k}{n}$$

and

$$f\left(\frac{k}{n}\right)-f\left(\frac{k-\alpha}{n}\right)=f'(c_{nk})\frac{\alpha}{n},\quad g\left(\frac{k}{n}\right)-g\left(\frac{k-\beta}{n}\right)=g'(d_{nk})\frac{\beta}{n}$$

Then we have

$$\begin{aligned}
n\Delta_n &= n\int_0^1 f(x)g(x)\,dx-\sum_{k=1}^n f\left(\frac{k-\alpha}{n}\right)g\left(\frac{k-\beta}{n}\right)\\
&= nD_n+\sum_{k=1}^n\left[f\left(\frac{k}{n}\right)g\left(\frac{k}{n}\right)-f\left(\frac{k-\alpha}{n}\right)g\left(\frac{k-\beta}{n}\right)\right]\\
&= nD_n+\sum_{k=1}^n\left[f\left(\frac{k}{n}\right)-f\left(\frac{k-\alpha}{n}\right)\right]g\left(\frac{k}{n}\right)\\
&\quad+\sum_{k=1}^n f\left(\frac{k-\alpha}{n}\right)\left[g\left(\frac{k}{n}\right)-g\left(\frac{k-\beta}{n}\right)\right]\\
&= nD_n+\frac{\alpha}{n}\sum_{k=1}^n f'\left(c_{nk}\right)g\left(\frac{k}{n}\right)+\sum_{k=1}^n f\left(\frac{k-\alpha}{n}\right)g'\left(d_{nk}\right)
\end{aligned}$$

which by Theorem 3.2 converges to

$$\frac{f(0)g(0)-f(1)g(1)}{2}+\alpha\int_0^1 f'(x)g(x)\,dx+\beta\int_0^1 f(x)g'(x)\,dx$$

□

3.2 Improper Integrals

We now turn our attention to improper integrals. Recall that a function $f:(a,b]\to\mathbb{R}$ has an **improper integral** if $\int_c^b f$ exists for all $c\in(a,b)$ and $\lim_{c\to a^+}\int_c^b f$ exists and is finite. A similar definition exists for intervals of the form $[a,b)$ and (a,b). For example, if $0<r<1$, the function $f(x)=1/x^r$ on $(0,1]$ is improperly integrable as we can see by

$$\lim_{\delta\to 0^+}\int_\delta^1 f=\lim_{\delta\to 0^+}\frac{1}{1-r}(1-\delta^{1-r})=\frac{1}{1-r}$$

We now ask, our the approximations that we have been using for the integral of a function still valid for functions that are improperly integrable, but not Riemann integrable. For example, if $f:(0,1]\to\mathbb{R}$ is improperly integrable, we can form the sum

$$\tilde{R}_n=\frac{1}{n}\sum_{k=1}^n f\left(\frac{k}{n}\right)$$

We place the tilde over the R to remind us that this is a different type of quantity than the right-endpoint approximation from Section 1.1, since we do not know that f is Riemann integrable. The next theorem shows that, under some conditions, this function does, in fact, converge to the value of the integral.

Theorem 3.4. Let $f:(0,1]\to\mathbb{R}$ be non-increasing and improperly integrable and let $\tilde{R}_n$ be as above. Then $\tilde{R}_n\to\int_0^1 f$ as $n\to\infty$.

Proof. Fix any $\epsilon>0$. Since $\int_c^1 f\to\int_0^1 f$ as $c\to 0^+$, we can find a number $a>0$ such that $\left|\int_0^a f\right|=\left|\int_0^1 f-\int_a^1 f\right|<\epsilon$. For every positive integer n, let j_n be the unique integer such that $j_n/n\le a<(j_n+1)/n$. Then $\tilde{R}_n$ is a right-endpoint approximation of $\int_{j/n}^1 f$, so we may find a value N such that $\left|\int_{j/n}^1 f-\tilde{R}_n\right|<\epsilon$ whenever $n>N$. Observe that since f is decreasing, we have

$$\frac{1}{n}\sum_{k=1}^j f(k/n)<\int_0^{j/n} f.$$

Then whenever $n > N$,

$$
\begin{aligned}
\left|\int_0^1 f - \tilde{R}_n\right| &\le \left|\int_0^a f\right| + \left|\int_a^1 f - \tilde{R}_n\right| \\
&\le \epsilon + \left|\frac{1}{n}\sum_{k=1}^{j_n-1} f(k/n)\right| + \left|\int_{j_n/n}^1 f - \frac{1}{n}\sum_{k=j_n}^{n} f(k/n)\right| + \left|\int_{j_n}^a f\right| \\
&\le 4\epsilon.
\end{aligned}
$$

Then it follows that $\tilde{R_n} \to \int_0^1 f$ as $n \to \infty$.

$\square$

The ideas used in the above proof can easily be extended to more general situations, that is for functions on general intervals (a, b) and for the left endpoint method as well.

What can be said about the order of convergence of this approximation? The author is not aware of a general answer to this question and it appears to be more complicated that the case of Riemann integrable functions. We provide an example to illustrate the situation

Theorem 3.5. Suppose that $0 < r < 1$ and let $g : (0, 1] \to \mathbb{R}$ be given by $g(x) = 1/x^r$. Then there exists a non-negative constant C such that

$$\int_0^1 g - \tilde{R}_n \sim \frac{C}{n^{1-r}}$$

where $\tilde{R}_n$ is the same as in the previous theorem.

To prove this theorem, we need the following lemma

Lemma 3.1. Suppose that $f : [1, \infty) \to \mathbb{R}$ is a non-negative differentiable function whose derivative is negative and increasing on $(1, \infty)$. For each positive integer k, let A_k be the absolute value of the error of a single subdivision trapezoid approximation of the integral of f on $[k, k + 1]$. Then the series $\sum_{k=1}^{\infty} A_k$ converges.

Proof. First, for each positive integer k, define the functions

$$L_k(x) = (f(k) - f(k-1))(x - k) + f(k), \quad U_k(x) = (f(k + 1) - f(k))(x - k) + f(k)$$

Observe that f' is increasing therefore f is convex, which implies that $L_k(x) \le f(x) \le U_k(x)$ for all $x \in [k, k + 1]$. Also note that, since $U_k(x)$ is a secant line connecting $(k, f(k))$ and $k + 1, f(k + 1)$, so

$$A_k = \int_k^{k+1} (U_k - f)$$

If we let $D_k = f(k + 1) - f(k)$ for $k \ge 2$, a simple computation shows that

$$\int_k^{k+1} (U_k - L_k) = \frac{1}{2}(D_{k-1} - D_k))$$

from which we get

$$0 \le A_k = \int_k^{k+1} (U_k - f) \le \int_k^{k+1} (U_k - L_k) = \frac{1}{2}(D_{k-1} - D_k)$$

Observe that since f' is negative and increasing, it converges to a finite limit as $x \to \infty$. Therefore, by the mean value theorem, $D_k = f(k + 1) - f(k)$ converges to a finite limit as $k \to \infty$. Then telescoping series $\sum_{k=2}^{\infty} D_k$ converges and the smaller series $\sum_{k=1}^{\infty} A_k$ converges as well. $\square$

We now have the tools to prove Theorem 3.4.

Proof. We want to show that the sequence

$$n^{1-r}\left(\int_0^1 \frac{dx}{x^r} - \tilde{R}_n\right) = \frac{n^{1-r}}{1-r} - \sum_{k=1}^{n} \frac{1}{k^r}$$

converges. Let $f(x) = 1/x^r$ for $1 \le x < \infty$. Then by applying Lemma 3.2 to f, we that get

$$\begin{aligned}\sum_{k=1}^{n} A_k &= \int_1^n \frac{dx}{x^r} - \sum_{k=1}^{n} f(k) + \frac{1}{2}\left(f(1) + f(n)\right) \\ &= \frac{n^{1-r}}{1-r} - \sum_{k=1}^{n} \frac{1}{x^r} + O(1)\end{aligned}$$

converges. Therefore the sequence we are interested in converges as well and the result is proven. $\square$

3.3 Higher Order Error Expansion and the Euler-Maclaurin Formula

So far, the results we have about the order of convergence of an approximation to the Riemann integral have looked like

$$\Delta_n \sim \frac{C}{n^p}$$

for some constant C and and some integer p. Another way of writing this would be

$$\Delta_n = \frac{C}{n^p} + o\left(\frac{1}{n^p}\right).$$

We might hope to add some more terms to this asymptotic expansion. A general asymptotic expansion of the error term would look like

$$\Delta_n = \frac{C_1}{n} + \frac{C_2}{n^2} + \cdots + \frac{C_k}{n^k} + o\left(\frac{1}{n^k}\right)$$

for some fixed positive integer k. To accomplish this goal, we must introduce a power tool that is used to connect sums and integrals called the Euler-Maclaurin summation formula. We will draw this dicussion from [4] and adopt that paper's notation.

We begin by defining the Bernoulli numbers and polynomials, which have many useful properties. We define the nth **Bernoulli polynomial** recursively, for each integer $n \ge 0$ by

$$B_0(x) = 1, \quad B_n'(x) = nB_{n-1}(x), \text{ and } \int_0^1 B_n(t)\,dt = 0$$

for all values of x. Note that the polynomials are well-defined, since we have

$$B_n(x) = n\int_0^x B_{n-1}(t)\,dt + B_n$$

for some choice of B_n. There is a unique choice of B_n that makes the above equation hold. We call B_n ($n \ge 0$) the nth **Bernoulli number**. Obeserve that $B_n = B_n(0)$ for $n \ge 1$. For n larger than 2, we have additionally

$$\begin{aligned}B_n(1) &= B_n(0) + (B_n(1) - B_n(0)) \\ &= B_n(0) + \int_0^1 B_n' = B_n(0) = B_n.\end{aligned}$$

The first several Bernoulli polynomials and numbers are

$$\begin{array}{ll}
B_0(x) = 1 & B_0 = 1 \\
B_1(x) = x - \frac{1}{2} & B_1 = -\frac{1}{2} \\
B_2(x) = x^2 - x + \frac{1}{6} & B_2 = \frac{1}{6} \\
B_3(x) = x^3 - \frac{3}{2}x^2 + \frac{1}{2}x & B_3 = 0 \\
B_4(x) = x^4 - 2x^3 + x^2 - \frac{1}{30} & B_4 = \frac{1}{30} \\
B_5(x) = x^5 - \frac{5}{2}x^4 + \frac{5}{3}x^3 - \frac{1}{6}x & B_5 = 0 \\
B_6(x) = x^6 - 3x^5 + \frac{5}{2}x^4 - \frac{1}{2}x^2 + \frac{1}{42} & B_6 = \frac{1}{42}
\end{array}$$

For every positive integer, we may define the nth **periodic Bernoulli polynomial** by

$$P_n(x) = B_n(x - \lfloor x \rfloor)$$

for all real numbers $\mathbb{R}$. Since $B_n(0) = B_n(1)$ for $n \geq 2$, P_n is continuous for $n \geq 2$. Using these new tools, we now state and prove the Euler-Maclaurin summation formula.

Theorem 3.6. Suppose $f : [0, n] \to \mathbb{R}$ is $m + 1$ times continuously differentiable. Then we have

$$\begin{aligned}\int_0^n f = \sum_{k=0}^{n} f(k) - \frac{1}{2}(f(n) + f(0)) + \sum_{j=1}^{m} (-1)^j \frac{B_j}{j!} \left(f^{(j-1)}(n) - f^{(j-1)}(0) \right) \\ + (-1)^{m+1} \frac{1}{(m+1)!} \int_0^n f^{(m+1)}(x) P_m(x)\, dx\end{aligned}$$

where B_k is the kth Bernoulli number and P_k is the kth periodic Bernoulli polynomial.

Proof. We will procede by induction on m. We begin with the base case $m = 1$. Fix a number k with $1 \leq k \leq n$. Then recalling that $P_0(x) = 1$ for all x, we have $\int_{k-1}^k f = \int_{k-1}^k f(x) P_0(x)\, dx$. By performing integration by parts with

$$\begin{aligned} u &= f(x) & dv &= P_0(x)\, dx \\ du &= f'(x)\, dx & v &= P_1(x) = x - \tfrac{1}{2} \end{aligned}$$

we get

$$\begin{aligned}\int_{k-1}^k f &= \Big[f(x) P_1(x) \Big]_{k-1}^k - \int_{k-1}^k f'(x) P_1(x)\, dx \\ &= \frac{1}{2}(f(k-1) + f(k)) - \int_{k-1}^k f'(x) P_1(x)\, dx\end{aligned}$$

We now perform integration by parts again on $\int_{k-1}^k f'(x) P_1(x)\, dx$ with

$$\begin{aligned} u &= f'(x) & dv &= P_1(x)\, dx \\ du &= f''(x)\, dx & v &= P_2(x)/2 \end{aligned}$$

to get

$$\int_{k-1}^k f'(x) P_1(x)\, dx = \frac{B_2}{2} \left(f'(k) - f'(k-1) \right) - \frac{1}{2} \int_{k-1} f''(x) P_2(x)\, dx.$$

Then we have

$$\int_{k-1}^k f = \frac{1}{2}(f(k-1) + f(k)) - \frac{B_2}{2} \left(f'(k) - f'(k-1) \right) + \frac{1}{2} \int_{k-1}^k f''(x) P_2(x)\, dx$$

which we may sum from $k = 1$ to n to get the $m = 2$ base case. Now fix some positive integer m and suppose that the mth case of our proposition holds. Consider the remainder term

$$\frac{1}{m!} \int_0^n f^{(m)}(x) P_m(x)\, dx$$

We again consider only the part of the integral on $[k - 1, k]$ and perform integration by parts with

$$
\begin{aligned}
u &= f^{(m)}(x) & dv &= P_m(x)\,dx \\
du &= f^{(m+1)}(x)\,dx & v &= P_{m+1}(x)/(m+1)
\end{aligned}
$$

to get

$$
\begin{aligned}
\frac{1}{m!}\int_0^n f^{(m)}(x)P_m(x)\,dx = &\frac{B_{m+1}}{(m+1)}(f^{(m)}(k) - f^{(m)}(k-1)) \\
&- \frac{1}{(m+1)!}\int_{k-1}^k f^{(m+1)}(x)P_{m+1}(x)\,dx
\end{aligned}
$$

Summing this from $k = 1$ to n and substituting it for the error term in the mth case yields the $(m+1)$th case, completing the proof. □

This formula allows us to extend our asymptotic expansion of the error term of our integral approximation as far out as we have derivatives. For example, suppose that $f : [0, n] \to \mathbb{R}$ is seven times differentiable. The Euler Maclaurin formula gives us

$$
\begin{aligned}
\int_0^n f - \sum_{k=1}^n \frac{f(k-1)+f(k)}{2} &= \sum_{j=2}^7 (-1)^j \frac{B_j}{j!}(f^{j-1}(n) - f^{j-1}(0)) \\
&= \frac{1}{12}(f'(n) - f'(0)) - \frac{1}{120}(f'''(n) - f'''(0)) + \frac{1}{42\cdot 6!}(f^{(5)}(n) - f^{(5)}(0)) + \rho
\end{aligned}
$$

where

$$
|\rho| \le \frac{1}{7!}\int_0^n |f^{(7)}(x)||P_7(x)|\,dx \le \frac{1}{7!}\int_0^n |f^{(7)}(x)|\,dx
$$

where we have used the fact that $|P_7(x)| \le 1$ for all x. Suppose now that $g : [0, 1] \to \mathbb{R}$ is seven times differentiable. We want to expand the error term $\int_0^1 g - T_n$ using our new formula. By letting $f(x) = ng(nx)$ in the above formula, we have

$$
\int_0^1 f - T_n = \frac{1}{12n^2}(f'(1) - f'(0)) - \frac{1}{120n^4}(f'''(1) - f'''(0)) + \frac{1}{42\cdot 6!}(f^{(5)}(1) - f^{(5)}(0)) + O(1/n^7)
$$

which is a much more precise expansion than our original

$$
\int_0^1 f - T_n \sim \frac{1}{12n^2}(f'(1) - f'(0))
$$

Acknowledgments

I would like to thank Russ Gordon for his help and patience with the writing of this paper. I would also like to thank Albert Schueller for organizing this project and Jamie Edison and Evan Kleiner for providing edits and feedback throughout the semester

References

[1] Merinescu, D. & Monea, M. (2011) An extension of a result about the order of convergence *Bulletin of Mathematical Analysis and Applications* Vol. 3, Iss. 3, 25-34

[2] Polya, G. & Szego, G. (1972) Problems and theorems in analysis. Berlin: Springer.

[3] Gordon, R. (2002) Real analysis: a first course. Addison-Wesley Higher Mathematics

[4] Apostol, T. (1999) An Elementary View of Euler's Summation Formula *American Mathematical Monthly* Vol. 106, No. 5

[5] Rudin, W. (1986) Real and Complex Analysis. 3rd Ed. McGraw-Hill Science/Engineering/Math