Proving Taylor's Theorem from the Fundamental Theorem of Calculus by Fixed-point Iteration

Christopher Thron

Texas A&M University - Central Texas, Killeen TX
*Corresponding author: thron@tamuct.edu

Abstract

Taylor's theorem (and its variants) is widely used in several areas of mathematical analysis, including numerical analysis, functional analysis, and partial differential equations. This article explains how Taylor's theorem in its most general form can be proved simply as an immediate consequence of the Fundamental Theorem of Calculus (FTOC). The proof shows the deep connection between the Taylor expansion and fixed-point iteration, which is a foundational concept in numerical and functional analysis. One elegant variant of the proof also demonstrates the use of combinatorics and symmetry in proofs in mathematical analysis. Since the proof emphasizes concepts and techniques that are widely used in current science and industry, it can be a valuable addition to the undergraduate mathematics curriculum.

Keywords: Taylor's theorem, fundamental theorem of calculus, fixed point, iteration, linear operator, integration, differentiation, volume integral.
MSC2010:26A18

1   Introduction

Taylor's theorem is one of the most important results that are taught in basic calculus classes. Its importance is both theoretical and practical. Taylor's theorem is a foundational result in the field of numerical analysis: many error estimates for numerical solutions to algebraic or differential equations are based on the Taylor expansion of the solution. Taylor series is also fundamental to the theory and application of differential equations, in which series solutions play a large role. In fact, the Cauchy-Kowalesky theorem that establishes the existence and uniqueness of solutions to partial differential equations relies on Taylor expansion of the solution [1].

It has been noted that the proofs of Taylor's theorem given in many textbooks are not well motivated [2]. Most often, the theorem is derived either using Rolle's theorem [3], or as a consequence of the mean value theorem [4], or by repeated integration by parts [5]. A less common proof uses induction [6]. It would seem that such a key theorem should have a stronger motivation, and should not be derived as an incidental result from other theorems.

The purpose of this paper is to show that Taylor's theorem is in fact an immediate consequence of the fundamental theorem of calculus; and furthermore, the proof is a straightforward application

---

101

of fixed point iteration, which is one of the most basic techniques in numerical and functional
analysis. We thus both provide motivation for Taylor's theorem, and show its deep relationship
with other areas of mathematical analysis.

The paper is organized as follows. Section 2 gives necessary background for the mathematical
concepts and notation used in the proof. Section 3 presents the new proof of Taylor's theorem.
Finally, Section 4 discusses the possible instructional uses for this new proof.

2   Background

This section introduces notation and concepts necessary for understanding the proof of Taylor's
theorem in Section 3. These concepts are indispensible in modern applied mathematics.

2.1   Theorem statements

The fundamental theorem of calculus expresses the inverse relationship between differentiation and
integration. There are many alternative statements of the theorem that differ slightly from one
another. We will use the following formulation [7]:

Theorem 2.1. *(Fundamental theorem of calculus for the Lebesgue integral)   A function $f : [a, b] \to$
$\mathbb{R}$ is absolutely continuous if and only if it is differentiable almost everywhere, its derivative $f' \in$
$L^1[a, b]$ and , for each $x \in [a, b]$,*

$$f(x) = f(a) + \int_a^x f'(t)\, dt. \tag{2.1}$$

Several alternative proofs of this theorem are cited in [7].
Taylor's theorem may be stated as follows:

Theorem 2.2. *(Taylor's theorem)   Given a function $f : [a, b] \to \mathbb{R}$ such that the Nth derivative
$f^{(N)}$ is absolutely continuous on $[a, b]$, then for all $x \in [a, b]$*

$$f(x) = f(a) + (x - a)f'(a) + \frac{(x - a)^2}{2!}f''(a) + \dots$$
$$\dots + \frac{(x - a)^N}{N!}f^{(N)} + R_n(x), \tag{2.2}$$

where $R_n(x)$ satisfies the inequality

$$|R_n(x)| \leq \sup_{c \in [a,x]} |f^{(N+1)}(c)| \frac{(x - a)^{N+1}}{(N + 1)!}, \tag{2.3}$$

and $R_n(x)$ is given exactly by the expression

$$R_n(x) = \int_a^x \frac{(x - t)^N}{N!} f(N + 1)(t) dt. \tag{2.4}$$

2.2   Linear operators and some basic properties

The concept of a *linear operator* is not typically mentioned in basic calculus classes: however, there
are good reasons for changing this practice. Modern computational mathematics is fundamentally
based on the connection between linear algebra and analysis, and the idea of linear operator is a
key aspect of this connection. Linear operators play the same role in calculus that matrices play in
linear algebra, as we shall now explain in more detail.

Let V, W be finite-dimensional real vector spaces, and let $L : V \to W$ be a mapping (i.e. a function) with domain V and codomain W. The mapping L is *linear* if for all vectors $\mathbf{u}, \mathbf{v} \in V$ and constants $a, b \in \mathbb{R}$ we have

$$L(a\mathbf{u} + b\mathbf{v}) = aL(\mathbf{u}) + bL(\mathbf{v}) \tag{2.5}$$

It is a well-known result in linear algebra that every linear mapping L between finite-dimensional vector spaces can be associated with a unique matrix (which we denote here as M) such that $L(\mathbf{v}) = M\mathbf{v}$.

In order to draw the connection between matrices and linear operators, we observe that function spaces (such as $L^1[a, b]$ or the set of absolutely continuous functions) are infinite-dimensional vector spaces. Therefore given two function spaces A, B, we may consider the set of mappings from A to B. Such mappings are called *operators*. An operator $\mathcal{L} : A \to B$ is *linear* if for any functions $f, g \in A$ we have

$$\mathcal{L}(af + bg) = a\mathcal{L}(f) + b\mathcal{L}(g), \tag{2.6}$$

which obviously corresponds exactly with (2.5). Because of the close connection between matrices and linear operators, the action of a linear operator is often written as if it were a multiplication: for example, $\mathcal{L}(f)$ is written instead as $\mathcal{L}f$, so that (2.6) is written as

$$\mathcal{L}(af + bg) = a\mathcal{L}f + b\mathcal{L}g. \tag{2.7}$$

Note that $\mathcal{L}f$ is in itself a real-valued function, which is a mapping from $\mathbb{R}$ to $\mathbb{R}$: thus $\mathcal{L}f(x)$ denotes the value of the function $\mathcal{L}f$ evaluated at the point x.

Two important examples of operators are the *differentiation* (or *derivative*) *operator*

$$\mathcal{D}f(x) := \lim_{\delta \to 0} \frac{f(x + \delta) - f(x)}{\delta}. \tag{2.8}$$

where $\mathcal{D}$ is defined on the set of absolutely continuous functions; and the *integral operator*

$$\mathcal{I}_a g(x) := \int_a^x g(t)dt, \tag{2.9}$$

which is defined for all $g \in L^1[a, b]$.

Another operator that is defined for absolutely continuous functions is the *evaluation operator*:

$$f \to f(a)\mathbf{1}, \tag{2.10}$$

where a is any number in the domain of f, and $\mathbf{1}$ denotes the constant function that takes the value 1 for all values of x in the domain of f.

Operators can be composed to produce other operators. In the subsequent discussion, the *associativity* of operator composition will play an important role: if $\mathcal{J}, \mathcal{K}, \mathcal{L}$ are compatible operators, then for any function f in the domain of $\mathcal{L}$ we have

$$\mathcal{J}(\mathcal{K}\mathcal{L}f) = (\mathcal{J}\mathcal{K})\mathcal{L}f. \tag{2.11}$$

It follows that parentheses are unnecessary when writing operator compositions: for example, we may write $\mathcal{I}_a(\mathcal{I}_a(\mathcal{D}(\mathcal{D}f)))$ as $\mathcal{I}_a^2\mathcal{D}^2 f$.

A linear combination of linear operators is also a linear operator. In Section 3 we will make use of the following linear operator, which is defined as a linear combination:

$$\mathcal{L}f := f(a)\mathbf{1} + \mathcal{I}_a\mathcal{D}f, \tag{2.12}$$

$\mathcal{L}f$ is defined on the set of absolutely continuous real-valued functions defined on any interval containing a.

In Section 3 we will make use of the *monotonicity* property of the integral operator: if $g, h \in L^1[a, x]$, then

$$g \le h \implies \mathcal{I}_a g \le \mathcal{I}_a h, \tag{2.13}$$

where the inequalities in (2.13) hold pointwise throughout the interval $[a, x]$. Since

$$- \sup_{c \in [a,x]} |g(c)| \cdot \mathbf{1} \le g \le \sup_{c \in [a,x]} |g(c)| \cdot \mathbf{1} \tag{2.14}$$

it follows that from the linearity of $\mathcal{I}_a$ that

$$- \sup_{c \in [a,x]} |g(c)| \cdot \mathcal{I}_a \mathbf{1} \le \mathcal{I}_a g \le \sup_{c \in [a,x]} |g(c)| \cdot \mathcal{I}_a \mathbf{1}$$
$$\implies |\mathcal{I}_a g| \le \sup_{c \in [a,x]} |g(c)| \dot{\mathcal{I}}_a \mathbf{1}. \tag{2.15}$$

By iterating (2.13) n times and applying $\mathcal{I}_a^n$ to (2.14), we may generalize (2.15) to

$$|\mathcal{I}_a^n g| \le \sup_{c \in [a,x]} |g(c)| \dot{\mathcal{I}}_a^{\,n} \mathbf{1}. \tag{2.16}$$

2.3 Fixed points and fixed-point iteration

A *fixed point* is any mathematical object x that satisfies an equation of the form $x = g(x)$, where g is a function. The concept of fixed point play a hugely important role in mathematical analysis, both theoretically and computationally. The online resource Mathworld lists nine prominent fixed-point theorems that appear in diverse areas of mathematics, including the famous Brouwer and Banach fixed point theorems for topological and metric spaces respectively [8].

Fixed point ideas also figure in other important theoretical results. The central limit theorem in probability is connected to the fact that the standard normal density function is a fixed point of the modified convolution operator $f(x) \to f(\sqrt{2}x) * f(\sqrt{2}x)$ [9]. Given a time-dependent differential equation $\frac{df}{dt} = A(f, x)$, then a steady-state solution satisfies $A(f, x) = 0$ which can also be written as a fixed point equation: $f = f + A(f, x)$. Furthermore, periodic solutions correspond to fixed points of the Poincare map [10]. In general, any algebraic equation can be written as a fixed point equation: for example, we have for any real-valued function f

$$f(x) = 0 \iff x = x + f(x). \tag{2.17}$$

On the practical side, one of the most versatile and widely-applied techniques in numerical analysis is *fixed point iteration*. One example is Newton's method for finding roots of functions. This method is based on the fact that any root x^* of a differentiable function $f : \mathbb{R} \to \mathbb{R}$, is a locally stable fixed point of the function $g(x) := x - \frac{f(x)}{f'(x)}$ as long as $f'(x^*) \ne 0$. As a result, x^* can be estimated numerically by iteration: that is, given an initial guess x_0 that is sufficiently close to x^* then the series $g(x_0), g(g(x_0)), g(g(g(x_0))), \ldots g^{(n)}(x_0) \ldots$ converges to x^*. Another example is the power method for finding the largest eigenvector-eigenvalue pair of a matrix (or more generally, a linear operator). Given a matrix M, the unit eigenvector $\mathbf{v}$ corresponding to the eigenvalue with largest absolute value is a solution to the fixed-point equation $\mathbf{v} = F(\mathbf{v})$ where $F(\mathbf{v}) := \frac{M\mathbf{v}}{|M\mathbf{v}|}$. It follows that $\mathbf{v}$ can be estimated numerically by repeatedly iterating the function F: $F^{(n)}(\mathbf{v}) \xrightarrow[n \to \infty]{} \mathbf{v}^*$ [11]. In physics, perturbation series are often obtained through fixed-point iteration. A prominent example is the *Born series* in electromagnetic and quantum scattering [12].

3 Method

3.1 Basic proof of Taylor's theorem

It is not commonly recognized in calculus textbooks that the fundamental theorem of calculus is simply a fixed-point equation:

$$f = \mathcal{L}f = f(a)\mathbf{1} + \mathcal{I}_a \mathcal{D}f, \tag{3.1}$$

where $\mathcal{L}$ is defined in (2.12). According to Theorem 2.1, (3.1) holds almost everywhere for all absolutely continuous functions $f : [a, b] \to \mathbb{R}$.

To obtain Taylor's theorem, we apply a slight modification of the technique of fixed-point iteration described in Section 2.3. If the derivative $\mathcal{D}f$ is absolutely continuous, then by (3.1) we have $\mathcal{D}f = \mathcal{L}\mathcal{D}f = \mathcal{D}f(a)\mathbf{1} + \mathcal{I}_a\mathcal{D}(\mathcal{D}f)$. Making this replacement in (3.1) gives

$$
\begin{aligned}
f &= f(a)\mathbf{1} + \mathcal{I}_a\mathcal{D}f \\
&= f(a)\mathbf{1} + \mathcal{I}_a(\mathcal{D}f(a)\mathbf{1} + \mathcal{I}_a\mathcal{D}(\mathcal{D}f)) \\
&= f(a)\mathbf{1} + \mathcal{D}f(a)\mathcal{I}_a\mathbf{1} + \mathcal{I}_a(\mathcal{I}_a\mathcal{D}(\mathcal{D}f)) \\
&= f(a)\mathbf{1} + \mathcal{D}f(a)\mathcal{I}_a\mathbf{1} + \mathcal{I}_a^2\mathcal{D}^2 f.
\end{aligned}
\tag{3.2}
$$

The last two equalities in (3.2) follow from the linearity and associativity properties of operators described in Section 2.2.

We can now continue the iterative process. If $\mathcal{D}^2 f$ is absolutely continuous, then by (3.1) we may similarly replace $\mathcal{D}^2 f$ in (3.2) with $\mathcal{D}^2 f(a)\mathbf{1} + \mathcal{I}_a\mathcal{D}(\mathcal{D}^2 f)$. This gives

$$
\begin{aligned}
f &= f(a)\mathbf{1} + \mathcal{D}f(a)\mathcal{I}_a\mathbf{1} + \mathcal{I}_a^2(\mathcal{D}^2 f(a) + \mathcal{I}_a\mathcal{D}(\mathcal{D}^2 f)) \\
&= f(a)\mathbf{1} + \mathcal{D}f(a)\mathcal{I}_a\mathbf{1} + \mathcal{D}^2 f(a)\mathcal{I}_a^2\mathbf{1} + \mathcal{I}_a^3\mathcal{D}^3 f.
\end{aligned}
\tag{3.3}
$$

Clearly we may continue the same process for $\mathcal{D}^3 f, \mathcal{D}^4 f, \ldots \mathcal{D}^N f$ as long as all of these derivatives exist and are absolutely continuous (in fact, if $\mathcal{D}^N f$ is absolutely continuous, then all of the lower-order derivatives will be absolutely continuous as well). We may summarize the result as follows:

$$
f = f(a)\mathbf{1} + \mathcal{D}f(a)\mathcal{I}_a\mathbf{1} + \mathcal{D}^2 f(a)\mathcal{I}_a^2\mathbf{1} + \ldots + \mathcal{D}^N f(a)\mathcal{I}_a^N\mathbf{1} + \mathcal{I}_a^{N+1}\mathcal{D}^{N+1} f.
\tag{3.4}
$$

Equation (3.4) is actually Taylor's theorem in disguise. To see this, we only need to rewrite the functions $\mathcal{I}_a^n\mathbf{1}$ in conventional notation, and evaluate the integrals successively:

$$
\begin{aligned}
\mathcal{I}_a\mathbf{1}(x) &= \int_a^x dt_1 = (x - a); \\
\mathcal{I}_a^2\mathbf{1}(x) &= \int_a^x \mathcal{I}_a\mathbf{1}(t_1)dt_1 = \int_a^x (t_1 - a)dt_1 = \frac{(x-a)^2}{2}; \\
\mathcal{I}_a^3\mathbf{1}(x) &= \int_a^x \mathcal{I}_a^2\mathbf{1}(t_1)dt_1 = \int_a^x \frac{(t_1-a)^2}{2}dt_1 = \frac{(x-a)^3}{3!}; \\
&\cdots\cdots\cdots \\
\mathcal{I}_a^N\mathbf{1}(x) &= \int_a^x \mathcal{I}_a^{N-1}\mathbf{1}(x)dt_1 = \int_a^x \frac{(t_1-a)^{N-1}}{(N-1)!}dt_1 = \frac{(x-a)^N}{N!}.
\end{aligned}
\tag{3.5}
$$

It remains to evaluate the final term in (3.4), which is $\mathcal{I}_a^{N+1}\mathcal{D}^{N+1} f$. In this case, the integrand is not the constant function $\mathbf{1}$ as in the other integrals. However, we may use (2.16) to give an upper bound on this term:

$$
\begin{aligned}
|\mathcal{I}_a^{N+1}\mathcal{D}^{N+1} f| &\leq \sup_{c \in [a,x]} |\mathcal{D}^{N+1} f(c)|\mathcal{I}_a^{N+1}\mathbf{1} \\
&\leq \sup_{c \in [a,x]} |f^{(N+1)}(c)|\frac{(x-a)^{N+1}}{(N+1)!}
\end{aligned}
\tag{3.6}
$$

which is the bound on R_N in (2.3). For the exact evaluation of R_N, we may repeatedly use the fact that integral order can be exchanged:

$$
\int_a^{t_k} \int_a^{t_j} g(t_i, t_j)dt_i\, dt_j = \int_a^{t_k} \int_{t_i}^{t_k} g(t_i, t_j)dt_j\, dt_i.
\tag{3.7}
$$

Notice what happens to the integral limits under exchange: the limits of the outer integral do not change; the lower limit of the inner integral becomes the outer integration variable; and the upper

limit of the inner integral is the same as the upper limit of the outer integral. We may apply this rule first to exchange the dt_N and dt_{N+1} integrals (shown in parentheses):

$$
\begin{aligned}
\mathcal{I}_a^{N+1} & \mathcal{D}^{N+1} f \\
&= \int_a^x \ldots \int_a^{t_{N-2}} \left(\int_a^{t_{N-1}} \int_a^{t_N} f^{(n+1)}(t_{N+1}) dt_{N+1} dt_N \right) dt_{N-1} \ldots dt_1 \\
&= \int_a^x \ldots \int_a^{t_{N-2}} \left(\int_a^{t_{N-1}} \int_{t_{N+1}}^{t_{N-1}} f^{(n+1)}(t_{N+1}) dt_N dt_{N+1} \right) dt_{N-1} \ldots dt_1,
\end{aligned}
\tag{3.8}
$$

Next, we exchange the dt_{N-1} and dt_{N+1} integrals in similar fashion:

$$
\begin{aligned}
\mathcal{I}_a^{N+1} & \mathcal{D}^{N+1} f \\
&= \int_a^x \ldots \left(\int_a^{t_{N-2}} \int_a^{t_{N-1}} \left[\int_{t_{N+1}}^{t_{N-1}} f^{(n+1)}(t_{N+1}) dt_N \right] dt_{N+1} dt_{N-1} \right) \ldots dt_1 \\
&= \int_a^x \ldots \left(\int_a^{t_{N-2}} \int_{t_{N+1}}^{t_{N-2}} \left[\int_{t_{N+1}}^{t_{N-1}} f^{(N+1)}(t_{N+1}) dt_N \right] dt_{N-1} dt_{N+1} \right) \ldots dt_1,
\end{aligned}
\tag{3.9}
$$

where the integral in square brackets plays the role of $g(t_i, t_j)$ in (3.7), and the integrals in parentheses have been exchanged. In the same way we may exchange the dt_{N+1} integral successively with $dt_{N-2}, dt_{N-3}, \ldots dt_1$. The final result is

$$
\begin{aligned}
\mathcal{I}_a^{N+1} & \mathcal{D}^{N+1} f \\
&= \int_a^x \int_{t_{N+1}}^{t_1} \ldots \int_{t_{N+1}}^{t_{N-1}} f^{(N+1)}(t_{N+1}) dt_N \ldots dt_1 dt_{N+1} \\
&= \int_a^x f^{(N+1)}(t) \left(\int_t^{t_1} \ldots \int_t^{t_{N-1}} dt_N \ldots dt_1 \right) dt,
\end{aligned}
\tag{3.10}
$$

where the last expression in (3.10) is obtained by replacing t_{N+1} with t, and by moving $f^{(n+1)}(t)$ outside of the integrals over $dt_1, \ldots dt_N$. The N-fold integral in parentheses is identical to the final integral in (3.5), except that the lower limit is t instead of a. It follows that

$$
\mathcal{I}_a^{N+1} \mathcal{D}^{N+1} f = \int_a^x f^{(n+1)}(t) \frac{(x-t)^N}{N!} dt,
\tag{3.11}
$$

which agrees with (2.4).

3.2 Alternative evaluation using volume integrals

An alternative, elegant evaluation of the integrals in (3.10) is acheived by interpreting the integrals as volume integrals over regions in $\mathbb{R}^n, 1 \leq n \leq N$. The n-dimensional integral $\int_a^x \int_a^{t_1} \ldots \int_a^{t_n} dt_n \ldots dt_2 dt_1$ can be interpreted as the volume of the set S_n in $\mathbb{R}^n$, where

$$
\mathsf{S}_n := a \leq t_n \leq t_{n-1} \leq \ldots \leq t_1 \leq x.
\tag{3.12}
$$

But the ordering $t_n \leq t_{n-1} \leq \ldots \leq t_1$ is simply one of $n!$ possible orderings of the n variables $t_1, \ldots t_n$. Since these variables are dummy variables that are integrated over, the value of the integral does not depend on the ordering of the variables. For example, in the case where $n = 3$ we have $3! = 6$ different orderings of the variables, namely:

$$
\begin{aligned}
& a \leq t_3 \leq t_2 \leq t_1 \leq x; \quad a \leq t_3 \leq t_1 \leq t_2 \leq x; \quad a \leq t_2 \leq t_3 \leq t_1 \leq x; \\
& a \leq t_2 \leq t_1 \leq t_3 \leq x; \quad a \leq t_1 \leq t_3 \leq t_2 \leq x; \quad a \leq t_1 \leq t_2 \leq t_3 \leq x.
\end{aligned}
\tag{3.13}
$$

But these 3! orderings correspond to 3! disjoint sets that together make up the cube $[a, x]^3$ [1]. The volumes of these 3! sets are equal: therefore the volume of $\mathtt{S}_3$ is $1/3!$ of the volume of the cube, giving the same result as (3.5) for $N = 3$. The argument generalizes directly to n dimensions $(1 \leq n \leq N)$, giving the result:

$$\int_a^x \int_a^{t_1} \dots \int_a^{t_n} dt_n \dots dt_2 dt_1 = (\text{Volume of } \mathtt{S}_n) = \frac{(x-a)^n}{n!}, \tag{3.14}$$

which is the same as (3.5). To evaluate the final integral (3.11), we note that

$$\mathtt{S}_{N+1} \cap \{t_{N+1} = t\} = \mathtt{S}'_N(t), \tag{3.15}$$

where

$$\mathtt{S}'_N(t) := t \leq t_N \leq t_{N-1} \leq \dots \leq t_1 \leq x \tag{3.16}$$

so that the volume integral $\mathcal{I}_a^{N+1} \mathcal{D}^{N+1} f$ can be expressed as:

$$\begin{aligned}
\mathcal{I}_a^{N+1} \mathcal{D}^{N+1} f &= \int_a^x \left(\int_{\mathtt{S}'_N(t)} f^{(N+1)}(t) dt_N \dots dt_2 dt_1 \right) dt \\
&= \int_a^x f^{(N+1)}(t) \left(\int_{\mathtt{S}'_N(t)} dt_N \dots dt_2 dt_1 \right) dt \\
&= \int_a^x f^{(N+1)}(t) \frac{(t-a)^N}{N!} dt,
\end{aligned} \tag{3.17}$$

where we have used the volume formula (3.14) with $n \to N$ and $x \to t$. This result agrees with (3.11).

4 Discussion

The role of mathematics within science and society is changing rapidly. Because of the digital revolution, math is having an increasingly weighty impact on all aspects of society, including business and government. This impact comes through the need for mathematically-based algorithms and procedures in communication, classification, modeling, prediction, and control. Many of these results are framed in terms of linear algebraic concepts such as vectors, matrices, and tensors, while calculus and functional analysis enter in cases where the systems under study can be approximated as continuous.

The idea of linear operator (or functional) as a function of functions is an important bridge between linear algebra and mathematical analysis. It is thus one of the cornerstones of modern applied mathematics. In order to adequately prepare students to deal with current mathematical challenges in science and technology, the teaching of mathematics should make appropriate adaptations. The proof of Taylor's theorem in Section 3 emphasizes the relation between numerical methods and analytical theory, and is thus suitable for this purpose.

One possible objection to the proof is that it involves long equations (such as (3.9)) that appear very complicated. Many students may find such long equations difficult to deal with, because they lack the technical facility in algebraic manipulations. But in applied mathematics (particularly areas of mathematics related to computation, such as linear algebra and numerical analysis) these types of equations are very common. Although they appear complicated, in actuality these equations are built up from the repeated application of very simple ideas. Furthermore, such equations may readily handled by looking at simple cases first, and then building up to the general case. The proof in Section 3 is a good example of this approach. For this reason, we assert that the new perspective on Taylor's theorem provided in this paper can be a valuable addition to the undergraduate mathematics curriculum.

---

[1]There is a technical issue here in that the surfaces of these sets are not disjoint, but comprise a set of measure 0. However, it is intuitively clear that the volumes of these sets should add to the volume of the cube. In three dimensions, this can be demonstrated using a model.

Acknowledgements

Competing financial interests

The author declares no competing financial interests.

References

[1] Folland, Gerald Introduction to Partial Differential Equations (2nd ed.). *Princeton University Press* (1995).

[2] Zvonimir Šikić Taylor's theorem. *International Journal of Mathematical Education in Science and Technology* **21**, 111–115 (1990).

[3] Thomas, George Brinton and Weir, Maurice D and Hass, Joel Thomas' Calculus: Single Variable. *Pearson* (2013).

[4] Leithold, Louis The Calculus 7. *HarperCollins College Publishing* (1996).

[5] Courant, Richard and John, Fritz Introduction to Calculus and Analysis I. *Springer* (1999).

[6] Brauer, Fred A simplification of Taylor's theorem. *The American Mathematical Monthly* **94**, 453–455 (1987).

[7] Bárcenas, Diómedes The fundamental theorem of calculus for Lebesgue integral. *Divulgaciones matemáticas* **8**, 75–85 (2000).

[8] Weisstein, E.W. Fixed Point Theorem'. From Mathworld–a Wolfram Web Resource. *https://mathworld.wolfram.com/FixedPointTheorem.html* (Online; accessed 13-August-2022).

[9] Feller, William Probability Theory and its Applications (3rd ed.), vol II. *John Wiley & Sons* (1968).

[10] Angenent, Sigurd The Poincare Map. *https://people.math.wisc.edu/ angenent/519.2016s/notes/poincare-map.htmlperiodic-solutions* (2016).

[11] Panju, Maysum Iterative methods for computing eigenvalues and eigenvectors. *arXiv preprint arXiv:1105.1185* (2011).

[12] Newton, Roger G. Scattering theory of waves and particles. *Springer Science & Business Media* (2013).