Calculus: the greatest hits

Andrew Putman

Abstract

We give the most efficient proofs we know of a number of basic results in calculus: Taylor's theorem with remainder, L'Hôpital's rule, the change of variables theorem (both in single-variable calculus and in multi-variable calculus), and the two fundamental theorems of calculus.

1 Introduction

Over the years, I have tried to collect efficient but non-standard proofs of many results in elementary mathematics. The goal of these notes is to assemble some of these for calculus. The outline is as follows:

- In §2, I will prove Taylor's Theorem with the remainder in Lagrange form.
- In §3, I will prove L'Hôpital's rule.
- In §4, I will prove the change of variable theorem for integrals. This has two parts: in §4.1 I will deal with the one-variable case, and in §4.2 I will deal with the multi-variable case.
- In §5, I will prove both fundamental theorems of calculus.

I will try to give sources when possible, but I do not always remember where I learned these from. I would like to thank Anand Goyal for comments on a previous version.

2 Taylor's Theorem

I will start with Taylor's Theorem with the remainder written in Lagrange form. The idea of this proof came from Tim Gowers's blog [1], but also shares some ideas from a comment on that blog by Pieter-Jan De Smet.

Theorem 2.1 (Taylor's Theorem, Lagrange Form)**.** *Let $f\colon U \to \mathbb{R}$ be a function defined on an open connected set $U \subset \mathbb{R}$. Assume that f is $(n+1)$-times differentiable and let $x_0 \in U$. Then for all $x \in U \setminus \{x_0\}$, there exists some ζ in the interval between x_0 and x such that*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} \cdot (x-x_0)^k + \frac{f^{(n+1)}(\zeta)}{(n+1)!} \cdot (x-x_0)^{n+1}.$$

Proof. To simplify the notation, we will assume that $x > x_0$; the other case is similar. Since $(x - x_0)^{n+1} \neq 0$, there exists some $\lambda \in \mathbb{R}$ such that

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} \cdot (x - x_0)^k + \lambda \cdot (x - x_0)^{n+1}.$$

Our goal is to find some $\zeta \in (x_0, x)$ such that

$$\lambda = \frac{f^{(n+1)}(\zeta)}{(n+1)!}. \tag{2.1}$$

We will do this by repeatedly applying Rolle's Theorem. Define $g\colon U \to \mathbb{R}$ via the formula

$$g(t) = f(t) - \left(\sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} \cdot (t - x_0)^k + \lambda \cdot (t - x_0)^{n+1} \right).$$

We thus have $g(0) = g(x) = 0$, so by Rolle's Theorem there exists some $\zeta_1 \in (x_0, x)$ such that $g'(\zeta_1) = 0$. Since $g'(0) = g'(\zeta_1) = 0$, we can apply Rolle's Theorem again to find some $\zeta_2 \in (x_0, \zeta_1)$ such that $g''(\zeta_2) = 0$. Repeating this process, we produce a decreasing sequence $\zeta_1 > \zeta_2 > \cdots > \zeta_{n+1}$ of points in (x_0, x) such that $g^{(k)}(\zeta_k) = 0$ for all $1 \leq k \leq n+1$. In particular, setting $\zeta = \zeta_{n+1}$ we have

$$0 = g^{(n+1)}(\zeta_{n+1}) = f^{(n+1)}(\zeta_{n+1}) - \lambda \cdot (n+1)!.$$

Rearranging this formula yields (2.1). □

3 L'Hôpital's rule

I now turn to L'Hôpital's rule. To focus on the main idea here, I will only discuss limits that are finite real numbers. Also, I will not worry about minimizing the degree of differentiability of the functions involved. The following result will be what we prove:

Theorem 3.1 (L'Hôpital's rule)**.** *Let $f, g\colon \mathbb{R} \to \mathbb{R}$ be infinitely differentiable functions. For some $a \in \mathbb{R}$, assume that $f^{(k)}(a) = g^{(k)}(a) = 0$ for all $0 \leq k < n$ and that $g^{(n)}(a) \neq 0$. Then*

$$\lim_{x \mapsto a} \frac{f(x)}{g(x)} = \frac{f^{(n)}(a)}{g^{(n)}(a)}.$$

This is often derived from Cauchy's generalization of the mean value theorem, but I think it is more enlightening to derive it from Taylor's theorem (this is similar to how it was first proved by Bernoulli; see the discussion [3] on MathOverflow). The main idea is one that is familiar to anyone who has taught limits.

We first illustrate it in the special case where $f(x)$ and $g(x)$ are polynomials. The assumption that $f^{(k)}(a) = g^{(k)}(a) = 0$ for $0 \leq k < n$ implies that we can factor them as

$$f(x) = (x-a)^n F(x) \quad \text{and} \quad g(x) = (x-a)^n G(x) \tag{3.1}$$

for polynomials $F(x)$ and $G(x)$. We can then cancel the common factors $(x-a)^n$ and see that

$$\lim_{x \mapsto a} \frac{f(x)}{g(x)} = \lim_{x \mapsto a} \frac{(x-a)^n F(x)}{(x-a)^n G(x)} = \lim_{x \mapsto a} \frac{F(x)}{G(x)}.$$

Differentiating $f(x) = (x-a)^n F(x)$ a total of n times using the product rule and plugging in $x = a$ leads to $f^{(n)}(a) = F(a)$. Similarly, $g^{(n)}(a) = G(a)$. Using this, we see that

$$\lim_{x \mapsto a} \frac{f(x)}{g(x)} = \lim_{x \mapsto a} \frac{F(x)}{G(x)} = \frac{F(a)}{G(a)} = \frac{f^{(n)}(a)}{g^{(n)}(a)},$$

as desired.

Where did we use the fact that $f(x)$ and $g(x)$ were polynomials? The only step where that was relevant was the factorizations (3.1). The following theorem shows that this holds for general infinitely differentiable functions, so in fact the above argument is completely general:

Theorem 3.2. *Let $f\colon \mathbb{R} \to \mathbb{R}$ be an infinitely differentiable function. For some $a \in \mathbb{R}$, assume that $f^{(k)}(a) = 0$ for all $0 \leq k < n$. We can then write $f(x) = (x-a)^n F(x)$ for an infinitely differentiable function $F(X)$.*

Proof. Immediate from Taylor's theorem. □

4 Change of variables

We now turn to the change of variables theorem for integrals.

4.1 The one-variable case

The one-variable change of variables theorem is often derived from the second fundamental theorem of calculus, but I think the following direct proof is enlightening. Doing it this way also allows this theorem to be used in the proof of the first fundamental theorem of calculus, which is necessary for the traditional proof (otherwise, it is not clear that all functions have anti-derivatives!). I came up with this proof myself, but I doubt that I was the first to notice it.

Theorem 4.1 (Change of variables, one variable)**.** *Let $\phi\colon [a,b] \to [c,d]$ be a continuously differentiable orientation-preserving homeomorphism and let $f\colon [c,d] \to \mathbb{R}$ be a Riemann integrable function. Then*

$$\int_c^d f(x)\,\mathrm{dx} = \int_a^b f(\phi(y))\phi'(y)\,\mathrm{dy}\,.$$

Proof. Set $I = \int_c^d f(x)\,\mathrm{dx}$ and $J = \int_a^b f(\phi(y))\phi'(y)\,\mathrm{dy}$. Fix some $\epsilon > 0$. Since $f(\phi(y))\phi'(y)$ is Riemann integrable, we can find some partition

$$a = y_1 < y_2 < \cdots < y_n = b \tag{4.1}$$

that is fine enough such that for all choices of $z_i \in [y_i, y_{i+1}]$ for $1 \leq i < n$, we have

$$\|J - \sum_{i=1}^{n-1} f(\phi(z_i))\phi'(z_i)(y_{i+1} - y_i)\| < \epsilon.$$

Moreover, since $f(x)$ is Riemann integrable and ϕ is an orientation-preserving homeomorphism, we can make (4.1) even finer and ensure that the partition

$$c = \phi(y_1) < \phi(y_2) < \cdots < \phi(y_n) = d$$

of $[c, d]$ is fine enough such that for all choices of $w_i \in [\phi(y_i), \phi(y_{i+1})]$ for $1 \leq i < n$, we have

$$\|I - \sum_{i=1}^{n-1} f(w_i)(\phi(y_{i+1}) - \phi(y_i))\| < \epsilon.$$

The key trick is now to observe that by the mean value theorem, we can find some $z_i \in [y_i, y_{i+1}]$ for $1 \leq i < n$ such that

$$\phi'(z_i) = \frac{\phi(y_{i+1}) - \phi(y_i)}{y_{i+1} - y_i}.$$

Setting $w_i = \phi(z_i)$, we then have that

$$\sum_{i=1}^{n-1} f(\phi(z_i))\phi'(z_i)(y_{i+1} - y_i) = \sum_{i=1}^{n-1} f(w_i)(\phi(y_{i+1}) - \phi(y_i)).$$

The left hand side is within ϵ of I, and the right hand side is within ϵ of J. Since $\epsilon > 0$ was arbitrary, we conclude that $I = J$, as desired. □

4.2 The multi-variable case

We now turn to the multi-variable change of variables theorem. Here we give a proof that is inspired by the proof of Sard's Theorem in Milnor's differentiable topology textbook [2]. To avoid trying to pin down the precise conditions under which the theorem holds, we will only deal with compactly supported continuous functions. More general functions can be dealt with using a limiting process.

Theorem 4.2 (Change of variables, multi-variable)**.** *Let $U, V \subset \mathbb{R}^n$ be open sets and let $\phi\colon U \to V$ be an orientation-preserving diffeomorphism. Then for all compactly supported continuous functions $f\colon V \to \mathbb{R}$ we have*

$$\int_V f\,\mathrm{dx} = \int_U (f \circ \phi) \cdot J(\phi)\,\mathrm{dx},$$

where $J(\phi)\colon U \to \mathbb{R}$ is the determinant of the Jacobian of ϕ.

Proof. The proof will be by induction on n. We dealt with the base case $n = 1$ in the previous section, so assume that $n > 1$ and that the theorem is true for all smaller n.

The first step is to observe that for all points $p_0 \in U$, it is enough to prove that there exists a cube C containing p_0 such that

$$\int_{\phi(C)} f \,\mathrm{dx} = \int_C (f \circ \phi) \cdot J(\phi) \,\mathrm{dx} \quad \text{for all compactly supported functions } f \colon U \to \mathbb{R}.$$

Fix such a point p_0 and write ϕ in coordinates as $\phi = (\phi_1, \ldots, \phi_n)$. Permuting the coordinates if necessary and multiplying some of them by -1 if necessary (remember that we require ϕ to be orientation-preserving!), we can assume that $\frac{\partial \phi_1}{\partial x_1} > 0$. Define $\psi \colon U \to \mathbb{R}^n$ via the formula

$$\psi(x_1, \ldots, x_n) = (\phi_1(x_1, \ldots, x_n), x_2, \ldots, x_n) \quad \text{for } (x_1, \ldots, x_n) \in U.$$

By the inverse function theorem, ψ is a local diffeomorphism at p_0. We can thus find a cube C containing p_0 such that setting $D = \psi(C)$, we can restrict ψ to C and get a diffeomorphism $\psi \colon C \to D$. Define $\zeta \colon D \to V$ by setting $\zeta = \phi \circ \psi^{-1}$. The key fact about ζ is that

$$\zeta(x_1, \ldots, x_n) = (x_1, \zeta_2(x_1, \ldots, x_n), \ldots, \zeta_n(x_1, \ldots, x_n)) \quad \text{for } (x_1, \ldots, x_n) \in D.$$

We have $\phi(C) = \zeta(D)$, and the map $\phi \colon C \to f(C)$ factors as

$$C \xrightarrow{\psi} D \xrightarrow{\zeta} h(D).$$

Since

$$J(\phi)(q) = J(\psi \circ \zeta)(q) = J(\psi)(\zeta(q)) \cdot J(\zeta)(q) \quad \text{for } q \in C,$$

proving that the theorem holds for the diffeomorphisms ψ and ζ will imply that it also holds for ϕ. This is an easy application of Fubini's theorem. For both cases you slice the domain using $\mathbb{R}^1 \times \mathbb{R}^{n-1}$. For ψ, you first integrate the $\mathbb{R}^1$ factor using the one-variable case, and for ζ you first integrate the $\mathbb{R}^{n-1}$ factor using the inductive hypothesis. □

5 The fundamental theorems of calculus

We conclude by proving the two fundamental theorems of calculus. It is traditional to derive the second from the first, but I think it is very enlightening to prove both independently. I learned these proofs from Tao's book on measure theory [4].

Theorem 5.1 (First fundamental theorem of calculus)**.** *Let $f \colon [a, b] \to \mathbb{R}$ be a continuous function. Then*

$$\frac{d}{dx} \int_a^x f(t) \,\mathrm{dt} = f(x) \quad \textit{for } x \in [a, b].$$

Proof. Define $F \colon [a, b] \to \mathbb{R}$ via the formula

$$F(x) = \int_a^x f(t) \,\mathrm{dt} \quad \text{for } x \in [a, b].$$

For any $x \in [a,b]$ and $h \neq 0$ such that $x+h \in [a,b]$, we have

$$\frac{F(x+h)-F(x)}{h} = \frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{dt} = \int_0^1 f(x+th)\,\mathrm{dt},$$

where the last equality follows from a change of coordinates. As $h \mapsto 0$, the function $t \mapsto f(x+th)$ uniformly approaches the constant function $f(x)$. It follows that

$$\lim_{h\mapsto 0} \frac{F(x+h)-F(x)}{h} = \lim_{h\mapsto 0}\int_0^1 f(x+th)\,\mathrm{dt} = \int_0^1 f(x)\,\mathrm{dt} = f(x). \qquad \square$$

Theorem 5.2 (Second fundamental theorem of calculus)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be a differentiable function whose derivative $f'(x)$ is Riemann integrable. Then*

$$\int_a^b f'(x)\,\mathrm{dx} = f(b) - f(a).$$

Proof. Set $C = \int_a^b f'(x)\,\mathrm{dx}$, and fix some $\epsilon > 0$. We can then find a sufficiently fine partition

$$a = x_1 < x_2 < \cdots < x_n = b$$

such that for all choices of $y_i \in [x_i, x_{i+1}]$ for $1 \leq i < n$, we have

$$\left| C - \sum_{i=1}^{n-1} f'(y_i)\cdot(x_{i+1}-x_i)\right| < \epsilon. \tag{5.1}$$

By the mean value theorem, for all $1 \leq i < n$ we can find some $z_i \in [x_i, x_{i+1}]$ such that

$$f'(z_i) = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1}-x_i}.$$

We therefore see that we have a telescoping sum

$$\sum_{i=1}^{n-1} f'(z_i)\cdot(x_{i+1}-x_i) = \sum_{i=1}^{n-1}(f(x_{i+1})-f(x_i)) = f(x_n) - f(x_1) = f(b) - f(a).$$

Plugging this into (5.1), we get that

$$|C - (f(b)-f(a))| < \epsilon.$$

Since $\epsilon > 0$ was arbitrary, we conclude that

$$f(b) - f(a) = C = \int_a^b f'(x)\,\mathrm{dx}\,. \qquad \square$$

References

[1] T. Gowers, Taylor's theorem with the Lagrange form of the remainder, blogpost on Gower's Weblog, `https://tinyurl.com/yd5ob6y3`.

[2] J. W. Milnor, *Topology from the differentiable viewpoint*, revised reprint of the 1965 original, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997.

[3] John Palmieri (https://mathoverflow.net/users/4194/john-palmieri), How did Bernoulli prove L'Hôpital's rule?, URL (version: 2014-11-22): https://mathoverflow.net/q/51685

[4] T. Tao, *An introduction to measure theory*, Graduate Studies in Mathematics, 126, American Mathematical Society, Providence, RI, 2011.

Andrew Putman
Department of Mathematics
University of Notre Dame
164 Hurley Hall
Notre Dame, IN 46556
andyp@nd.edu