# Taylor's Formula

G. B. Folland

There's a lot more to be said about Taylor's formula than the brief discussion on pp.113–4 of Apostol. Let me begin with a few definitions.

*Definitions.* A function $f$ defined on an interval $I$ is called *$k$ times differentiable* on $I$ if the derivatives $f', f'', \ldots, f^{(k)}$ exist and are finite on $I$, and $f$ is said to be of *class $C^k$* on $I$ if these derivatives are all continuous on $I$. (Note that if $f$ is $k$ times differentiable, the derivatives $f', \ldots, f^{(k-1)}$ are necessarily continuous, by Theorem 5.3; the only question is the continuity of $f^{(k)}$.) If $f$ is (at least) $k$ times differentiable on an open interval $I$ and $c \in I$, its *$k$th order Taylor polynomial about $c$* is the polynomial

$$P_{k,c}(x) = \sum_{j=0}^{k} \frac{f^{(j)}(c)}{j!}(x-c)^j$$

(where, of course, the "zeroth derivative" $f^{(0)}$ is $f$ itself), and its *$k$th order Taylor remainder* is the difference

$$R_{k,c}(x) = f(x) - P_{k,c}(x).$$

*Remark 1.* The $k$th order Taylor polynomial $P_{k,c}(x)$ is a polynomial of degree at most $k$, but its degree may be less than $k$ because $f^{(k)}(c)$ might be zero.

*Remark 2.* We have $P_{k,c}(c) = f(c)$, and by differentiating the formula for $P_{k,c}(x)$ repeatedly and then setting $x = c$ we see that $P_{k,c}^{(j)}(c) = f^{(j)}(c)$ for $j \le k$. That is, *$P_{k,c}$ is the polynomial of degree $\le k$ whose whose derivatives of order $\le k$ at $c$ agree with those of $f$.*

For future reference, here are a few frequently used examples of Taylor polynomials:

$$f(x) = e^x; \qquad P_{k,0}(x) = \sum_{0 \le j \le k} \frac{x^j}{j!}$$

$$f(x) = \cos x; \qquad P_{k,0}(x) = \sum_{0 \le j \le k/2} \frac{(-1)^j x^{2j}}{(2j)!}$$

$$f(x) = \sin x; \qquad P_{k,0}(x) = \sum_{0 \le j < k/2} \frac{(-1)^j x^{2j+1}}{(2j+1)!}$$

$$f(x) = \log x; \qquad P_{k,1}(x) = \sum_{1 \le j \le k} \frac{(-1)^{j-1}(x-1)^j}{j}$$

Note that (for example) $1 - \frac{1}{2}x^2$ is both the 2nd order and the 3rd order Taylor polynomial of $\cos x$, because the cubic term in its Taylor expansion vanishes. (Also note that in higher mathematics the natural logarithm function is almost always called log rather than ln.)

For $k = 1$ we have $P_{1,c}(x) = f(c) + f'(c)(x-c)$; this is the linear function whose graph is the tangent line to the graph of $f$ at $x = c$. Just as this tangent line is the straight line that best approximates the graph of $f$ near $x = c$, we shall see that $P_{k,c}(x)$ is the polynomial of degree $\leq k$ that best approximates $f(x)$ near $x = c$. To justify this assertion we need to see that the remainder $R_{k,c}(x)$ is suitably small near $x = c$, and there are several ways of making this precise. The first one is simply this: the remainder $R_{k,c}(x)$ tends to zero as *$x \to c$ faster than any nonzero term in the polynomial $P_{k,c}(x)$*, that is, faster than $(x-c)^k$. Here is the result:

**Theorem 1.** *Suppose $f$ is $k$ times differentiable in an open interval $I$ containing the point $c$. Then*

$$\lim_{x \to c} \frac{R_{k,c}(x)}{(x-c)^k} = \lim_{x \to c} \frac{f(x) - P_{k,c}(x)}{(x-c)^k} = 0.$$

*Proof.* Since $f$ and its derivatives up to order $k$ agree with $P_{k,c}$ and its derivatives up to order $k$ at $x = c$, the difference $R_{k,c}$ and its derivatives up to order $k$ vanish at $x = c$. Moreover, $(x-c)^k$ and its derivatives up to order $k-1$ also vanish at $x = c$, so we can apply l'Hôpital's rule $k$ times to obtain

$$\lim_{x \to c} \frac{R_{k,c}(x)}{(x-c)^k} = \lim_{x \to c} \frac{R^{(k)}_{k,c}(x)}{k(k-1)\cdots 1(x-c)^0} = \frac{0}{k!} = 0.$$

∎

There is a convenient notation to describe the situation in Theorem 1: we say that

$$R_{k,c}(x) = o((x-c)^k) \text{ as } x \to c,$$

meaning that $R_{k,c}(x)$ is *of smaller order than* $(x-c)^k$ as $x \to c$. More generally, if $g$ and $h$ are two functions, we say that $h(x) = o(g(x))$ as $x \to c$ (where $c$ might be $\pm\infty$) if $h(x)/g(x) \to 0$ as $x \to c$. The symbol $o(g(x))$ is pronounced "little oh of $g$ of $x$"; it does not denote any particular function, but rather is a shorthand way of describing any function that is of smaller order than $g(x)$ as $x \to c$. For example, Corollary 1 of l'Hôpital's rule (see the notes on l'Hôpital's rule) says that for any $a > 0$, $x^a = o(e^x)$ and $\log x = o(x^a)$ as $x \to \infty$, and $\log x = o(x^{-a})$ as $x \to 0+$. Another example: saying that $h(x) = o(1)$ as $x \to c$ simply means that $\lim_{x \to c} h(x) = 0$.

In order to simplify notation, in the following discussion we shall assume that $c = 0$ and write $P_k$ instead of $P_{k,c}$. (The Taylor polynomial $P_k = P_{k,0}$ is often called the $k$th order *Maclaurin polynomial* of $f$.) There is no loss of generality in doing this, as one can always reduce to the case $c = 0$ by making the change of variable $\widetilde{x} = x - c$ and regarding all functions in question as functions of $\widetilde{x}$ rather than $x$.

The conclusion of Theorem 1, that $f(x) - P_k(x) = o(x^k)$, actually characterizes the Taylor polynomial $P_{k,c}$ completely:

**Theorem 2.** *Suppose $f$ is $k$ times differentiable on an open interval $I$ containing 0. If $Q$ is a polynomial of degree $\leq k$ such that $f(x) - Q(x) = o(x^k)$ as $x \to 0$, then $Q = P_k$.*

*Proof.* Since $f-Q$ and $f-P_k$ are both of smaller order than $x^k$, so is their difference $P_k-Q$. Let $P_k(x)=\sum_0^k a_jx^j$ (of course $a_j=f^{(j)}(0)/j!$) and $Q(x)=\sum_0^k b_jx^j$. Then

$$(a_0-b_0)+(a_1-b_1)x+\cdots+(a_k-b_k)x^k=P_k(x)-Q(x)=o(x^k).$$

Letting $x\to 0$, we see that $a_0-b_0=0$. This being the case, we have

$$(a_1-b_1)+(a_2-b_2)x+\cdots+(a_k-b_k)x^{k-1}=\frac{P_k(x)-Q(x)}{x}=o(x^{k-1}).$$

Letting $x\to 0$ here, we see that $a_1-b_1=0$. But then

$$(a_2-b_2)+(a_3-b_3)x+\cdots+(a_k-b_k)x^{k-2}=\frac{P_k(x)-Q(x)}{x^2}=o(x^{k-2}),$$

which likewise gives $a_2-b_2=0$. Proceeding inductively, we find that $a_j=b_j$ for all $j$ and hence $P_k=Q$. ∎

Theorem 2 is very useful for calculating Taylor polynomials. It shows that using the formula $a_k=f^{(k)}(0)/k!$ is not the only way to calculate $P_k$; rather, if by *any* means we can find a polynomial $Q$ of degree $\le k$ such that $f(x)=Q(x)+o(x^k)$, then $Q$ must be $P_k$. Here are two important applications of this fact.

*Taylor Polynomials of Products.* Let $P_k^f$ and $P_k^g$ be the $k$th order Taylor polynomials of $f$ and $g$, respectively. Then

$$\begin{aligned} f(x)g(x) &= \left[P_k^f(x)+o(x^k)\right]\left[P_k^g(x)+o(x^k)\right] \\ &= \left[\text{terms of degree} \le k \text{ in } P_k^f(x)P_k^g(x)\right]+o(x^k). \end{aligned}$$

Thus, to find the $k$th order Taylor polynomial of $fg$, simply multiply the $k$th Taylor polynomials of $f$ and $g$ together, discarding all terms of degree $>k$.

EXAMPLE 1. What is the 6th order Taylor polynomial of $x^3e^x$? Solution:

$$x^3e^x=x^3\left[1+x+\frac{x^2}{2}+\frac{x^3}{6}+o(x^3)\right]=x^3+x^4+\frac{x^5}{2}+\frac{x^6}{6}+o(x^6),$$

so the answer is $x^3+x^4+\frac{1}{2}x^5+\frac{1}{6}x^6$.

EXAMPLE 2 What is the 5th order Taylor polynomial of $e^x\sin 2x$? Solution:

$$\begin{aligned} e^x\sin 2x &= \left[1+x+\frac{x^2}{2}+\frac{x^3}{6}+\frac{x^4}{24}+\frac{x^5}{120}+o(x^5)\right]\left[2x-\frac{(2x)^3}{6}+\frac{(2x)^5}{120}+o(x^5)\right] \\ &= 2x+2x^2+x^3\left[\frac{2}{2}-\frac{8}{6}\right]+x^4\left[\frac{2}{6}-\frac{8}{6}\right]+x^5\left[\frac{2}{24}-\frac{8}{12}+\frac{32}{120}\right]+o(x^5), \end{aligned}$$

so the answer is $2x+2x^2-\frac{1}{3}x^3-x^4-\frac{19}{60}x^5$.

*Taylor Polynomials of Compositions.* If $f$ and $g$ have derivatives up to order $k$, and $g(0) = 0$, we can find the $k$th Taylor polynomial of $f \circ g$ by substituting the Taylor expansion of $g$ into the Taylor expansion of $f$, retaining only the terms of degree $\le k$. That is, suppose

$$f(x) = a_0 + a_1 x + \cdots + a_k x^k + o(x^k).$$

Since $g(0) = 0$ and $g$ is differentiable, we have $g(x) \approx g'(0)x$ for $x$ near 0, so anything that is $o(g(x)^k)$ is also $o(x^k)$ as $x \to 0$. Hence,

$$f(g(x)) = a_0 + a_1 g(x) + \cdots + a_k g(x)^k + o(x^k).$$

Now plug in the Taylor expansion of $g$ on the right and multiply it out, discarding terms of degree $> k$.

EXAMPLE 3. What is the 16th order Taylor polynomial of $e^{x^6}$? Solution:

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + o(x^3) \qquad \Longrightarrow \qquad e^{x^6} = 1 + x^6 + \frac{x^{12}}{2} + \frac{x^{18}}{6} + o(x^{18}).$$

The last two terms are both $o(x^{16})$, so the answer is $1 + x^6 + \frac{1}{2}x^{12}$.

EXAMPLE 4. What is the 4th order Taylor polynomial of $e^{\sin x}$? Solution:

$$e^{\sin x} = 1 + \sin x + \frac{\sin^2 x}{2} + \frac{\sin^3 x}{6} + \frac{\sin^4 x}{24} + o(x^4)$$

since $|\sin x| \le |x|$. Now substitute $x - \frac{1}{6}x^3 + o(x^4)$ for $\sin x$ on the right (yes, the error term is $o(x^4)$ because the 4th degree term in the Taylor expansion of $\sin x$ vanishes) and multiply out, throwing all terms of degree $> 4$ into the "$o(x^4)$" trash can:

$$e^{\sin x} = 1 + \left[x - \frac{x^3}{6}\right] + \frac{1}{2}\left[x^2 - \frac{x^4}{3}\right] + \frac{x^3}{6} + \frac{x^4}{24} + o(x^4),$$

so the answer is $1 + x + \frac{1}{2}x^2 - \frac{1}{8}x^4$. (To appreciate how easy this is, try finding this polynomial by computing the first four derivatives of $e^{\sin x}$.)

*Taylor Polynomials and l'Hôpital's Rule.* Taylor polynomials can often be used effectively in computing limits of the form 0/0. Indeed, suppose $f$, $g$, and their first $k-1$ derivatives vanish at $x = 0$, but their $k$th derivatives do not both vanish. The Taylor expansions of $f$ and $g$ then look like

$$f(x) = \frac{f^{(k)}(0)}{k!}x^k + o(x^k), \qquad g(x) = \frac{g^{(k)}(0)}{k!}x^k + o(x^k).$$

Taking the quotient and cancelling out $x^k/k!$, we get

$$\frac{f(x)}{g(x)} = \frac{f^{(k)}(0) + o(1)}{g^{(k)}(0) + o(1)} \to \frac{f^{(k)}(0)}{g^{(k)}(0)} \text{ as } x \to 0.$$

This is in accordance with l'Hôpital's rule, but the devices discussed above for computing Taylor polynomials may lead to the answer more quickly than a direct application of l'Hôpital.

EXAMPLE 5. What is $\lim_{x\to 0}(x^2 - \sin^2 x)/x^2 \sin^2 x$? Solution:

$$\sin^2 x = \left[x - \frac{x^3}{6} + o(x^4)\right]^2 = x^2 - \frac{x^4}{3} + o(x^4),$$

so $x^2 \sin^2 x = x^4 + o(x^4)$, and

$$\frac{x^2 - \sin^2 x}{x^2 \sin^2 x} = \frac{\frac{1}{3}x^4 + o(x^4)}{x^4 + o(x^4)} = \frac{\frac{1}{3} + o(1)}{1 + o(1))} \to \frac{1}{3}.$$

(Again, to appreciate how easy this is, try doing it by l'Hôpital's rule.)

EXAMPLE 6. Evaluate

$$\lim_{x\to 1}\left[\frac{1}{\log x} + \frac{x}{x-1}\right].$$

Solution: Here we need to expand in powers of $x - 1$. First of all,

$$\frac{1}{\log x} - \frac{x}{x-1} = \frac{x - 1 - x\log x}{(x-1)\log x} = \frac{(x-1) - (x-1)\log x - \log x}{(x-1)\log x}.$$

Next, $\log x = (x-1) - \frac{1}{2}(x-1)^2 + o((x-1)^2)$, and plugging this into the numerator and denominator gives

$$\frac{(x-1) - (x-1)^2 - \left[(x-1) - \frac{1}{2}(x-1)^2\right] + o((x-1)^2)}{(x-1)^2 + o((x-1)^2)} = \frac{-\frac{1}{2} + o(1)}{1 + o(1)} \to -\frac{1}{2}.$$

Theorem 1 tells us a lot about the remainder $R_{k,c}(x) = f(x) - P_{k,c}(x)$ for small $x$, but sometimes one wants a more precise quantitative estimate of it. The most common ways of obtaining such an estimate involve slightly stronger conditions on $f$; namely, instead of just being $k$ times differentiable we assume that it is $k+1$ times differentiable, or perhaps of class $C^{k+1}$, and the estimates we obtain involve bounds on the derivative $f^{(k+1)}$. There are several formulas for $R_{k,c}(x)$ that lead to such estimates; we shall present the two that are most often encountered. The first one is the one presented in Apostol. (It's Theorem 5.19, with the change of variable $k = n-1$. Apostol states the hypotheses in a slightly more general, but also more complicated, form; the version below usually suffices.)

**Theorem 3** (Lagrange's Form of the Remainder)**.** *Suppose $f$ is $k+1$ times differentiable on an open interval $I$ and $c \in I$. For each $x \in I$ there is a point $x_1$ between $c$ and $x$ such that*

$$R_{k,c}(x) = \frac{f^{(k+1)}(x_1)}{(k+1)!}(x-c)^{k+1}. \tag{1}$$

For the proof of Theorem 3 we refer to Apostol. The other popular form of the remainder requires a slightly stronger hypothesis, that $f^{(k+1)}$ not only exists but is continuous. (Actually, it's enough for it to be Riemann integrable, but these minor variations in the assumptions are usually of little importance.) I suspect the reason that Apostol doesn't mention it is that it involves an integral, and he doesn't want to discuss integrals until later.

**Theorem 4** (Integral Form of the Remainder)**.** *Suppose $f$ is of class $C^{k+1}$ on an open interval $I$ and $c \in I$. If $x \in I$, then*

$$R_{k,c}(x) = \frac{1}{(k+1)!}\int_c^x (x-t)^k f^{(k+1)}(t)\,dt. \tag{2}$$

*Proof.* Recalling the definition of $R_{k,c}$, we can restate (2) as

$$f(x) = \sum_{j=0}^{k} \frac{f^{(j)}(c)}{j!}(x-c)^j + \frac{1}{(k+1)!}\int_c^x (x-t)^k f^{(k+1)}(t)\,dt. \tag{3}$$

For $k = 0$, this simply says that

$$f(x) = f(c) + \int_c^x f'(t)\,dt, \tag{4}$$

which is true by the fundamental theorem of calculus. Next, we integrate (4) by parts, taking

$$u = f'(t), \quad du = f''(t)\,dt; \qquad dv = dt, \quad v = t - x.$$

Notice the twist: normally if $dv = dt$ we would simply take $v = t$, but we are free to add a constant of integration, and we take that constant to be $-x$. (The number $x$, like $c$, is fixed in this discussion; the variable of integration is $t$.) The result is

$$\begin{aligned} f(x) &= f(c) + (t-x)f'(t)\Big|_c^x - \int_c^x (t-x)f''(t)\,dt \\ &= f(c) + (x-c)f'(c) + \int_c^x (x-t)f''(t)\,dt, \end{aligned}$$

which is (3) with $k = 1$. Another integration by parts, with

$$u = f''(t), \quad du = f'''(t)\,dt; \qquad dv = (x-t)\,dt, \quad v = -\tfrac{1}{2}(x-t)^2,$$

(again, instead of taking $v = xt - \frac{1}{2}t^2$ we take $v = -\frac{1}{2}(x-t)^2 = -\frac{1}{2}x^2 + xt - \frac{1}{2}t^2$) gives

$$\begin{aligned} f(x) &= f(c) + (x-c)f'(c) - \tfrac{1}{2}(x-t)^2 f''(t)\Big|_c^x + \int_c^x \tfrac{1}{2}(x-t)^2 f'''(t)\,dt \\ &= f(c) + (x-c)f'(c) + \frac{(x-c)^2}{2!}f''(c) + \frac{1}{2!}\int_c^x (x-t)^2 f'''(t)\,dt, \end{aligned}$$

which is (3) with $k = 2$. The pattern should now be clear: a $k$-fold integration by parts starting from (4) yields (3). The formal inductive proof is left to the reader. ∎

Let's be clear about the significance of Theorems 3 and 4. They are almost never used to find the exact value of the remainder term (which amounts to knowing the exact value of the original $f(x)$); one doesn't know just where the point $x_1$ in (1) is, and the integral in

(2) is usually hard to evaluate. Instead, the philosophy is that Taylor polynomials $P_{k,c}$ are used as (simpler) approximations to (complicated) functions $f$ near $c$, and the remainders $R_{k,c}$ are regarded as junk to be disregarded. For this to work one needs some assurance that $R_{k,c}(x)$ is small enough that one can safely neglect it or an estimate of the magnitude of the error one makes in doing so. The main purpose of Theorems 3 and 4 is to provide such information via the following result.

**Corollary 1.** *Suppose $f$ is $k+1$ times differentiable on an interval $I$ and that $|f^{(k+1)}(x)| \leq C$ for $x \in I$. Then for any $x, c \in I$ we have*

$$|R_{k,c}(x)| \leq C\frac{|x-c|^{k+1}}{(k+1)!}. \tag{5}$$

*Proof.* The estimate (5) is clearly an immediate consequence of (1). It also follows easily from (2): if $x > c$,

$$|R_{k,c}(x)| \leq \frac{C}{k!}\int_c^x (x-t)^k\,dt = -\frac{C}{k!}\frac{(x-t)^{k+1}}{k+1}\bigg|_c^x = C\frac{(x-c)^{k+1}}{(k+1)!},$$

and if $x < c$,

$$|R_{k,c}(x)| \leq \frac{C}{k!}\left|\int_c^x (x-t)^k\,dt\right| = \frac{C}{k!}\int_x^c (t-x)^k\,dt = \frac{C}{k!}\frac{(c-x)^{k+1}}{k+1} = C\frac{|x-c|^{k+1}}{(k+1)!}.$$

∎

Observe that Corollary 1 is a more precise and quantitative version of Theorem 1 (under slightly stronger hypotheses on $f$): Theorem 1 says that $R_{k,c}(x)$ vanishes faster than $(x-c)^k$ as $x \to c$; Corollary 1 says that it vanishes at least at a rate proportional to $(x-c)^{k+1}$ and gives a good estimate for the proportionality constant. The best estimate is obtained by taking $C$ to be the *least* upper bound for $|f^{(k+1)}|$ on $I$, but it is usually not crucial to compute this optimal value for $C$. What *is* crucial, however, and what some people find easy to forget, is the use of absolute values. It's the size of $R_{k,c}(x)$ that matters.

A typical use of Taylor polynomials is to evaluate integrals of functions that don't have an elementary antiderivative. Here's an example.

EXAMPLE 7. The function $f(x) = e^{-x^2}$ has no elementary antiderivative. However, we can do a Taylor approximation of $e^{-x^2}$ and integrate the resulting polynomial. The efficient way to proceed is to consider the Taylor approximations of $e^{-y}$ (easier to compute with!) and then set $y = x^2$. Since $|(d/dy)^j e^{-y}| = |(-1)^j e^{-y}| \leq 1$ for $y \geq 0$, the estimate (5) shows that

$$e^{-y} = 1 - y + \frac{y^2}{2} - \cdots + (-1)^k\frac{y^k}{k!} + R_{k,0}(y), \text{ where } |R_{k,0}(y)| \leq \frac{y^{k+1}}{(k+1)!} \text{ for } y \geq 0.$$

Setting $y = x^2$ yields

$$e^{-x^2} = 1 - x^2 + \frac{x^4}{2!} - \cdots + (-1)^k\frac{x^{2k}}{k!} + R_{k,0}(x^2), \text{ where } |R_{k,0}(x^2)| \leq \frac{x^{2k+2}}{(k+1)!}.$$

Therefore,

$$\int_0^x e^{-t^2}\,dt = x - \frac{x^3}{3} + \frac{x^5}{5\cdot 2!} - \cdots + (-1)^k \frac{x^{2k+1}}{(2k+1)\cdot k!} + \text{ error},$$

where

$$|\text{error}| \le \left| \int_0^x \frac{t^{2k+2}}{(k+1)!}\,dt \right| = \frac{|x|^{2k+3}}{(2k+3)\cdot(k+1)!}.$$

For instance, if $x = 1$, we can take $k = 4$ and obtain

$$\int_0^1 e^{-t^2}\,dt = 1 - \frac{1}{3} + \frac{1}{5\cdot 2} - \frac{1}{7\cdot 3!} + \frac{1}{9\cdot 4!} = 0.7382 \text{ with error less than } 0.0008.$$

*A Few Concluding Remarks.* Although Theorems 3 and 4 are most commonly used through Corollary 1, there are other things that can be done with them. There's a nice application of Theorem 3 on p.376 of Apostol, which we'll discuss toward the end of the quarter. For an extra twist on Theorem 4 that yields more estimates, as well as a sharper form of Theorem 1, see my paper "Remainder estimates in Taylor's theorem," *American Mathematical Monthly* **97** (1990), 233–235 (available online through the UW Libraries site).