Taylor’s theorem plays a central role in numerical analysis, providing a method for approximating arbitrary functions by polynomials, and estimating the errors. There are several common proofs. The one that follows, from James Wolfe, “A proof of Taylor’s formula,” American mathematical monthly, 60 (1953), 415-416, is appropriate because it depends on an extension of Rolle’s theorem, like some other results fundamental to this subject.

According to Rolle’s theorem, if a function \( g \) is differentiable on an open interval \( I = (a, b) \), continuous at \( a \) and \( b \), and \( g(a) = g(b) = 0 \), then \( g'(\xi) = 0 \) for some \( \xi \) in \( I \). The extension needed here is concerned with higher derivatives:

if \( g \) has \( n \) continuous derivatives on the closed interval \([a, b]\), \( g^{(n+1)} \) exists on its interior \( I \), \( g^{(j)}(a) = 0 \) for \( j = 0 \) to \( n \), and \( g(b) = 0 \), then \( g^{(n+1)}(\xi) = 0 \) for some \( \xi \) in \( I \).

The proof is simple. First, \( g \) satisfies the hypotheses of Rolle’s theorem on the interval \([a, b]\), so \( g'(x_1) = 0 \) for some \( x_1 \) in \( I \). Now apply the same argument to \( g' \) on the interval \( I_1 = [a, x_1] \) to determine \( x_2 \) in \( I_1 \) with \( g''(x_2) = 0 \). Continue the process until you’ve found \( \xi = x_{n+1} \) in \( I_n \) with \( g^{(n+1)}(x_{n+1}) = 0 \).

With this preliminary step completed, you can turn to approximation by polynomials. Consider, as an example, a fourth degree polynomial and its derivatives:

\[
\begin{align*}
P(x) &= P^{(0)}(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + a_3(x-a)^3 + a_4(x-a)^4 \\
P'(x) &= P^{(1)}(x) = a_1 + 2a_2(x-a) + 3a_3(x-a)^2 + 4a_4(x-a)^3 \\
P''(x) &= P^{(2)}(x) = 2a_2 + 2\cdot3a_3(x-a) + 3\cdot4a_4(x-a)^2 \\
P^{(3)}(x) &= 2\cdot3a_3 + 2\cdot3\cdot4a_4(x-a) \\
P^{(4)}(x) &= 2\cdot3\cdot4a_4 
\end{align*}
\]

This should convince you that, in general, for an \( n \)th degree polynomial

\[
P(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + \ldots + a_n(x-a)^n \quad (*)
\]

the formula

\[
P^{(j)}(a) = j! a_j
\]

holds for \( j = 0 \) to \( n \). You can use it in reverse: to construct a polynomial \((*)\) with prescribed derivatives \( P^{(j)}(a) = p_j \), set \( a_j = p_j / j! \) for \( j = 0 \) to \( n \).
Now let \( f \) be any function with \( n \) derivatives at a point \( a \). Use the method of the previous paragraph to construct a polynomial \( T_{n,a}(x) \) with the same derivatives there:

\[
T_{n,a}(x) = a_0 + a_1(x - a) + a_2(x - a)^2 + \ldots + a_n(x - a)^n
\]

\[
a_j = \frac{f^{(j)}(a)}{j!}
\]

for \( j = 0 \) to \( n \).

\( T_{n,a}(a) = f^{(j)}(a) \) is called the \( n \)th degree Taylor polynomial for \( f \) at \( a \). When \( a = 0 \), it’s also called a Maclaurin polynomial.

If \( f \) is sufficiently smooth, you can approximate values \( f(x) \) by those of Taylor polynomials even at arguments \( x \) distant from \( a \). Taylor’s theorem lets you determine how good an approximation it is:

Suppose \( f \) has \( n \) continuous derivatives on the closed interval \([a, b]\), and the \( n+1 \)st exists on its interior \( I \). Then there exists \( \xi \) in \( I \) such that

\[
f(b) = T_{n,a}(b) + \frac{f^{(n+1)}(\xi)}{(n + 1)!}(b - a)^{n+1}.
\]

The right hand term is called the error term: it tells how far the Taylor polynomial deviates from the function value at \( x \). There are several versions of Taylor’s theorem with different but equivalent error terms; this one is known as Lagrange’s. If you can estimate the size of \( f^{(n+1)}(\xi) \), then you can estimate the accuracy of the Taylor approximation. Note that Taylor’s theorem for \( n = 0 \) is simply the mean-value theorem.

To follow Wolfe’s proof, define

\[
k = \frac{f(b) - T_{n,a}(b)}{(b - a)^{n+1}}
\]

\[
g(x) = f(x) - T_{n,a}(x) - k(x - a)^{n+1}.
\]

for all \( x \). Then for \( j = 0 \) to \( n \),

\[
g^{(j)}(a) = f^{(j)}(a) - T^{(j)}_{n,a}(a) - k \frac{d^j}{dx^j}(x - a)^{n+1} \bigg|_{x=a}
\]

\[= f^{(j)}(a) - f^{(j)}(a) - k \cdot 0 = 0.\]
Moreover, \( g(b) = f(b) - T_{n,a}(b) - k(b - a)^{n+1} = 0 \) by the definition of \( k \). By the extended Rolle’s theorem, \( g^{(n+1)}(\xi) = 0 \) for some \( \xi \) in \( I \). That means

\[
0 = g^{(n+1)}(\xi) \quad f^{(n+1)}(\xi) - T_{n,a}^{(n+1)}(x) - k \frac{d^{n+1}}{dx^{n+1}}(x - a)^{n+1}\bigg|_{x=a} \\
= f^{(n+1)}(\xi) - (n+1)! k \\
0 = f^{(n+1)}(\xi) - (n+1)! \frac{f(b) - T_{n,a}(b)}{(b-a)^{n+1}}.
\]

The equation in Taylor’s theorem is just a rearrangement of this last one.

As an example, you can construct some Maclaurin polynomials for \( f(x) = \sin x \) at \( a = 0 \). Note that

\[
f(0) = 0 \quad f'(0) = 1 \quad f''(0) = 0 \quad f^{(3)}(0) = -1
\]

and subsequent derivatives repeat these values cyclically. Each polynomial of odd degree \( n \) is identical to the next one, hence its error term has degree \( n + 2 \). Since the derivatives of the sine are positive or negative sines or cosines, you can easily find error bounds. For \( n = 7 \), for instance, here’s the Maclaurin polynomial, its error term, and an error bound:

\[
T_{7,0}(x) = T_{8,0}(x) = x - x^3/3! + x^5/5! - x^7/7! \\
\left|\frac{f^{(9)}(\xi)}{9!}\right| x^9 \leq \frac{|x|^9}{9!}.
\]

The following table shows the values of these bounds for \( x = 3\pi/2 \) and \( n = 1 \) to \( 13 \). The polynomial approximations to \( \sin x = -1 \) are also shown, as well as the actual error.
Figure 1 compares graphs of all but the last of these polynomials with that of the sine.

Because of the size of the errors, Maclaurin polynomials do not provide a reasonable way of calculating $\sin x$ except for very small values of $x$. Other methods are used in practice.

**Figure 1** $\sin x$ and its first few Maclaurin polynomials