Here’s some reflection on the proof(s) of Taylor’s theorem. First we recall the (derivative form) of the theorem:

**Theorem 1** (Taylor’s theorem). Suppose $f: (a, b) \to \mathbb{R}$ is a function on $(a, b)$, where $a, b \in \mathbb{R}$ with $a < b$. Assume that for some positive integer $n$, $f$ is $n$-times differentiable on the open interval $(a, b)$, and that $f, f', f'', \ldots, f^{(n-1)}$ all extend continuously to the closed interval $[a, b]$ (the extended functions will still be called $f, f', f'', \ldots, f^{(n-1)}$ respectively). Then there exists $c \in (a, b)$ such that

$$f(b) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (b - a)^k + \frac{f^{(n)}(c)}{n!} (b - a)^n.$$ 

A key observation is that when $n = 1$, this reduces to the ordinary mean-value theorem. This suggests that we may modify the proof of the mean value theorem, to give a proof of Taylor’s theorem.

The proof of the mean-value theorem comes in two parts: first, by subtracting a linear (i.e. degree 1) polynomial, we reduce to the case where $f(a) = f(b) = 0$. Next, the special case where $f(a) = f(b) = 0$ follows from Rolle’s theorem.

In the proof of the Taylor’s theorem below, we mimic this strategy.

The key is to observe the following generalization of Rolle’s theorem:

**Proposition 2.** Suppose $F: (a, b) \to \mathbb{R}$ is a function on $(a, b)$, where $a, b \in \mathbb{R}$ with $a < b$. Assume that for some positive integer $n$, $F$ is $n$-times differentiable on the open interval $(a, b)$, and that $F, F', F'', \ldots, F^{(n-1)}$ all extend continuously to the closed interval $[a, b]$ (the extended functions will still be called $F, F', F'', \ldots, F^{(n-1)}$ respectively). If in addition

$$F(a) = F'(a) = \cdots = F^{(n-1)}(a) = 0, \quad \text{and} \quad F(b) = 0,$$

then there exists $c \in (a, b)$ such that

$$F^{(n)}(c) = 0.$$ 

**Proof.** The proof of this proposition follows readily from an $n$-fold application of Rolle’s theorem: Since $F(a) = F(b) = 0$, by Rolle’s theorem applied to $F$ on $[a, b]$, there exists $c_1 \in (a, b)$ such that

$$F'(c_1) = 0.$$ 

Next, since $F'(a) = F'(c_1) = 0$, by Rolle’s theorem applied to $F$ on $[a, c_1]$, there exists $c_2 \in (a, c_1)$ such that

$$F''(c_2) = 0.$$ 

Repeat, then we get $c_1, \ldots, c_n$ such that

$$a < c_n < c_{n-1} < \cdots < c_1 < b,$$

with

$$F^{(k)}(c_k) = 0 \quad \text{for } k = 1, 2, \ldots, n.$$
In particular, setting \( c = c_n \), we have \( c \in (a, b) \), and 
\[
F^{(n)}(c) = 0.
\]

\( \square \)

**First proof of Theorem 1.** Now we apply the proposition to prove Theorem 1. The key is to construct a degree \( n \) polynomial, that allows us to reduce to the case in Proposition 2. The fact that such polynomial exists follows by a dimension counting argument in linear algebra. But we will need the explicit expression of the polynomial, so let’s construct the polynomial explicitly.

Indeed, let \( f \) be as in Theorem 1. Let 
\[
P(x) = \sum_{k=0}^{n} a_k (x-a)^k.
\]
(This is a convenient form of expressing a polynomial of degree \( k \), since we will need to compute high order derivatives of this polynomial at the point \( a \).) We will find coefficients \( a_0, a_1, \ldots, a_n \), such that \( F(x) := f(x) - P(x) \) satisfies the conditions of Proposition 2. Indeed, for \( k = 0, 1, \ldots, n - 1 \), we have
\[
F^{(k)}(a) = f^{(k)}(a) - k! a_k,
\]
so in order for \( F(a) = F'(a) = \cdots = F^{(n-1)}(a) = 0 \), it suffices to set 
\[
a_k = \frac{f^{(k)}(a)}{k!} \quad \text{for} \quad k = 0, 1, \ldots, n - 1.
\]

It remains then to determine \( a_n \). But this is determined by the equation \( F(b) = 0 \): indeed 
\[
F(b) = f(b) - \sum_{k=0}^{n} a_k (b-a)^k = f(b) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (b-a)^k - a_n (b-a)^n,
\]
so setting \( F(b) = 0 \), we get 
\[
a_n = \frac{1}{(b-a)^n} \left( f(b) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (b-a)^k \right).
\]

Now we have found a polynomial \( P \) such that \( F(x) := f(x) - P(x) \) satisfies the conditions of Proposition 2. Hence there exists \( c \in (a, b) \) such that \( F^{(n)}(c) = 0 \). But 
\[
F^{(n)}(c) = f^{(n)}(c) - P^{(n)}(c) = f^{(n)}(c) - n! a_n = f^{(n)}(c) - \frac{n!}{(b-a)^n} \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (b-a)^k.
\]
Since \( F^{(n)}(c) = 0 \), it follows that 
\[
0 = f^{(n)}(c) - \frac{n!}{(b-a)^n} \left( f(b) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (b-a)^k \right),
\]
i.e.
\[
f(b) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{f^{(n)}(c)}{n!} (b-a)^n
\]
as desired. This completes the proof of Taylor’s theorem. 

\( \square \)

We can also give a second proof, based on the Cauchy mean-value theorem.
Second proof of Theorem 1. Let $f$ be as in Theorem 1, and
$$F(x) = f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (x-a)^k.$$ 

Let also
$$G(x) = (x-a)^n.$$ 

Then both $F$ and $G$ vanishes to order $(n-1)$ at $a$, in the sense that $F, F', F'', \ldots, F^{(n-1)}$ and $G, G', G'', \ldots, G^{(n-1)}$ all extends continuously to $[a,b]$, and the extended functions satisfy
$$F(a) = F'(a) = \cdots = F^{(n-1)}(a) = 0,$$
$$G(a) = G'(a) = \cdots = G^{(n-1)}(a) = 0.$$ 

Note also that $G', G'', \ldots, G^{(n)}$ all never vanishes on $(a,b)$. Hence we may apply Cauchy’s mean-value theorem $n$ times: the first time we obtain
$$F(b) = F'(c_1) = \frac{F(b) - F(a)}{G(b) - G(a)} \cdot \frac{G'(c_1)}{G'(c_1)}$$
for some $c_1 \in (a,b)$. Next we can repeat this argument, on the interval $[a,c_1]$ instead of $[a,b]$: we then obtain
$$F''(c_1) = \frac{F''(c_1) - F'(a)}{G''(c_1) - G'(a)} \cdot \frac{G''(c_1)}{G''(c_1)}$$
for some $c_2 \in (a,c_1)$. Repeating, we obtain $c_1, \ldots, c_n$ such that
$$a < c_n < c_{n-1} < \cdots < c_1 < b,$$
with
$$F(b) = F'(c_1) = F''(c_2) = \cdots = F^{(n)}(c_n).$$

In particular, setting $c = c_n$, we have $c \in (a,b)$, and
$$\frac{F(b)}{G(b)} = \frac{F^{(n)}(c)}{G^{(n)}(c)}.$$ 

This is equivalent to saying that
$$f(b) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (b-a)^k = \frac{f^{(n)}(c)}{n!} (b-a)^n,$$
which upon rearranging yields
$$f(b) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{f^{(n)}(c)}{n!} (b-a)^n$$
as desired. This completes the second proof of Taylor’s theorem. □