Proof of Taylor's Theorem

Comments on notation: Suppose $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ is a multi-index. The *length* of α is $|\alpha| = \alpha_1 + \cdots + \alpha_n$, and $\alpha!$ is defined to be $\alpha_1! \cdots \alpha_n!$. For $\mathbf{h} \in \mathbb{R}^n$, $\mathbf{h}^\alpha$ is the monomial $h_1^{\alpha_1} \cdots h_n^{\alpha_n}$. (By calling $\mathbf{h}^\alpha$ a "monomial", we mean in particular that $\alpha_i = 0$ implies $h_i^{\alpha_i} = 1$, even if $h_i = 0$.)

If $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ is a C^k-function and $|\alpha| \leq k$, then we use $\partial^\alpha f$ to denote the (mixed) partial derivative

$$\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f.$$

We also define $\partial^{(0,\ldots,0)} f = f$.

For notational convenience, we make use of the fact that the operators $\partial/\partial x_i$ can be used like monomials (as long as we put them to the left of the function they're operating on); e.g.,

$$\left(\frac{\partial}{\partial x_i} + \frac{\partial}{\partial x_j} \right) f = \frac{\partial}{\partial x_i} f + \frac{\partial}{\partial x_j} f.$$

Taylor's theorem. Let U be an open subset of $\mathbb{R}^n$ and let $f \in C^k = C^k(U, \mathbb{R})$. Let $\mathbf{x} \in U$, and let $\mathbf{h} \in \mathbb{R}^n$ be any vector such that $\mathbf{x} + t\mathbf{h} \in U$ for all $t \in [0, 1]$.

(i) f satisfies the Taylor formula with integral remainder term:

$$f(\mathbf{x} + \mathbf{h}) = \sum_{|\alpha| < k} \partial^\alpha f(\mathbf{x}) \frac{\mathbf{h}^\alpha}{\alpha!} + k \sum_{|\alpha| = k} \left(\int_0^1 (1-t)^{k-1} \partial^\alpha f(\mathbf{x} + t\mathbf{h}) \, \mathrm{d}t \right) \frac{\mathbf{h}^\alpha}{\alpha!}$$

Proof. Since $C^1 \supset C^2 \supset C^3 \supset \cdots$, we use induction on k.

For $f \in C^1$, set $g(t) = f(\mathbf{x} + t\mathbf{h})$. Then the fundamental theorem of calculus gives

$$g(1) = g(0) + \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} g(t) \, \mathrm{d}t.$$

By the chain rule, $\frac{\mathrm{d}}{\mathrm{d}t} g(t) = \left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n} \right) f(\mathbf{x} + t\mathbf{h})$. We conclude that the formula is true, i.e.,

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \sum_{i=1}^n \left(\int_0^1 \frac{\partial}{\partial x_i} f(\mathbf{x} + t\mathbf{h}) \, \mathrm{d}t \right) h_i.$$

Now assume $f \in C^{k+1}$ and that the formula is true for f in C^k. Integration by parts gives

$$\begin{aligned}
&\int_0^1 (1-t)^{k-1} \partial^\alpha f(\mathbf{x} + t\mathbf{h}) \, \mathrm{d}t \\
&= -\frac{(1-t)^k}{k} \partial^\alpha f(\mathbf{x} + t\mathbf{h}) \Big|_{t=0}^1 + \int_0^1 \frac{(1-t)^k}{k} \frac{\mathrm{d}}{\mathrm{d}t} \partial^\alpha f(\mathbf{x} + t\mathbf{h}) \, \mathrm{d}t \\
&= \frac{1}{k} \partial^\alpha f(\mathbf{x}) + \frac{1}{k} \int_0^1 (1-t)^k \left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n} \right) \partial^\alpha f(\mathbf{x} + t\mathbf{h}) \, \mathrm{d}t.
\end{aligned}$$

It is consistent with our previous notation to let

$$(\mathbf{h}\partial)^\alpha f = \left(h_1^{\alpha_1}\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}}h_2^{\alpha_2}\frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}}\cdots h_n^{\alpha_n}\frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}\right) f = \mathbf{h}^\alpha \partial^\alpha f,$$

so we adopt this convention. We recall the multinomial formula

$$(a_1 + a_2 + \cdots + a_n)^m = \sum_{b_1+\cdots+b_n=m} \binom{m}{b_1,\ldots,b_n} a_1^{b_1}\cdots a_n^{b_n}$$

which, if we let $a = (a_1, \ldots, a_n)$ and $\beta = (b_1, \ldots, b_n)$, we can write as

$$\frac{1}{m!}(a_1 + \cdots + a_n)^m = \sum_{|\beta|=m} \frac{a^\beta}{\beta!}.$$

Using the multinomial formula and the fact that mixed partials of order $\leq k+1$ commute for functions in C^{k+1}, we have

$$\begin{aligned}
&\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right)\sum_{|\alpha|=k}\frac{\mathbf{h}^\alpha}{\alpha!}\partial^\alpha f(\mathbf{x}+t\mathbf{h}) \\
&= \left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right)\sum_{|\alpha|=k}\frac{(\mathbf{h}\partial)^\alpha}{\alpha!} f(\mathbf{x}+t\mathbf{h}) \\
&= \frac{1}{k!}\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right)\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right)^k f(\mathbf{x}+t\mathbf{h}) \\
&= \frac{k+1}{(k+1)!}\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right)^{k+1} f(\mathbf{x}+t\mathbf{h}) \\
&= (k+1)\sum_{|\alpha|=k+1}\frac{(\mathbf{h}\partial)^\alpha}{\alpha!} f(\mathbf{x}+t\mathbf{h}) = (k+1)\sum_{|\alpha|=k+1}\frac{\mathbf{h}^\alpha}{\alpha!}\partial^\alpha f(\mathbf{x}+t\mathbf{h}).
\end{aligned}$$

Now we apply the gathered equations to the formula in the case $f \in C^k$, and get

$$\begin{aligned}
&f(\mathbf{x}+\mathbf{h}) \\
&= \sum_{|\alpha|<k}\partial^\alpha f(\mathbf{x})\frac{\mathbf{h}^\alpha}{\alpha!} + k\sum_{|\alpha|=k}\left(\int_0^1 (1-t)^{k-1}\partial^\alpha f(\mathbf{x}+t\mathbf{h})\,\mathrm{d}t\right)\frac{\mathbf{h}^\alpha}{\alpha!} \\
&= \sum_{|\alpha|<k}\partial^\alpha f(\mathbf{x})\frac{\mathbf{h}^\alpha}{\alpha!} + k\sum_{|\alpha|=k}\left(\frac{1}{k}\,\partial^\alpha f(\mathbf{x}) + \frac{1}{k}\int_0^1 (1-t)^k\frac{\mathrm{d}}{\mathrm{d}t}\partial^\alpha f(\mathbf{x}+t\mathbf{h})\,\mathrm{d}t\right)\frac{\mathbf{h}^\alpha}{\alpha!} \\
&= \sum_{|\alpha|\leq k}\partial^\alpha f(\mathbf{x})\frac{\mathbf{h}^\alpha}{\alpha!} + (k+1)\sum_{|\alpha|=k+1}\left(\int_0^1 (1-t)^k\partial^\alpha f(\mathbf{x}+t\mathbf{h})\,\mathrm{d}t\right)\frac{\mathbf{h}^\alpha}{\alpha!}
\end{aligned}$$

which is what we wanted. (Note that we used the fact that integrals commute with finite sums.) $\square$

(ii) The Taylor formula in (i) implies that

$$f(\mathbf{x}+\mathbf{h}) = \sum_{|\alpha|\le k} \partial^\alpha f(\mathbf{x}) \frac{\mathbf{h}^\alpha}{\alpha!} + o(|\mathbf{h}|^k)$$

as $\mathbf{h} \to \mathbf{0}$.

Proof. The above equation means

$$f(\mathbf{x}+\mathbf{h}) - \sum_{|\alpha|\le k} \partial^\alpha f(\mathbf{x}) \frac{\mathbf{h}^\alpha}{\alpha!} = o(|\mathbf{h}|^k) \qquad \text{as } \mathbf{h} \to \mathbf{0},$$

i.e., that the ratio of the left-hand side to $|\mathbf{h}|^k$ goes to zero as $\mathbf{h} \to \mathbf{0}$. Using the Taylor formula from (i), this means we need to show that

$$\lim_{\mathbf{h}\to\mathbf{0}} \sum_{|\alpha|=k} \frac{\mathbf{h}^\alpha}{|\mathbf{h}|^k \alpha!} \left[k \left(\int_0^1 (1-t)^{k-1} \partial^\alpha f(\mathbf{x}+t\mathbf{h})\, \mathrm{d}t \right) - \partial^\alpha f(\mathbf{x}) \right] = 0.$$

It suffices to prove this limit for each term of the sum, because there are finitely many terms. Because $\alpha!$ is a constant, we can ignore it. Also, $|\mathbf{h}^\alpha| = |h_1|^{\alpha_1} \cdots |h_n|^{\alpha_n} \le h_0^k \le |\mathbf{h}|^k$, where $h_0 = \max\{|h_i| : 1 \le i \le n\}$, because $\alpha_1 + \cdots + \alpha_n = k$. Thus we only need to show that the limit of the expression in brackets is zero. Set $F = \partial^\alpha f$. Then F is continuous on U, because $f \in C^k$. The expression in brackets becomes

$$k \left(\int_0^1 (1-t)^{k-1} F(\mathbf{x}+t\mathbf{h})\, \mathrm{d}t \right) - F(\mathbf{x}).$$

By the continuity of F, this expression is continuous in $\mathbf{h}$ (as long as $\mathbf{h}$ satisfies $\mathbf{x}+t\mathbf{h} \in U$ for all $t \in [0,1]$). Thus to find the limit we can simply evaluate at $\mathbf{h} = \mathbf{0}$:

$$\begin{aligned} & k \left(\int_0^1 (1-t)^{k-1} F(\mathbf{x}+t\mathbf{0})\, \mathrm{d}t \right) - F(\mathbf{x}) \\ &= F(\mathbf{x}) \left(\int_0^1 k(1-t)^{k-1}\, \mathrm{d}t \right) - F(\mathbf{x}) \\ &= F(\mathbf{x}) - F(\mathbf{x}) = 0. \end{aligned}$$

This proves the desired formula. $\square$