Solutions Manual to Walter Rudin's *Principles of Mathematical Analysis*

Roger Cooke, University of Vermont

Chapter 5

Differentiation

Exercise 5.1 Let f be defined for all real x, and suppose that

$$|f(x) - f(y)| \leq (x-y)^2$$

for all real x and y. Prove that f is constant.

Solution. Dividing by $x - y$, and letting $x \to y$, we find that $f'(y) = 0$ for all y. Hence f is constant.

Exercise 5.2 Suppose $f'(x) > 0$ in (a, b). Prove that f is strictly increasing in (a, b), and let g be its inverse function. Prove that g is differentiable, and that

$$g'(f(x)) = \frac{1}{f'(x)} \qquad (a < x < b).$$

Solution. For any c, d with $a < c < d < b$ there exists a point $p \in (c, d)$ such that $f(d) - f(c) = f'(p)(d-c) > 0$. Hence $f(c) < f(d)$.

We know from Theorem 4.17 that the inverse function g is continuous. (Its restriction to each closed subinterval $[c, d]$ is continuous, and that is sufficient.) Now observe that if $f(x) = y$ and $f(x+h) = y + k$, we have

$$\frac{g(y+k) - g(y)}{k} - \frac{1}{f'(x)} = \frac{1}{\frac{f(x+h)-f(x)}{h}} - \frac{1}{f'(x)}.$$

Since we know $\lim \frac{1}{\varphi(t)} = \frac{1}{\lim \varphi(t)}$ provided $\lim \varphi(t) \neq 0$, it follows that for any $\varepsilon > 0$ there exists $\eta > 0$ such that

$$\left| \frac{1}{\frac{f(x+h)-f(x)}{h}} - \frac{1}{f'(x)} \right| < \varepsilon$$

if $0 < |h| < \eta$. Since $h = g(y+k) - g(y)$, there exists $\delta > 0$ such that $0 < |h| < \eta$ if $0 < |k| < \delta$. The proof is now complete.

Exercise 5.3 Suppose g is a real function on R^1 with bounded derivative (say $|g'| \leq M$). Fix $\varepsilon > 0$, and define $f(x) = x + \varepsilon g(x)$. Prove that f is one-to-one if ε is small enough. (A set of admissible values of ε can be determined which depends only on M.)

Solution. If $0 < \varepsilon < \dfrac{1}{M}$, we certainly have

$$f'(x) \geq 1 - \varepsilon M > 0,$$

and this implies that $f(x)$ is one-to-one, by the preceding problem.

Exercise 5.4 If

$$C_0 + \frac{C_1}{2} + \cdots + \frac{C_{n-1}}{n} + \frac{C_n}{n+1} = 0,$$

where $C_0, \ldots, C_n$ are real constants, prove that the equation

$$C_0 + C_1 x + \cdots + C_{n-1}x^{n-1} + C_n x^n = 0$$

has at least one real root between 0 and 1.

Solution. Consider the polynomial

$$p(x) = C_0 x + \frac{C_1}{2}x^2 + \cdots + \frac{C_{n-1}}{n}x^n + \frac{C_n}{n+1}x^{n+1},$$

whose derivative is

$$p'(x) = C_0 + C_1 x + \cdots + C_{n-1}x^{n-1} + C_n x^n.$$

It is obvious that $p(0) = 0$, and the hypothesis of the problem is that $p(1) = 0$. Hence Rolle's theorem implies that $p'(x) = 0$ for some x between 0 and 1.

Exercise 5.5 Suppose f is defined and differentiable for every $x > 0$, and $f'(x) \to 0$ as $x \to +\infty$. Put $g(x) = f(x+1) - f(x)$. Prove that $g(x) \to 0$ as $x \to +\infty$.

Solution. Let $\varepsilon > 0$. Choose x_0 such that $|f'(x)| < \varepsilon$ if $x > x_0$. Then for any $x \geq x_0$ there exists $x_1 \in (x, x+1)$ such that

$$f(x+1) - f(x) = f'(x_1).$$

Since $|f'(x_1)| < \varepsilon$, it follows that $|f(x+1) - f(x)| < \varepsilon$, as required.

Exercise 5.6 Suppose

(a) f is continuous for $x \geq 0$,

(b) $f'(x)$ exists for $x > 0$,

(c) $f(0) = 0$,

(d) f' is monotonically increasing.

Put

$$g(x) = \frac{f(x)}{x} \quad (x > 0)$$

and prove that g is monotonically increasing.

Solution. By the mean-value theorem

$$f(x) = f(x) - f(0) = f'(c)x,$$

for some $c \in (0, x)$. Since f' is monotonically increasing, this result implies that $f(x) < xf'(x)$. It therefore follows that

$$g'(x) = \frac{xf'(x) - f(x)}{x^2} > 0,$$

so that g is also monotonically increasing.

Exercise 5.7 Suppose $f'(x)$ and $g'(x)$ exist, $g'(x) \neq 0$, and $f(x) = g(x) = 0$. Prove that

$$\lim_{t \to x} \frac{f(t)}{g(t)} = \frac{f'(x)}{g'(x)}.$$

(This holds also for complex functions.)

Solution. Since $f(x) = g(x) = 0$, we have

$$\begin{aligned} \lim_{t \to x} \frac{f(t)}{g(t)} &= \lim_{t \to x} \frac{\dfrac{f(t) - f(x)}{t - x}}{\dfrac{g(t) - g(x)}{t - x}} \\ &= \frac{\lim\limits_{t \to x} \dfrac{f(t) - f(x)}{t - x}}{\lim\limits_{t \to x} \dfrac{g(t) - g(x)}{t - x}} \\ &= \frac{f'(x)}{g'(x)}. \end{aligned}$$

Exercise 5.8 Suppose f' is continuous on $[a,b]$ and $\varepsilon > 0$. Prove that there exists $\delta > 0$ such that

$$\left|\frac{f(t)-f(x)}{t-x} - f'(x)\right| < \varepsilon$$

whenever $0 < |t-x| < \delta$, $a \le x \le b$, $a \le t \le b$. (This could be expressed by saying that f is *uniformly differentiable* on $[a,b]$ if f' is continuous on $[a,b]$.) Does this hold for vector-valued functions too?

Solution. Let δ be such that $|f'(x)-f'(u)| < \varepsilon$ for all $x, u \in [a,b]$ with $|x-u| < \delta$. Then if $0 < |t-x| < \delta$ there exists u between t and x such that

$$\frac{f(t)-f(x)}{t-x} = f'(u),$$

and hence, since $|u-x| < \delta$,

$$\left|\frac{f(t)-f(x)}{t-x} - f'(x)\right| = |f'(u) - f'(x)| < \varepsilon.$$

Since this result holds for each component of a vector-valued function $\mathbf{f}(x)$, it must hold also for $\mathbf{f}$.

Exercise 5.9 Let f be a continuous real function on R^1, of which it is known that $f'(x)$ exists for all $x \neq 0$ and that $f'(x) \to 3$ as $x \to 0$. Does it follow that $f'(0)$ exists?

Solution. Yes. By L'Hospital's rule

$$\lim_{t\to 0} \frac{f(t)-f(0)}{t} = \lim_{t \to 0} f'(t) = 3,$$

and this by definition means that $f'(0) = 3$.

Exercise 5.10 Suppose f and g are complex differentiable functions on $(0,1)$, $f(x) \to 0$, $g(x) \to 0$, $f'(x) \to A$, $g'(x) \to B$ as $x \to 0$, where A and B are complex numbers, $B \neq 0$. Prove that

$$\lim_{x \to 0} \frac{f(x)}{g(x)} = \frac{A}{B}.$$

Compare with Example 5.18. *Hint:*

$$\frac{f(x)}{g(x)} = \left\{\frac{f(x)}{x} - A\right\} \cdot \frac{x}{g(x)} + A \cdot \frac{x}{g(x)}.$$

Apply Theorem 5.13 to the real and imaginary parts of $f(x)/x$ and $g(x)/x$.

Solution. We can make f and g continuous on $[0,1)$ by simply defining $f(0) = 0 = g(0)$. Then Exercise 9 applied to the real and imaginary parts of f and g show that $f'(0) = A$ and $g'(0) = B$. (These are one-sided derivatives, since f and g are not defined for negative values of x; however, we could extend them as odd functions, since both are 0 at 0). We could then apply Exercise 7, whose proof does not use anything but the definition of the derivative and some general facts about limits. In this way we get the result without resorting to the combinatorial trick referred to in the hint. This result shows that many of the facts ordinarily proved for real functions by use of the mean-value theorem and L'Hospital's rule remain true for complex-valued functions, even though, as Example 5.18 shows, these theorems are not true for complex-valued functions.

Exercise 5.11 Suppose f is defined in a neighborhood of x, and suppose $f''(x)$ exists. Show that

$$\lim_{h\to 0} \frac{f(x+h)+f(x-h)-2f(x)}{h^2} = f''(x).$$

Solution. For a real-valued function this is a routine application of L'Hospital's rule:

$$\begin{aligned}
\lim_{h\to 0} \frac{f(x+h)+f(x-h)-2f(x)}{h^2} &= \lim_{h\to 0} \frac{f'(x+h)-f'(x-h)}{2h} \\
&= \frac{1}{2}\lim_{h\to 0} \frac{f'(x+h)-f'(x)}{h} + \\
&\qquad + \frac{f'(x)-f'(x-h)}{h} \\
&= f''(x).
\end{aligned}$$

For complex-valued functions the result follows from separate consideration of real and imaginary parts.

The limit will be zero at $x = 0$ for any odd function f whatsoever, even if the function is not continuous. For example we could take $f(x) = \mathrm{sgn}\,(x)$, which is $+1$ for $x > 0$, 0 for $x = 0$, and -1 for $x < 0$

Exercise 5.12 If $f(x) = |x|^3$, compute $f'(x)$, $f''(x)$ for all real x, and show that $f^{(3)}(0)$ does not exist.

Solution. For $x > 0$ we have $f'(x) = 3x^2$, $f''(x) = 6x$, and for $x < 0$ $f'(x) = -6x^2$, $f''(x) = -6x$, i.e., $f'(x) = 3x|x|$, and $f''(x) = 6|x|$ for $x \neq 0$. By Exercise 9, it therefore follows that $f'(0)$ exists and equals 0, and then another application of Exercise 9 shows that $f''(0)$ also exists and equals 0. However

$$\frac{f''(x)-f''(0)}{x} = 6\mathrm{sgn}\,(x),$$

which has no limit at 0. Hence $f^{(3)}(0)$ does not exist.

Exercise 5.13 Suppose a and c are real numbers, $c > 0$, and f is defined on $[-1, 1]$ by

$$f(x) = \begin{cases} x^a \sin(x^{-c}) & (\text{if } x \neq 0), \\ 0 & (\text{if } x = 0). \end{cases}$$

Prove the following statements:

(a) f is continuous if and only if $a > 0$.

(b) $f'(0)$ exists if and only if $a > 1$.

(c) f' is bounded if and only if $a \geq 1 + c$.

(d) f' is continuous if and onlyl if $a > 1 + c$.

(e) $f''(0)$ exists if and only if $a > 2 + c$.

(f) f'' is bounded if and only if $a \geq 2 + 2c$.

(g) f'' is continuous if and only if $a > 2 + 2c$.

Solution. We remark editorially that there are two difficulties with this problem. One is that we haven't yet introduced the function sin. To overcome that problem we can rely on our intuitive notion or use the Taylor series if we have to. The second problem is more serious, however: What do x^a and x^{-c} mean when $x < 0$? In general these will be complex-valued functions. It might be better to use absolute values in both cases. Thus we shall amend the problem by defining $f(x) = |x|^a \sin(|x|^{-c})$ when $x \neq 0$.

(a) Since f is infinitely differentiable except at $x = 0$, the only question of continuity is at $x = 0$. Let $t_n = 2\pi(n + \frac{1}{8})$, $x_n = t_n^{-\frac{1}{c}}$ and $y_n = \dfrac{1}{\sqrt{2}} t_n^{-\frac{a}{c}}$. Notice that $f(x_n) = y_n$ and that y_n tends to $\dfrac{1}{\sqrt{2}}$ if $a = 0$ and to $+\infty$ if $a < 0$. Hence the function cannot be continuous if $a \leq 0$. On the other hand, we have

$$|f(x) - f(0)| = |f(x)| \leq |x|^a,$$

so that if $a > 0$ and ε is given, we can choose $\delta = \varepsilon^{\frac{1}{a}}$, and then $|x - 0| < \delta$ implies $|f(x) - f(0)| < \varepsilon$, i.e., $f(x)$ is continuous at $x = 0$.

(b) If $f'(0)$ exists, then f is continuous at 0, so that $a > 0$. Notice that

$$\frac{f(x_n) - f(0)}{x_n} = \frac{y_n}{x_n} = \frac{1}{\sqrt{2}} t_n^{\frac{1-a}{c}}.$$

which tends to $\dfrac{1}{\sqrt{2}}$ if $a = 1$ and to $+\infty$ if $0 < a < 1$. Hence $f'(0)$ does not exist if $a \leq 1$. On the other hand if $a > 1$ we have

$$0 \leq \frac{f(x) - f(0)}{x} < |x|^{a-1} \to 0,$$

and so $f'(0) = 0$.

(c) For $x \neq 0$ we have

$$f'(x) = \operatorname{sgn}(x)|x|^{a-1}\left[a \sin(|x|^{-c}) - c|x|^{-c} \cos(|x|^{-c})\right].$$

Hence $f'(x_n) = \frac{1}{\sqrt{2}}\big[ax_n^{a-1} - cx_n^{-c+a-1}\big]$, which tends to $-\infty$ if $a < 1 + c$. On the other hand we have

$$|f'(x)| \le |a||x|^{a-1} + c|x|^{a-1-c},$$

which is certainly bounded on $[-1, 1]$ if $a \ge 1 + c$.

(d) If f' is continuous, it is bounded, and so $a \ge 1 + c$. However if $a = 1 + c$, then

$$f'(x_n) = \frac{1}{\sqrt{2}}\big[(1+c)t_n^{-1} - c\big]$$

which tends to $-\frac{c}{\sqrt{2}}$ as $n \to \infty$, while $x_n \to 0$. Hence f' is not continuous at 0 unless $a > 1 + c$. If $a > 1 + c$, the inequality

$$|f'(x)| \le |a||x|^{a-1} + c|x|^{a-1-c},$$

implies that $f(x) \to 0$ as $x \to 0$, and so f' is continuous.

(e) If $f''(0)$ exists, then f' must be continuous at 0, and so $a \ge 1 + c$. Now for $x \ne 0$

$$\frac{f'(x) - f'(0)}{x} = \operatorname{sgn}(x)\big[a|x|^{a-2}\sin(|x|^{-c}) - c|x|^{a-c-2}\cos(|x|^{-c}).$$

Taking $x = x_n$, we find that this difference quotient equals

$$\frac{1}{\sqrt{2}}\big[at_n^{\frac{2-a}{c}} - ct_n^{\frac{c+2-a}{c}}\big],$$

which tends to $\frac{1}{\sqrt{2}}$ if $a = c + 2$ and to $-\infty$ if $a < c + 2$. Hence $f''(0)$ exists only if $a > c + 2$.

On the other hand, if $a > c + 2$, we have the inequality

$$\left|\frac{f'(x) - f'(0)}{x}\right| \le a|x|^{a-2} + c|x|^{a-c-2},$$

from which it follows immediately that $f''(0) = 0$.

(f) For $x \ne 0$ we have

$$\begin{aligned} f''(x) = \operatorname{sgn}(x)\big[a(a-1)|x|^{a-2} - c^2|x|^{a-2c-2}\big]\sin(|x|^{-c}] \\ - c(2a - c - 1)|x|^{a-c-1}\cos(|x|^{-c}]. \end{aligned}$$

In particular

$$f''(x_n) = \frac{1}{\sqrt{2}}\big[a(a-1)t_n^{\frac{2-a}{c}} - c^2 t_n^{\frac{2+2c-a}{c}} - c(2a - c - 1)t_n^{\frac{c+1-a}{c}}\big],$$

which tends to $-\infty$ if $a < 2 + 2c$. On the other hand, we have the inequality

$$|f''(x)| \le |a||a-1||x|^{a-2} + c^2|x|^{a-2c-2} + c|2a - c - 1||x|^{a-c-1},$$

and the right-hand side is certainly bounded if $a \geq 2+2c$.

(g) If f'' is continuous, then it is bounded, and hence $a \geq 2+2c$. If $a = 2+2c$, we have

$$f''(x_n) = \frac{1}{\sqrt{2}}[(2c+2)(2c+1)t_n^{-2} - c^2 - c(3+3c)t_n^{\frac{-c-1}{c}}],$$

which tends to $-\frac{c^2}{\sqrt{2}}$, so that f'' is not continuous at 0. On the other hand, if $a > 2+2c$, the inequality

$$|f''(x)| \leq |a||a-1||x|^{a-2} + c^2|x|^{a-2c-2} + c|2a-c-1||x|^{a-c-1},$$

shows that $f''(x) \to 0$ as $x \to 0$, and hence f'' is continuous.

Exercise 5.14 Let f be a differentiable real function defined in (a,b). Prove that f is convex if and only if f' is monotonically increasing. Assume next that $f''(x)$ exists for every $x \in (a,b)$, and prove that f is convex if and only if $f''(x) \geq 0$ for all $x \in (a,b)$.

Suppose first that f' is montonically increasing, and that $x < y$. We wish to show that if $0 < \lambda < 1$, then

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).$$

Letting $z = \lambda x + (1-\lambda)y$, we have $\lambda = \frac{y-z}{y-x}$, $1-\lambda = \frac{z-x}{y-x}$, and $x < z < y$. Now the required inequality can be written

$$(1-\lambda)[f(y)-f(z)] \geq \lambda[f(z)-f(x)],$$

which, when we insert the values of λ and $1-\lambda$, and multiply by the positive number $\frac{y-x}{(z-x)(y-z)}$, becomes

$$\frac{f(y)-f(z)}{y-z} \geq \frac{f(z)-f(x)}{z-x}.$$

Since the left-hand side is $f'(d)$ for some $d \in (z,y)$, the right-hand side is $f'(c)$ for some $c \in (x,z)$, and f' is nondecreasing, we have the required inequality.

By Exercise 23 of Chapter 4 we know that if f is convex on (a,b) and $a < c < d < p < q < b$, then

$$\frac{f(d)-f(c)}{d-c} \leq \frac{f(p)-f(d)}{p-d} \leq \frac{f(q)-f(p)}{q-p}.$$

Hence, if f' exists, letting $d \to c$ and $q \to p$, we find

$$f'(c) \leq f'(p),$$

so that f' is nondecreasing.

Finally if f'' exists, we know that f' is nondecreasing if and only if $f''(x) \geq 0$ for all $x \in (a,b)$. Hence f is convex if and only if $f''(x) \geq 0$ for all $x \in (a,b)$.

Exercise 5.15 Suppose $a \in R^1$, f is a twice-differentiable real function on (a, ∞), and M_0, M_1, M_2 are the least upper bounds of $|f(x)|$, $|f'(x)|$, $|f''(x)|$, respectively, on (a, ∞). Prove that

$$M_1^2 \leq 4M_0M_2.$$

Hint: If $h > 0$, Taylor's theorem shows that

$$f'(x) = \frac{1}{2h}[f(x+2h) - f(x)] - hf''(\xi)$$

for some $\xi \in (x, x+2h)$. Hence

$$|f(x)| \leq hM_2 + \frac{M_0}{h}.$$

To show that $M_1^2 = 4M_0M_2$ can actually happen, take $a = -1$, define

$$f(x) = \begin{cases} 2x^2 - 1, & (-1 < x < 0), \\ \dfrac{x^2-1}{x^2+1}, & (0 \leq x < \infty), \end{cases}$$

and show that $M_0 = 1$, $M_1 = 4$, $M_2 = 4$.

Does $M_1^2 \leq 4M_1M_2$ hold for vector-valued functions too?

Solution. The inequality is obvious if $M_0 = +\infty$ or $M_2 = +\infty$, so we shall assume that M_0 and M_2 are both finite. We need to show that

$$|f'(x)| \leq 2\sqrt{M_0M_2}$$

for all $x > a$. We note that this is obvious if $M_2 = 0$, since in that case $f'(x)$ is constant, $f(x)$ is a linear function, and the only bounded linear function is a constant, whose derivative is zero. Hence we shall assume from now on that $0 < M_2 < +\infty$ and $0 < M_0 < +\infty$.

Following the hint, we need only choose $h = \sqrt{\dfrac{M_0}{M_2}}$, and we obtain

$$|f'(x)| \leq 2\sqrt{M_0M_2},$$

which is precisely the desired inequality.

The case of equality follows, since the example proposed satisfies

$$f(x) = 1 - \frac{2}{x^2+1}$$

for $x \geq 0$. We see easily that $|f(x)| \leq 1$ for all $x > -1$. Now $f'(x) = \dfrac{4x}{(x^2+1)^2}$ for $x > 0$ and $f'(x) = 4x$ for $x < 0$. It thus follows from Exercise 9 above that $f'(0) = 0$, and that $f'(x)$ is continuous. Likewise $f''(x) = 4$ for $x < 0$

and $f''(x) = \dfrac{4-4x^2}{(x^2+1)^3} = -4\dfrac{x^2-1}{(x^2+1)^3}$. This shows that $|f''(x)| < 4$ for $x > 0$ and also that $\lim_{x\to 0} f''(x) = 4$. Hence Exercise 9 again implies that $f''(x)$ is continuous and $f''(0) = 4$.

On n-dimensional space let $\mathbf{f}(x) = (f_1(x), \ldots, f_n(x))$, $M_0 = \sup|\mathbf{f}(x)|$, $M_1 = \sup|\mathbf{f}'(x)|$, and $M_2 = \sup|\mathbf{f}''(x)|$. Just as in the numerical case, there is nothing to prove if $M_2 = 0$ or $M_0 = +\infty$ or $M_2 = +\infty$, and so we assume $0 < M_0 < +\infty$ and $0 < M_2 < \infty$. Let a be any positive number less than M_1, let x_0 be such that $|\mathbf{f}'(x_0)| > a$, and let $\mathbf{u} = \dfrac{1}{|\mathbf{f}'(x_0)|}\mathbf{f}'(x_0)$. Consider the real-valued function $\varphi(x) = \mathbf{u}\cdot\mathbf{f}(x)$. Let N_0, N_1, and N_2 be the suprema of $|\varphi(x)|$, $|\varphi'(x)|$, and $|\varphi''(x)|$ respectively. By the Schwarz inequality we have (since $|\mathbf{u}| = 1$) $N_0 \le M_0$ and $N_2 \le M_2$, while $N_1 \ge \varphi(x_0) = |\mathbf{f}'(x_0)| > a$. We therefore have $a^2 < 4N_0N_2 \le 4M_0M_2$. Since a was any positive number less than M_1, we have $M_1^2 \le 4M_0M_2$, i.e., the result holds also for vector-valued functions.

Equality can hold on any R^n, as we see by taking $\mathbf{f}(x) = (f(x), 0, \ldots, 0)$ or $\mathbf{f}(x) = (f(x), f(x), \ldots, f(x))$, where $f(x)$ is a real-valued function for which equality holds.

Exercise 5.16 Suppose f is twice-differentiable on $(0,\infty)$, f'' is bounded on $(0,\infty)$, and $f(x) \to 0$ as $x \to \infty$. Prove that $f'(x) \to 0$ as $x \to \infty$.

Solution. We shall prove an even stronger statement. If $f(x) \to L$ as $x \to \infty$ and $f'(x)$ is uniformly continuous on $(0,\infty)$, then $f'(x) \to 0$ as $x \to \infty$.

For, if not, let $x_n \to \infty$ be a sequence such that $f(x_n) \ge \varepsilon > 0$ for all n. (We can assume $f(x_n)$ is positive by replacing f with $-f$ if necessary.) Let δ be such that $|f'(x) - f'(y)| < \frac{\varepsilon}{2}$ if $|x-y| < \delta$. We then have $f'(y) > \frac{\varepsilon}{2}$ if $|y - x_n| < \delta$, and so

$$|f(x_n+\delta) - f(x_n-\delta)| \ge 2\delta\cdot\frac{\varepsilon}{2} = \delta\varepsilon.$$

But, since $\delta\varepsilon > 0$, there exists X such that

$$|f(x) - L| < \frac{1}{2}\delta\varepsilon$$

for all $x > X$. Hence for all large n we have

$$|f(x_n+\delta) - f(x_n-\delta)| \le |f(x_n+\delta) - L| + |L - f(x_n-\delta)| < \delta\varepsilon,$$

and we have reached a contradiction.

The problem follows from this result, since if f'' is bounded, say $|f''(x)| \le M$, then $|f'(x) - f'(y)| \le M|x-y|$, and f' is certainly uniformly continuous.

Exercise 5.17 Suppose f is a real, three times differentiable function on $[-1, 1]$, such that

$$f(-1) = 0, \quad f(0) = 0, \quad f(1) = 1, \quad f'(0) = 0.$$

Prove that $f^{(3)}(x) \geq 3$ for some $x \in (-1, 1)$.

Note that equality holds for $\frac{1}{2}(x^3 + x^2)$.

Hint: Use Theorem 5.15 with $\alpha = 1$ and $\beta = \pm 1$, to to show that there are $s \in (0, 1)$ and $t \in (-1, 0)$ such that

$$f^{(3)}(s) + f^{(3)}(t) = 6.$$

Solution. Following the hint, we observe that Theorem 5.15 (Taylor's formula with remainder) implies that

$$\begin{aligned} f(1) &= f(0) + f'(0) + \frac{1}{2}f''(0) + \frac{1}{6}f^{(3)}(s) \\ f(-1) &= f(0) - f'(0) + \frac{1}{2}f''(0) - \frac{1}{6}f^{(3)}(t) \end{aligned}$$

for some $s \in (0, 1)$, $t \in (-1, 0)$. By subtracting the second equation from the first and using the given values of $f(1)$, $f(-1)$, and $f'(0)$, we obtain

$$1 = \frac{1}{6}(f^{(3)}(s) + f^{(3)}(t)),$$

which is the desired result. Note that we made no use of the hypothesis $f(0) = 0$.

Exercise 5.18 Suppose f is a real function on $[a, b]$, n is a positive integer, and $f^{(n-1)}$ exists for every $t \in [a, b]$. Let α, β, and P be as in Taylor's theorem (5.15). Define

$$Q(t) = \frac{f(t) - f(\beta)}{t - \beta}$$

for $t \in [a, b]$, $t \neq \beta$, differentiate

$$f(t) - f(\beta) = (t - \beta)Q(t)$$

$n - 1$ times at $t = \alpha$, and derive the following version of Taylor's theorem:

$$f(\beta) = P(\beta) + \frac{Q^{n-1}(\alpha)}{(n-1)!}(\beta - \alpha)^n.$$

Solution. The function $Q(t)$ is differentiable $n-1$ times except possibly at $t = \beta$, so we don't have to worry when differentiating $n - 1$ times at $t = \alpha$. It is easy to prove by induction that

$$f^{(k)}(t) = (t - \beta)Q^{(k)}(t) + kQ^{(k-1)}(t)$$

for $0 < k \leq n-1$. Hence

$$\frac{1}{k!}f^{(k)}(\alpha)(\beta-\alpha)^k = -\frac{(\beta-\alpha)^{k+1}}{k!}Q^{(k)}(\alpha) + \frac{(\beta-\alpha)^k}{(k-1)!}Q^{(k-1)}(\alpha).$$

Then, because the sum telescopes, we find

$$\sum_{k=0}^{n-1}\frac{f^{(k)}(\alpha)}{k!}(\beta-\alpha)^k = f(\beta) - \frac{Q^{n-1}(\alpha)}{(n-1)!}(\beta-\alpha)^n,$$

which can be rewritten as

$$f(\beta) = P(\beta) + \frac{Q^{(n-1)}(\alpha)}{(n-1)!}(\beta-\alpha)^n.$$

Exercise 5.19 Suppose f is defined in $(-1,1)$ and $f'(0)$ exists. Suppose $-1 < \alpha_n < \beta_n < 1$, $\alpha_n \to 0$, and $\beta_n \to 0$ as $n \to \infty$. Define the difference quotients

$$Dn = \frac{f(\beta_n) - f(\alpha_n)}{\beta_n - \alpha_n}.$$

Prove the following statements:

(*a*) If $\alpha_n < 0 < \beta_n$, then $\lim D_n = f'(0)$.

(*b*) If $0 < \alpha_n < \beta_n$ and $\beta_n/(\beta_n - \alpha_n)$ is bounded, then $\lim D_n = f'(0)$.

(*c*) if f' is continuous in $(-1,1)$, then $\lim D_n = f'(0)$.

Give an example in which f is differentiable in $(-1,1)$ (but f' is not continuous at 0) and in which α_n, β_n tends to 0 in such a way that $\lim D_n$ exists but is different from $f'(0)$.

Solution. We assume that $\alpha_n\beta_n \neq 0$ throughout, i.e., that neither α_n nor β_n is zero.

(*a*) Write

$$\begin{aligned} D_n &= \frac{f(\beta_n) - f(0)}{\beta_n - \alpha_n} + \frac{f(0) - f(\alpha_n)}{\beta_n - \alpha_n} \\ &= \frac{\beta_n}{\beta_n - \alpha_n}\frac{f(\beta_n) - f(0)}{\beta_n} + \frac{-\alpha_n}{\beta_n - \alpha_n}\frac{f(\alpha_n) - f(0)}{\alpha_n}. \end{aligned}$$

Now let $\varepsilon > 0$. Choose $\delta > 0$ such that

$$\left|\frac{f(x) - f(0)}{x} - f'(0)\right| < \varepsilon$$

if $0 < |x| < \delta$. Then choose N so that $0 < \beta_n < \delta$ and $-\delta < \alpha_n < 0$ for $n > N$. Then for all $n > N$ we have

$$\begin{aligned} |D_n - f'(0)| &\le \frac{\beta_n}{\beta_n - \alpha_n}\left|\frac{f(\beta_n) - f(0)}{\beta_n} - f'(0)\right| + \\ &\qquad + \frac{-\alpha_n}{\beta_n - \alpha_n}\left|\frac{f(\alpha_n) - f(0)}{\alpha_n} - f'(0)\right| \\ &< \frac{\beta_n}{\beta_n - \alpha_n}\varepsilon + \frac{-\alpha_n}{\beta_n - \alpha_n}\varepsilon \\ &= \varepsilon. \end{aligned}$$

(*b*) If $\dfrac{\beta_n}{\beta_n - \alpha_n} \le M$ for all n, and $0 < \alpha_n < \beta_n$, then surely $\dfrac{\alpha_n}{\beta_n - \alpha_n} < M$ for all n. Hence if $\varepsilon > 0$ is given, choose N so that

$$\left|\frac{f(x) - f(0)}{x} - f'(0)\right| < \frac{\varepsilon}{2M}$$

if $0 < |x| < \delta$. Then choose N so that $0 < \beta_n < \delta$ (hence also $0 < \alpha_n < \delta$) for $n > N$. Then for all $n > N$ we have

$$\begin{aligned} |D_n - f'(0)| &\le \frac{\beta_n}{\beta_n - \alpha_n}\left|\frac{f(\beta_n) - f(0)}{\beta_n} - f'(0)\right| + \\ &\qquad + \frac{\alpha_n}{\beta_n - \alpha_n}\left|\frac{f(\alpha_n) - f(0)}{\alpha_n} - f'(0)\right| \\ &< \frac{\beta_n}{\beta_n - \alpha_n}\frac{\varepsilon}{2M} + \frac{\alpha_n}{\beta_n - \alpha_n}\frac{\varepsilon}{2M} \\ &< \varepsilon. \end{aligned}$$

(*c*) By the mean-value theorem there exists γ_n between α_n and β_n such that $D_n = f'(\gamma_n)$. Since $\gamma_n \to 0$ and f' is continuous, it follows that $D_n \to f'(0)$.

Let $f(x)$ be *any* function such that $f'(0)$ exists but $\lim_{x\to 0} f'(x)$ does not exist. We know that $f'(x)$ does not tend to infinity as $x \to 0$, since if it did, we would have $|f'(x)| > 1 + |f'(0)|$ for all sufficiently small nonzero x, and this contradicts the intermediate-value property of derivatives. Hence there is a sequence $x_n \to 0$, $x_n \ne 0$, such that $\lim_{n\to\infty} f'(x_n) = L \ne f'(0)$. Let $\beta_n = x_n$, and let y_n be such that $0 < |y_n - x_n| < \frac{1}{2}|x_n|$ and

$$\left|\frac{f(y_n) - f(x_n)}{y_n - x_n} - f'(x_n)\right| < \frac{|L - f'(0)|}{2n}$$

It is then immediate that

$$\lim_{n\to\infty} \frac{f(y_n) - f(x_n)}{y_n - x_n} = L \ne f'(0).$$

A suitable example of such a function $f(x)$ is

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x \ne 0, \\ 0, & x = 0. \end{cases}$$

In this case we can get the counterexample in a slightly different form by taking $x_n = \frac{1}{2\pi n}$ and $y_n = \frac{1}{2\pi(n+\frac{1}{4})}$. We then have $f'(0) = 0$ and

$$\frac{f(y_n) - f(x_n)}{y_n - x_n} = \frac{2n}{\pi(n+\frac{1}{4})} \to \frac{2}{\pi}.$$

Exercise 5.20 Formulate and prove an inequality which follows from Taylor's theorem and which remains valid for vector-valued functions.

Solution. There is a variety of possibilities, of which we choose just one: *Suppose $f(x)$ has continuous derivatives up to order n on $[a, b]$. Then there exists $c \in (a, b)$ such that*

$$|f(b) - P(b)| \le \left|\frac{f^n(c)}{n!}\right|(b-a)^n.$$

To prove this assertion true for a vector-valued function $\mathbf{f}$, we merely observe that it holds for each scalar-valued function $\mathbf{u}\cdot\mathbf{f}$ if $\mathbf{u}$ is any fixed vector of length 1. It is obviously true if $|\mathbf{f}(b) - \mathbf{P}(b)| = 0$, and in all other cases it follows by taking $\mathbf{u} = \frac{1}{|\mathbf{f}(b) - \mathbf{P}(b)|}(\mathbf{f}(b) - \mathbf{P}(b))$.

Exercise 5.21 Let E be a closed subset of R^1. We saw in Exercise 22, Chap. 4, that there is a real continuous function f on R^1 whose zero set is E. Is it possible, for each closed set E, to find such an f which is differentiable on R^1, or one which is n times differentiable, or even one which has derivatives of all orders on R^1?

Solution. Yes, it is possible. The proof depends on the following lemma:

Let a and b be any real numbers with $a < b$, and let $f(x)$ be defined for all real numbers x by the formulas

$$f(x) = \begin{cases} e^{\frac{1}{(x-a)(x-b)}}, & a < x < b, \\ \\ 0, & x \le a \text{ or } x \ge b. \end{cases}$$

Then f has derivatives of all orders on R^1.

It is obvious that f has derivatives of all orders at every point except possibly a and b. To prove that derivatives exist at these points we need two sublemmas: *For each nonnegative integer n there exists a polynomial $p_n(z, w)$ such that*

$$f^{(n)}(x) = p_n\Big(\frac{1}{x-a}, \frac{1}{x-b}\Big)e^{\frac{1}{(x-a)(x-b)}}$$

for $a < x < b$.

The proof of this sublemma uses only the partial-fraction decomposition

$$\frac{1}{(x-a)(x-b)} = \frac{1}{b-a}\left[\frac{1}{x-b} - \frac{1}{x-a}\right],$$

together with the chain rule and the fact that the partial derivative of a polynomial is again a polynomial. We omit the details.

The second sublemma is stated as a formula: *For every nonnegative integer n,*

$$\lim_{x\downarrow a} \frac{e^{\frac{1}{(x-a)(x-b)}}}{(x-a)^n} = 0.$$

Its proof is a consequence of Taylor's formula. To be specific, Taylor's formula with remainder implies the following result:

For each nonnegative integer k and each positive number t

$$e^t > \frac{1}{k!}t^k.$$

This last result follows easily since there is a point $t_k \in (0,t)$ for which

$$e^t = 1 + t + \frac{t^2}{2!} + \cdots + \frac{t^{k-1}}{(k-1)!} + \frac{e^{t_k}}{k!}t^k,$$

every term in this last sum is positive, and $e^{t_k} > 1$.

We now apply this result with $k = n$ and $t = \dfrac{1}{(x-a)(b-x)}$, to obtain

$$\begin{aligned} e^{\frac{1}{(x-a)(x-b)}} &= \frac{1}{e^{\frac{1}{(x-a)(b-x)}}} \\ &< n!(b-x)^n(x-a)^n \end{aligned}$$

for all $n = 0, 1, \ldots$. In particular

$$e^{\frac{1}{(x-a)(x-b)}} < n!(b-a)^n(x-a)^n = K_n(x-a)^n.$$

Since the kth derivative of $e^{\frac{1}{(x-a)(x-b)}}$ is a polynomial in $\dfrac{1}{x-a}$ and $\dfrac{1}{x-b}$, each derivative also satisfies such an estimate. It follows from this last result that

$$\lim_{x\downarrow a} p\left(\frac{1}{z-b}, \frac{1}{z-a}\right) e^{\frac{1}{(x-a)(x-b)}} = 0$$

for any polynomial $p(z,w)$, and hence that $f^{(n)}(a) = 0$ for all n. The proof that $f^{(n)}(b) = 0$ is similar. We observe that the zero set of $f(x)$ is the complement of the open interval (a,b).

Identical reasoning shows that the function

$$f(x) = \begin{cases} e^{\frac{1}{a-x}}, & x > a, \\ 0, & x \le a, \end{cases}$$

has derivatives of all orders, and its zero set is the complement of the semi-infinite open interval $(a, +\infty)$. A similar function can be constructed for a semi-infinite open interval $(-\infty, b)$.

Now let F be any non-empty closed set. The complement of F consists of a countable set of pairwise disjoint finite open intervals (a_k, b_k), together with possibly one or two semi-infinite open intervals. Define $f(x)$ to be zero on F, let $f(x) = e^{\frac{1}{(x-a_k)(x-b_k)}}$ in each finite open interval complementary to F with endpoints in F, $f(x) = e^{\frac{1}{a-x}}$ for $x > a$ if the complement of F contains a semi-infinite interval $(a, +\infty)$ with endpoint $a \in F$, and $f(x) = e^{\frac{1}{x-b}}$ if the complement of F contains a semi-infinite interval $(-\infty, b)$ with endpoint $b \in F$.

It is now obvious that f is zero precisely on F, and that F has derivatives of all orders at each point of the complement of F and at each interior point of F.

It remains to be shown that f has derivatives of all orders at each boundary point x of F. There are actually 4 cases to consider, but all are handled alike, and we shall settle for just one typical case, in which there is a decreasing sequence of points $x_p \in F$, $x_p \to x$, and a decreasing sequence of points $y_p \notin F$, $y_p \to x$, but no increasing sequence of points $z_p \in F$, $z_p \to x$. This means either $x = b_k$ for some k or $x = b$. Now for each y such that $x < y < x_1$ there is a complementary interval to F, say $(a_l, b_l) \subset (x, x_1)$, with $a_k < y < b_k$. Then for all nonnegative integers k and n we have

$$0 < f^{(k)}(y) < K_{n,k}(y - a_l)^n < K_{n,k}(y - x)^n$$

where $K_{n,k}$ is a positive constant independent of l, hence independent of y. It therefore follows, upon taking $n = 2$ that if $x_1 > y > x$, then

$$\left|\frac{f(y) - f(x)}{(y - x)}\right| \leq K_{2,0}(y - x)$$

(We have just proved this inequality for $y \notin F$, and $f(y) = f(x) = 0$ if $y \in F$.) Hence the right-handed derivative

$$f'_+(x) = \lim_{y \downarrow x} \frac{f(y) - f(x)}{y - x}$$

is zero. That the left-hand derivative is zero follows from the fact that $x = b_k$ or $x = b$. Hence $f'(x) = 0$. We now assume by induction that $f^{(k-1)}(x) = 0$ Then the inequality $f^{(k-1)}(y) \leq K_{2,k-1}(y - x)^2$ shows that

$$f_+^{(k)}(x) = \lim_{y \downarrow x} \frac{f^{(k-1)}(y)}{y - x} = 0.$$

Again, the left-hand kth derivative is zero since $x = b_k$ or $x = b$. It follows easily that $f^{(k)}(x)$ exists and equals zero for all k.

Exercise 5.22 Suppose f is a real function on $(-\infty, \infty)$. Call x a *fixed point* of f if $f(x) = x$.

(*a*) If f is differentiable and $f'(t) \neq 1$ for every real t, prove that f has at most one fixed point.

(*b*) Show that the function f defined by

$$f(t) = t + (1 + e^t)^{-1}$$

has no fixed point, although $0 < f'(t) < 1$ for all real t.

(*c*) However, if there is a constant $A < 1$ such that $|f'(t)| \leq A$ for all real t, prove that a fixed point x of f exists, and that $x = \lim x_n$, where x_1 is an arbitrary real number and

$$x_{n+1} = f(x_n)$$

for $n = 1, 2, 3, \ldots$.

(*d*) Show that the process described in (*c*) can be visualized by the zig-zag path

$$(x_1, x_2) \to (x_2, x_2) \to (x_2, x_3) \to (x_3, x_3) \to (x_3, x_4) \to \cdots.$$

Solution. (*a*) If a function $f(x)$ has two fixed points x and y, $x \neq y$, the mean-value theorem implies that there exists a point z between x and y such that

$$y - x = f(y) - f(x) = f'(z)(y - x),$$

so that $f'(z) = 1$.

(*b*) The equation $f(t) = t$ implies that $(1+e^t)^{-1} = 0$, which is clearly impossible, while $f'(t) = 1 - \dfrac{e^t}{(1+e^t)^2}$ always lies in $(0, 1)$.

(*c*) Since f' is bounded, f is uniformly continuous, and we observe that the sequence $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence. Indeed, if $n > m > N$, we have

$$|x_n - x_m| \leq |x_n - x_{n-1}| + |x_{n-1} - x_{n-2}| + \cdots + |x_{m+1} - x_m|.$$

Now it is easy to show by induction, using the mean-value theorem and the fact that $|f'(x)| \leq A$ for all x, that

$$|x_{n+1} - x_n| \leq A^{n-1}|x_2 - x_1|$$

for $n \geq 1$. We therefore have

$$\begin{aligned} |x_n - x_m| &\leq |x_2 - x_1|(A^{n-2} + A^{n-3} + \cdots + A^{m-1}) \\ &< \frac{1}{1-A} A^{m-1}|x_2 - x_1| \\ &\leq \frac{|x_2 - x_1|}{1-A} A^N. \end{aligned}$$

Since $0 \leq A < 1$, it follows that $A^N \to 0$ as $N \to \infty$, and so this is a Cauchy sequence. Let its limit be x. We claim that x is a fixed point. Indeed, $x =$

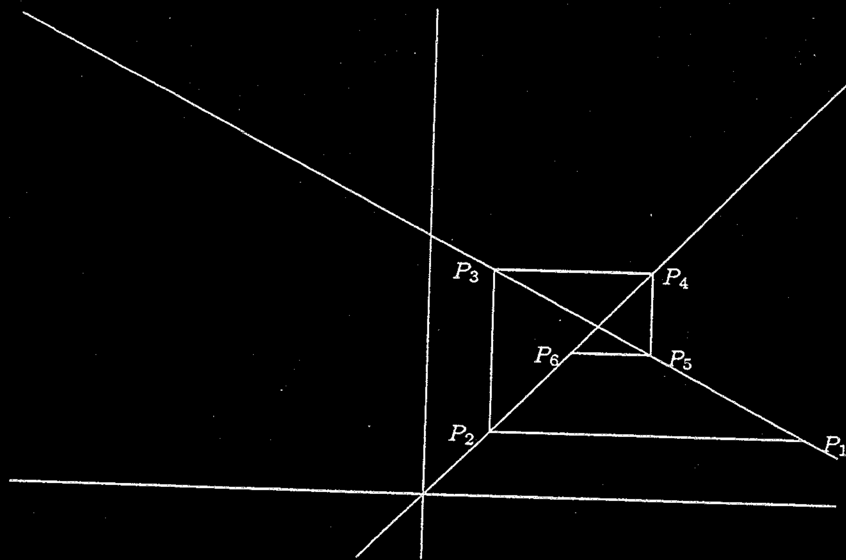

Figure 5.1: Finding a fixed point

$\lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n) = f(x)$, since f is continuous. There can of course be only one fixed point because of the result proved in (a).

(d) The procedure described can be depicted on the graph of the function f, i.e., the set of points $(x, f(x))$, as follows: Let x_1 be any abscissa; locate the point $(x_1, f(x_1))$ on the graph. Thereafter, for each point (x_n, y_n) located on the graph, let the abscissa of (x_{n+1}, y_{n+1}) be the ordinate of (x_n, y_n), i.e., $x_{n+1} = y_n$. Thus, from a point (x_n, y_n) on the graph of f we move horizontally to the line $y = x$, then vertically back to the graph of f. It is clear visually that this process leads to the point of intersection of the graph of f with the line $y = x$, as illustrated in Fig. 1 for the case of $f(x) = 2 - \frac{1}{2}x$, where $P_1 = (2, 1)$, $P_2 = (1, 1)$, $P_3 = (1, \frac{3}{2})$, $P_4 = (\frac{3}{2}, \frac{3}{2})$, $P_5 = (\frac{3}{2}, \frac{5}{4})$, and $P_6 = (\frac{5}{4}, \frac{5}{4})$. (The fixed point is $(\frac{4}{3}, \frac{4}{3})$, which is the point of intersection of the graph of f and the line $y = x$.)

Exercise 5.23 The function f defined by

$$f(x) = \frac{x^3 + 1}{3}$$

has three fixed points, say α, β, γ, where

$$-2 < \alpha < -1, \quad 0 < \beta < 1, \quad 1 < \gamma < 2.$$

For arbitrarily chosen x_1, define $\{x_n\}$ by setting $x_{n+1} = f(x_n)$.

(a) If $x_1 < \alpha$, prove that $x_n \to -\infty$ as $n \to \infty$.

(b) If $\alpha < x_1 < \gamma$, prove that $x_n \to \beta$ as $n \to \infty$.

(c) If $\gamma < x_1$, prove that $x_n \to +\infty$ as $n \to \infty$.

Thus β can be located by this method, but α and γ cannot.

Solution. We shall make use of the auxiliary functions

$$g(x) = f(x) - x = \frac{x^3+1}{3} - x$$

and

$$h(x) = \begin{cases} \dfrac{g(x)-g(\beta)}{x-\beta}, & x \neq \beta, \\ g'(\beta) & x = \beta, \end{cases}$$

i.e., $g(x) = \dfrac{x^2+\beta x+\beta^2}{3} - 1$. We observe that the fixed points of f are the zeros of g. Since $g(-2) = -\frac{1}{3} < 0$, $g(-1) = 1 > 0$, $g(0) = \frac{1}{3} > 0$, $g(1) = -\frac{1}{3} < 0$, and $g(2) = 1 > 0$, the intermediate value theorem shows that α, β, and γ are located in the intervals they are asserted to be in.

Since $g(\alpha) = g(\beta) = g(\gamma) = 0$, it follows that $h(\alpha) = h(\gamma) = 0$. Since h is a quadratic function, it has only the two zeros α and γ, and in particular $h(x)$ is negative for $\alpha < x < \gamma$. Now the minimum value of $h(x)$ is attained at $x = -\frac{\beta}{2}$, and this minimum value is c, where $c = \dfrac{\beta^2}{4} - 1$. Thus $-1 < c < 0$. In particular, for $\alpha < x < \gamma$ there is a number $r \in (0,1)$ such that

$$f(x) - x = r(\beta - x),$$

i.e.,

$$f(x) - \beta = s(x - \beta),$$

where $s = 1 - r$ is also in the interval $(0,1)$. This means that $f(x) - \beta$ and $x - \beta$ both have the same sign, but that $|f(x) - \beta| < |x - \beta|$. Thus $f(x)$ is always between β and x. Therefore the sequence $\{x_n\}$ is monotonic and converges to a fixed point in the interval whose endpoints are x_1 and β. Since the only fixed point in this interval is β, the sequence must converge to β.

If $x < \alpha$ (resp. $x > \gamma$), it is easy to see that $f(x) < x$ (resp. $f(x) > x$). Thus the sequence $\{x_n\}$ is monotonically decreasing (resp. increasing), and hence either tends to $-\infty$ (resp. $+\infty$) or converges to a fixed point δ in the interval $(-\infty, x_1)$ (resp. $(x_1, +\infty)$). Since there are no fixed points in this interval, it follows that $x_n \to -\infty$ (resp. $x_n \to +\infty$).

Exercise 5.24 The process described in part (*c*) of Exercise 22 can of course also be applied to functions that map $(0,\infty)$ to $(0,\infty)$.

Fix some $\alpha > 1$, and put

$$f(x) = \frac{1}{2}\left(x + \frac{\alpha}{x}\right), \quad g(x) = \frac{\alpha + x}{1+x}.$$

Both f and g have $\sqrt{\alpha}$ as their only fixed point in $(0, \infty)$. Try to explain, on the basis of properties of f and g, why the convergence in Exercise 16, Chap. 3, is so much more rapid than it is in Exercise 17. (Compare f' and g', draw the zig-zag suggested in Exercise 22.)

Do the same when $0 < \alpha < 1$.

Solution. We recall that in Chap. 3 we proved that the first function leads to $|x_n - \sqrt{\alpha}| \le Ar^{2^n}$ for some $r \in (0,1)$, while the second leads only to $|x_n - \sqrt{\alpha}| \le Ar^n$. The exact values of A and r depend on α and x_1.

The best explanation of the difference between the two methods is that

$$\begin{aligned} f(x) - \sqrt{\alpha} &= \frac{1}{2}\Big(1 - \frac{\sqrt{\alpha}}{x}\Big)(x - \sqrt{\alpha}), \\ g(x) - \sqrt{\alpha} &= \frac{1 - \sqrt{\alpha}}{1 + x}(x - \sqrt{\alpha}). \end{aligned}$$

The first of these makes it plain that if $x > \sqrt{\alpha}$, the same will be true of $f(x)$, though $f(x)$ will be closer to α than x by a factor that is at most $\frac{1}{2}$ and tends to zero as x tends to $\sqrt{\alpha}$, i.e., the relative improvement in accuracy itself improves as the recursion proceeds. The second equality shows that $g(x) - \alpha$ is on the opposite side of $\sqrt{\alpha}$ from x if $\alpha > 1$, though closer by a factor that is at least the absolute value of $\dfrac{1 - \sqrt{\alpha}}{1 + x_1}$. Hence the relative improvement in accuracy as the recursion proceeds is limited.

In terms of the zigzag pattern, when we use g, the zigzag keeps circulating around the point of intersection of the graph of g and the line $y = x$ instead of moving steadily toward it in a staircase pattern.

When $0 < \alpha < 1$, the zigzag does stay on one side of the point of intersection of the two curves. However, the relative improvement is still at best a factor of $\dfrac{1 - \sqrt{\alpha}}{2}$ when x is close to $\sqrt{\alpha}$.

Exercise 5.25 Suppose f is twice differentiable in $[a, b]$, $f(a) < 0$, $f(b) > 0$, $f'(x) \ge \delta > 0$, and $0 \le f''(x) \le M$ for all $x \in [a, b]$. Let ξ be the unique point in (a, b) at which $f(\xi) = 0$.

Complete the details in the following outline of *Newton's method* for computing ξ.

(*a*) Choose $x_1 \in (\xi, b)$, and define x_n by

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Interpret this geometrically, in terms of a tangent to the graph of f.

(*b*) Prove that $x_{n+1} < x_n$, and that

$$\lim_{n \to \infty} x_n = \xi.$$

(c) Use Taylor's theorem to show that

$$x_{n+1} - \xi = \frac{f''(t_n)}{2f'(x_n)}(x_n - \xi)^2$$

for some $t_n \in (\xi, x_n)$.

(d) If $A = M/2\delta$, deduce that

$$0 \le x_{n+1} - \xi \le \frac{1}{A}[A(x_1 - \xi)]^{2^n}.$$

(Compare with Exercises 16 and 18, Chap. 3.)

(e) Show that Newton's method amounts to finding a fixed point of the function g defined by

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

How does $g'(x)$ behave for x near ξ?

(f) Put $f(x) = x^{1/3}$ on $(-\infty, \infty)$ and try Newton's method. What happens?

Solution. We remark at the outset that x_1 can be found by trying $z_0 = \frac{a+b}{2}$. If $f(z_0) > 0$, take $x_1 = z_0$. Otherwise let $z_{n+1} = (b + z_n)/2$, and let x_1 be the first z_n for which $f(z_n) > 0$. (In a finite number of steps we must reach such a point since $z_n \uparrow b$ and $f(b) > 0$.)

(a) The tangent line to the graph of f at the point x_n has the equation $y - f(x_n) = f'(x_n)(x - x_n)$. Setting $y = 0$ in this equation and solving for x gives $x = x_{n+1}$. Thus the interpretation of Newton's method is that we approximate the point where the graph of f intersects the x-axis by the point at which its tangent line at $(x_n, f(x_n))$ intersects the x-axis.

(b) We can assume by induction that $f(x_n) > 0$, and hence, since $f'(x_n) > 0$, it follows immediately that $x_{n+1} < x_n$. Notice that there exists c between x_n and x_{n+1} such that $f(x_{n+1}) = f(x_n) - f'(c)(x_n - x_{n+1}) > f(x_n) - f'(x_n)(x_n - x_{n+1}) = 0$ since $f'(c) < f'(x_n)$ and $x_n - x_{n+1} > 0$. Thus it follows that $\xi < x_{n+1} < x_n$. Hence $\{x_n\}$ converges to a limit η satisfying $\eta \ge \xi$. Now, however, we have

$$\eta = \eta - \frac{f(\eta)}{f'(\eta)},$$

from which it follows that $f(\eta) = 0$, i.e., $\eta = \xi$.

(c) The required equality can be written as

$$x_n - \xi - \frac{f(x_n)}{f'(x_n)} = \frac{f''(t_n)}{2f'(x_n)}(x_n - \xi)^2,$$

while Taylor's theorem can be written as

$$f(\xi) = f(x_n) + f'(x_n)(\xi - x_n) + \frac{f''(t_n)}{2}(\xi - x_n)^2.$$

Since $f(\xi) = 0$, it is clear that these two equations are equivalent.

(*d*) Since $0 \le f''(t_n) \le M$ and $f'(x_n) > \delta$, we have

$$0 \le x_{n+1} - \xi \le A(x_n - \xi)^2.$$

In particular

$$0 \le x_2 - \xi \le A(x_1 - \xi)^2 = \frac{1}{A}[A(x_1 - \xi)]^2,$$

and then an easy induction gets the general result.

We found this kind of convergence in Exercises 16 and 18 of Chap. 3 with the recursion relation

$$x_{n+1} = \frac{p-1}{p}x_n + \frac{\alpha}{p}x_n^{-p+1}.$$

We now recognize this recursion as Newton's method for the function $f(x) = x^p - \alpha$ on the interval $[1, \sqrt{\alpha} + 1]$. Exercise 16 of Chap. 2 was the special case $p = 2$.

(*e*) Obviously the equation $g(x) = x$ is equivalent to the equation $f(x) = 0$.

Since $g'(x) = \dfrac{f(x)f''(x)}{[f'(x)]^2}$, we see that $g'(x)$ tends to zero as x tends to ξ, i.e., the graph of $g(x)$ meets the line $y = x$ at a 45° degree angle at the point (ξ, ξ).

(*f*) The fixed point of $f(x)$ is $x = 0$. However $f'(x) \to \infty$ as $x \to 0$, and $f'(0)$ does not exist. This destroys the convergence of Newton's method. In fact, if $x_n \ne 0$, then $x_{n+1} = -2x_n$, so that x_n oscillates wildly: $\limsup x_n = +\infty$, $\liminf x_n = -\infty$.

Exercise 5.26 Suppose f is differentiable on $[a, b]$, $f(a) = 0$, and there is a real number A such that $|f'(x)| \le A|f(x)|$ on $[a, b]$. Prove that $f(x) = 0$ for all $x \in [a, b]$. *Hint:* Fix $x_n \in [a, b]$, let

$$M_0 = \sup|f(x)|, \quad M_1 = \sup|f'(x)|$$

for $a \le x \le x_0$. For any such x,

$$|f(x)| \le M_1(x_0 - a) \le A(x_0 - a)M_n.$$

Hence $M_0 = 0$ if $A(x_0 - a) \le 1$. That is, $f = 0$ on $[a, x_0]$. Proceed.

Solution. If we anticipate the fundamental result that the function $f(x) = e^x$ satisfies $f'(x) = f(x)$, Exercise 2 above yields the result that $\ln x$ is differentiable and has derivative $\frac{1}{x}$. Hence by the chain rule for any positive differentiable function $f(x)$ the function $g(x) = \ln f(x)$ is differentiable and $g'(x) = \dfrac{f'(x)}{f(x)}$. (Unfortunately this fundamental result is not proved until Chapter 7, so we shall just have to wait. However, since certain other functions such as $\sin x$ and $\cos x$ have been introduced without any formal definition, and their derivatives have been assumed known, we might as well continue along this line of reasoning.)

Now suppose there is an interval $(c,d) \subset [a,b]$ such that $f(c) = 0$ but $f(x) \neq 0$ for $c < x < d$. By passing to consideration of $-f(x)$ if necessary, we can assume $f(x) > 0$ for $c < x < d$. The function $g(x) = \ln f(x)$ is then defined for $c < x < d$, and its derivative satisfies

$$|g'(x)| = \left|\frac{f'(x)}{f(x)}\right| \leq A.$$

The mean-value theorem them implies that

$$g(x) \geq g\Big(\frac{c+d}{2}\Big) - A\Big(\frac{d-c}{2}\Big)$$

for all $x \in (c,d)$. But this is a contradiction, since $g(x) \longrightarrow -\infty$ as $x \longrightarrow c$.

This finishes the proof, except that it assumes we know the derivative of e^x. If we don't assume that, we have to fall back on the hint. In that case, let $x_0 = a + \frac{1}{2A}$, and let $M_0 = \sup\{|f(x)| : a \leq x \leq x_0\}$. We then have

$$|f(x)| \leq M_1(x-a) \leq AM_0(x_0 - a) = \frac{1}{2}M_0$$

for all $x \in [a, x_0]$. But by definition of M_0 this implies $M_0 \leq \frac{1}{2}M_0$, so that $M_0 \leq 0$, i.e., $M_0 = 0$. We now start over with a replaced by x_0, $x_1 = x_0 + \dfrac{1}{2A}$. In a finite number of steps, we will have $b < x_n + \dfrac{1}{2A}$, so that $f(x) = 0$ for $a \leq x \leq b$.

Exercise 5.27 Let ϕ be a real function defined on a rectangle R in the plane, given by $a \leq x \leq b$, $\alpha \leq y \leq \beta$. A *solution* of the initial-value problem

$$y' = \phi(x,y), \quad y(a) = c \quad (\alpha \leq c \leq \beta)$$

is, by definition, a differentiable function f on $[a,b]$ such that $f(a) = c$, $\alpha \leq f(x) \leq \beta$, and

$$f'(x) = \phi(x, f(x)) \quad (a \leq x \leq b).$$

Prove that such a problem has at most one such solution if there is a constant A such that

$$|\phi(x,y_2) - \phi(x,y_1)| \leq A|y_2 - y_1|$$

whenever $(x,y_1) \in R$ and $(x,y_2) \in R$.

Hint: Apply Exercise 26 to the difference of two solutions. Note that this uniqueness theorem does not hold for the initial-value problem

$$y' = y^{1/2}, \quad y(0) = 0,$$

which has two solutions: $f(x) = 0$ and $f(x) = x^2/4$. Find all other solutions.

Solution. Following the hint, we observe that if $f(x) = f_2(x) - f_1(x)$, then

$$\begin{aligned} |g'(x)| &= |f_2'(x) - f_1'(x)| \\ &= |\phi(x, f_2(x)) - \phi(x, f_1(x))| \\ &\leq A|f_2(x) - f_1(x)| \\ &= A|g(x)|. \end{aligned}$$

By the initial condition $g(a) = f_2(a) - f_1(a) = c - c = 0$. Hence by the preceding exercise $g(x) = 0$ for all $x \in [a, b]$.

As for the equation $y' = \sqrt{y}$, if $f(x)$ is a solution and $f(x) > 0$ on an interval (a, b), while $f(a) = 0$, we observe that $g(x) = \sqrt{f(x)}$ satisfies $g'(x) = \frac{1}{2}(f(x))^{-1/2}f'(x) = \frac{1}{2}$, so that for some constant c we have $g(x) = \frac{1}{2}(x + c)$. Thus

$$f(x) = (g(x))^2 = \frac{1}{4}(x + c)^2.$$

Since $f(a) = 0$, it follows that $c = -a$, i.e., $f(x) = \dfrac{(x-a)^2}{4}$. Thus the only possible solutions are

$$f(x) = \begin{cases} 0, & 0 \leq x \leq a, \\ \dfrac{(x-a)^2}{4}, & a \leq x. \end{cases}$$

Here $a \geq 0$ is arbitrary.

Exercise 5.28 Formulate and prove an analogous uniqueness theorem for systems of differential equations of the form

$$y_j' = \phi_j(x, y_1, \ldots, y_k), \quad y_j = c_j \quad (j = 1, \ldots, k).$$

Note that this can be rewritten in the form

$$\mathbf{y}' = \phi(x, \mathbf{y}), \quad \mathbf{y}(a) = \mathbf{c},$$

where $\mathbf{y} = (y_1, \ldots, y_k)$ ranges over a k-cell, ϕ is the mapping of a $(k+1)$-cell into the Euclidean k-space whose components are the functions $\phi_1, \ldots, \phi_k$, and $\mathbf{c}$ is the vector $(c_1, \ldots, c_k)$. Use Exercise 26 for vector-valued functions.

Solution. The result is the following:

Let ϕ be a vector-valued function defined on a $(k+1)$-cell $D = [a, b] \times C$ in R^{k+1} whose range is contained in R^k, and suppose that there exists a constant A such that

$$|\phi(x, \mathbf{y}_2) - \phi(x, \mathbf{y}_1)| \leq A|\mathbf{y}_2 - \mathbf{y}_1|$$

for all $\mathbf{y}_1 \in C$, $\mathbf{y}_2 \in C$. *Then the initial-value problem*

$$\mathbf{y}' = \phi(x, \mathbf{y}) \quad \mathbf{y}(a) = \mathbf{c}$$

has at most one solution $\mathbf{y} : [a, b] \to C$.

The main tool needed to prove this result is the analogue of Exercise 26 for vector-valued functions, which does hold. Indeed the proof is identical, considering that the original proof depends only on the inequality $|f(d) - f(c)| \le |f'(r)|(d - c)$ for some $r \in (c, d)$, and this inequality is certainly valid for vector-valued functions. Once that result is obtained, the preceding exercise can be applied verbatim.

Exercise 5.29 Specialize Exercise 28 by considering the system

$$\begin{aligned} y_j' &= y_{j+1} \quad (j = 1, \ldots, k-1), \\ y_k' &= f(x) - \sum_{j=1}^{k} g_j(x) y_j, \end{aligned}$$

where $f, g_1, \ldots, g_k$ are continuous real functions on $[a, b]$, and derive a uniqueness theorem for solutions of the equation

$$y^{(k)} + g_k(x) y^{(k-1)} + \cdots + g_2(x) y' + g_1(x) y = f(x),$$

subject to initial conditions

$$y(a) = c_1, \quad y'(a) = c_2, \quad \ldots, \quad y^{(k-1)}(a) = c_k.$$

Solution. We let $\mathbf{y} = (y_1, y_2, y_3, \ldots, y_k) = (y, y', y'', \ldots, y^{(k-1)})$ and $\phi(x, \mathbf{y}) = \left(y_2, y_3, \ldots, y_k, f(x) - \sum_{j=1}^{k} g_j(x) y_j\right)$. We then observe that if $\mathbf{y}_i = (y_{i1}, \ldots, y_{ik})$, then

$$|\phi(x, \mathbf{y}_2) - \phi(x, \mathbf{y}_1)| = \left|\left(y_{22} - y_{12}, y_{23} - y_{13}, \ldots, \sum_{j=1}^{k} g_j(x)(y_{1j} - y_{2j})\right)\right|.$$

If $M = \sup\{|g_j(x)| : a \le x \le b,\ 1 \le j \le k\}$, we then have

$$|\phi(x, \mathbf{y}_2) - \phi(x, \mathbf{y}_1)| \le (M+1) \sum_{j=1}^{k} |y_{2j} - y_{1j}| \le k(M+1)|\mathbf{y}_2 - \mathbf{y}_1|.$$

This provides the hypothesis of the theorem for any $(k+1)$-cell $[a, b] \times C$ whatsoever in R^{k+1}. Hence there is at most one solution to this initial-value problem.