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

Roger Cooke, University of Vermont

Chapter 9

Functions of Several Variables

Exercise 9.1 If S is a nonempty subset of a vector space X, prove (as asserted in Sec. 9.1) that the span of S is a vector space.

Solution. We need only verify that the span of S is closed under the two vector space operations. All the other properties of a vector space hold in the span of S, since it is contained in a vector space in which they hold.

To that end, let $\mathbf{x}$ and $\mathbf{y}$ be elements of the span of S, and let c be any real number. By definition there are elements $\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{y}_1, \ldots, \mathbf{y}_n$, and scalars $c_1, \ldots, c_m, d_1, \ldots, d_n$ such that $\mathbf{x} = c_1\mathbf{x}_1 + \cdots + c_m\mathbf{x}_m$ and $\mathbf{y} = d_1\mathbf{y}_1 + \cdots + d_n\mathbf{y}_n$. We then have

$$\mathbf{x} + \mathbf{y} = c_1\mathbf{x}_1 + \cdots + c_m\mathbf{x}_m + d_1\mathbf{y}_1 + \cdots + d_n\mathbf{y}_n,$$

which is a finite linear combination of elements of S, hence belongs to the span of S. Likewise, by the distributive law,

$$c\mathbf{x} = c(c_1\mathbf{x}_1 + \cdots + c_m\mathbf{x}_m) = (cc_1)\mathbf{x}_1 + \cdots + (cc_m)\mathbf{x}_m,$$

which belongs to the span of S.

Exercise 9.2 Prove (as asserted in Sec. 9.6) that BA is linear if A and B are linear transformations. Prove also that A^{-1} is linear and invertible.

Solution. Let $A : X \to Y$ and $B : Y \to Z$ be linear transformations, and let $\mathbf{x}$ and $\mathbf{y}$ be any elements of A and c any scalar. Then $BA : X \to Z$ satisfies

$$\begin{aligned} BA(\mathbf{x}+\mathbf{y}) &= B\big(A(\mathbf{x}+\mathbf{y})\big) \\ &= B\big(A(\mathbf{x}) + A(\mathbf{y})\big) \\ &= B\big(A(x)\big) + B\big(A(\mathbf{y})\big) \\ &= BA(\mathbf{x}) + BA(\mathbf{y}). \end{aligned}$$

Similarly,

$$\begin{aligned}
BA(c\mathbf{x}) &= B(A(c\mathbf{x})) \\
&= B(cA(\mathbf{x})) \\
&= cB(A(\mathbf{x})) \\
&= cBA(\mathbf{x}).
\end{aligned}$$

If A is a one-to-one mapping of X onto Y, and $\mathbf{z}$ and $\mathbf{w}$ are any elements of Y, let $\mathbf{x} = A^{-1}(\mathbf{z})$ and $\mathbf{y} = A^{-1}(\mathbf{w})$. Then by definition $A(\mathbf{x}) = \mathbf{z}$ and $A(\mathbf{y}) = \mathbf{w}$. It therefore follows from the linearity of A that $A(\mathbf{x}+\mathbf{y}) = \mathbf{z}+\mathbf{w}$. Again, by definition, this means that $A^{-1}(\mathbf{z}+\mathbf{w}) = \mathbf{x}+\mathbf{y} = A^{-1}(\mathbf{z}) + A^{-1}(\mathbf{w})$, so that A^{-1} preserves vector addition. Similarly, $A(c\mathbf{x}) = cA(\mathbf{x}) = c\mathbf{z}$, so that $A^{-1}(c\mathbf{z}) = c\mathbf{x} = cA^{-1}(\mathbf{z})$, and hence A^{-1} also preserves scalar multiplication.

Exercise 9.3 Assume $A \in L(X,Y)$ and $A\mathbf{x} = \mathbf{0}$ only when $\mathbf{x} = \mathbf{0}$. Prove that A is then 1-1.

Solution. Suppose $A(\mathbf{x}) = A(\mathbf{y})$. It then follows that $A(\mathbf{x}-\mathbf{y}) = A(\mathbf{x}) - A(\mathbf{y}) = \mathbf{0}$. Hence by assumption $\mathbf{x}-\mathbf{y} = \mathbf{0}$, and so $\mathbf{x} = \mathbf{y}$; therefore A is one-to-one.

Exercise 9.4 Prove (as asserted in Sec. 9.30) that null spaces and ranges of linear transformations are vector spaces.

Solution. Let N be the null space of the linear transformation $A : X \to Y$, let $\mathbf{x}$ and $\mathbf{y}$ be elements of N, and let c be any scalar. By definition $A(\mathbf{x}) = \mathbf{0} = A(\mathbf{y})$, and $A(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}) + A(\mathbf{y}) = \mathbf{0}+\mathbf{0} = \mathbf{0}$, so that, by definition, $\mathbf{x}+\mathbf{y} \in N$. Likewise $A(c\mathbf{x}) = cA(\mathbf{x}) = c\mathbf{0} = \mathbf{0}$, and so $c\mathbf{x} \in N$. Therefore N is a subspace of X.

Let R be the range of A, let $\mathbf{z}$ and $\mathbf{w}$ be any elements of R, and let c be any scalar. By definition, there exist vectors $\mathbf{x} \in X$ and $\mathbf{y} \in X$ such that $z = A(\mathbf{x})$ and $w = A(\mathbf{y})$. Then $A(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}) + A(\mathbf{y}) = \mathbf{z}+\mathbf{w}$, and hence $\mathbf{z}+\mathbf{w} \in R$. Likewise $A(c\mathbf{x}) = cA(\mathbf{x}) = c\mathbf{z}$, so that $c\mathbf{z} \in R$. Therefore R is a subspace of Y.

Exercise 9.5 Prove that to every $A \in L(R^n, R^1)$ corresponds a unique $\mathbf{y} \in R^n$ such that $A\mathbf{x} = \mathbf{x}\cdot\mathbf{y}$. Prove also that $\|A\| = |\mathbf{y}|$.

Hint: Under certain conditions, equality holds in the Schwarz inequality.

Solution. Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis of R^n, and let $\mathbf{y} = A(\mathbf{e}_1)\mathbf{e}_1 + \cdots + A(\mathbf{e}_n)\mathbf{e}_n$. Then for any $\mathbf{x} = c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n$ we have

$$\begin{aligned}
A(\mathbf{x}) &= c_1A(\mathbf{e}_1) + \cdots + c_nA(\mathbf{e}_n) \\
&= \mathbf{y}\cdot\mathbf{x}.
\end{aligned}$$

There can be at most one such $\mathbf{y}$, since if $A(\mathbf{x}) = \mathbf{z}\cdot\mathbf{x}$, then $|\mathbf{y}-\mathbf{z}|^2 = \mathbf{y}\cdot\mathbf{y} - \mathbf{y}\cdot\mathbf{z} - \mathbf{z}\cdot\mathbf{y} + \mathbf{z}\cdot\mathbf{z} = A(\mathbf{y}) - A(\mathbf{y}) - A(\mathbf{z}) + A(\mathbf{z}) = 0$.

By the Schwarz inequality we have

$$|A(\mathbf{x})| = |\mathbf{y}\cdot\mathbf{x}| \le |\mathbf{y}|\,|\mathbf{x}|$$

for all $\mathbf{x}$, so that $\|A\| \le |\mathbf{y}|$. On the other hand $A(\mathbf{y}) = \mathbf{y}\cdot\mathbf{y} = |\mathbf{y}|^2$, so that $\|A\| \ge |\mathbf{y}|$.

Exercise 9.6 If $f(0,0) = 0$ and

$$f(x,y) = \frac{xy}{x^2+y^2} \quad \text{if } (x,y) \ne (0,0),$$

prove that $(D_1 f)(x,y)$ and $(D_2 f)(x,y)$ exist at every point of R^2, although f is not continuous at $(0,0)$.

Solution. At any point (x,y) except $(0,0)$ the differentiability of $f(x,y)$ follows from the rules for differentiation and the principles of Chapter 5. At $(0,0)$ it is a routine computation to verify that both partial derivatives equal zero:

$$(D_1 f)(0,0) = \lim_{h\to 0} \frac{f(h,0)-f(0,0)}{h} = 0.$$

However, $f(x,y)$ is not continuous at $(0,0$, since $f(x,x) = \frac{1}{2}$ for all $x \ne 0$, and hence $\lim_{x\to 0} f(x,x) = \dfrac{1}{2} \ne f(0,0)$.

Exercise 9.7 Suppose that f is a real-valued function defined in an open set $E \subset R^n$, and that the partial derivatives $D_1 f, \ldots, D_n f$ are bounded in E. Prove that f is continuous in E

Hint: Proceed as in the proof of Theorem 9.21.

Solution. Let $\varepsilon > 0$ be given, and let $\mathbf{x}^0 = (x_1^0, \ldots, x_n^0)$ be any point of E. First choose $\delta_0 > 0$ so that $\mathbf{y} \in E$ if $|\mathbf{y}-\mathbf{x}^0| < 2\delta_0$. Then, if $M = \max\limits_{\mathbf{x}\in E}\big((D_1 f)(\mathbf{x}), \ldots, (D_n f)(\mathbf{x})\big)$, choose $\delta = \min\Big(\delta_0, \dfrac{\varepsilon}{(n+1)M}\Big)$. It then follows that if $|\mathbf{y}-\mathbf{x}^0| < \delta$, we have

$$\begin{aligned}
|f(\mathbf{y}) - f(\mathbf{x}^0)| &= |f(y_1,\ldots,y_n) - f(x_1^0,\ldots,x_n^0)| \\
&\le |f(y_1,y_2,\ldots,y_n) - f(x_1^0,y_2,\ldots,y_n)| + \\
&\quad + |f(x_1^0,y_2),\ldots,y_n) - f(x_1^0,x_2^0,\ldots,y_n)| + \cdots \\
&\quad \cdots + |f(x_1^0,x_2^0,\ldots,x_{n-1}^0,y_n) - f(x_1^0,x_2^0,\ldots,x_{n-1}^0,x_n^0)|,
\end{aligned}$$

where the ellipsis indicates terms of the form

$$|f(x_1^0,x_2^0,\ldots,x_{k-1}^0,y_k,y_{k+1},\ldots,y_n) - f(x_1^0,x_2^0,\ldots,x_{k-1}^0,x_k^0,y_{k+1},\ldots,y_n)|.$$

By the mean-value theorem there is a number c_k between x_k^0 and y_k such that this last difference equals

$$|(D_k f)(x_1^0, x_2^0, \ldots, x_{k-1}^0, c_k, y_{k+1}, \ldots, y_n)(y_k - x_k^0)|,$$

which is at most $M\delta$. Since by definition $M\delta$ is at most $\dfrac{\varepsilon}{n+1}$, and there are only n such terms, it follows that $|f(\mathbf{x}^0) - f(\mathbf{y})| < \varepsilon$. Thus f is continuous.

Remark: We have actually shown that $f(\mathbf{x})$ satisfies a Lipschitz condition on any convex subset of E, i.e., that $|f(\mathbf{x}) - f(\mathbf{y})| \le nM|\mathbf{x} - \mathbf{y}|$ on each convex subset.

Exercise 9.8 Suppose that f is a differentiable real function in an open set $E \subset R^n$, and that f has a local maximum at a point $\mathbf{x} \in E$. Prove that $f'(\mathbf{x}) = 0$.

Solution. Let $\mathbf{y}$ be any element of R^n, and consider the real-valued function $\varphi(t) = f(\mathbf{x}+t\mathbf{y})$, defined near $t = 0$. This function is differentiable (by Theorem 9.15 $\varphi(t) = f'(\mathbf{x} + t\mathbf{y})(\mathbf{y})$). Since $\varphi(t)$ has a maximum at $t = 0$, it follows that $\varphi'(0) = 0$, i.e., that $f'(\mathbf{x})(\mathbf{y}) = 0$. Since $\mathbf{y}$ is arbitrary, it follows by definition of the zero linear transformation that $f'(\mathbf{x})$ is the zero linear transformation.

Exercise 9.9 If $\mathbf{f}$ is a differentiable mapping of a *connected* open set $E \subset R^n$ into R^m, and if $\mathbf{f}'(\mathbf{x}) = \mathbf{0}$ for for every $\mathbf{x} \in E$, prove that $\mathbf{f}$ is constant in E.

Solution. The mean-value argument given in Exercise 7 above, applied to each component of $\mathbf{f}$, shows that $\mathbf{f}$ is *locally* constant (the partial derivatives are all zero). Hence, if $\mathbf{x}^0$ is any point of E, the set of $\mathbf{x}$ such that $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}^0)$ is an open set. Since this set is also closed in E, and E is connected, it follows that it must be all of E.

Exercise 9.10 If f is a real function defined in a convex open set $E \subset R^n$, such that $(D_1 f)(\mathbf{x}) = 0$ for every $\mathbf{x} \in E$, prove that $f(\mathbf{x})$ depends only on $x_2, \ldots, x_n$.

Show that the convexity of E can be replaced by a weaker condition, but that some condition is required. For example, if $n = 2$ and E is shaped like a horseshoe, the statement may be false.

Solution. We need to show that $f(x_1^0, x_2, \ldots, x_n) = f(x_1^1, x_2, \ldots, x_n)$ whenever $\mathbf{x}^0 = (x_1^0, x_2, \ldots, x_n)$ and $\mathbf{x}^1 = (x_1^1, x_2, \ldots, x_n)$ both belong to E. Since E is convex, the line segment joining $\mathbf{x}^0$ and $\mathbf{x}^1$ is contained in E. The mean-value theorem applies on this line segment, showing that $f(\mathbf{x}^0) - f(\mathbf{x}^1) = (x_1^0 - x_1^1)(D_1 f)(\mathbf{x})$ for some point $\mathbf{x}$ on this interval. Hence the result now follows from the hypothesis.

Note that convexity is needed only on each line segment through E parallel to the x_1-axis. Thus if the intersection of E with each line parallel to the x_1-axis is an interval and $(D_1 f)(\mathbf{x}) = 0$ for all $\mathbf{x} \in E$, then f is independent of x_1.

If we define $f(x,y)$ on all of R^2 except the nonnegative portion of the y-axis by specifying

$$f(x,y) = \begin{cases} 0 & \text{if } y < 0 \text{ or } x < 0, \\ y^2 & \text{if } y \geq 0 \text{ and } x > 0, \end{cases}$$

then $f(x,y)$ is continuously differentiable on its domain, $(D_1 f)(x,y) = 0$ everywhere on that domain, yet $f(-1,1) = 0 \neq 1 = f(1,1)$, so that f is not independent of x.

Exercise 9.11 If f and g are differentiable real functions in R^n, prove that

$$\nabla(fg) = f\nabla g + g\nabla f$$

and that $\nabla(1/f) = -f^{-2}\nabla f$ wherever $f \neq 0$.

Solution. This is a routine computation applied to the ith component of the various quantities.

Exercise 9.12 Fix two real numbers a and b, $0 < a < b$. Define a mapping $\mathbf{f} = (f_1, f_2, f_3)$ of R^2 into R^3 by

$$\begin{aligned} f_1(s,t) &= (b + a\cos s)\cos t \\ f_2(s,t) &= (b + a\cos s)\sin t \\ f_3(s,t) &= a\sin s \end{aligned}$$

Describe the range K of $\mathbf{f}$. (It is a certain compact subset of R^3.)

(a) Show that there are exactly 4 points $\mathbf{p} \in K$ such that

$$(\nabla f_1)(\mathbf{f}^{-1}(\mathbf{p})) = \mathbf{0}.$$

Find these points.

(b) Determine the set of all $\mathbf{q} \in K$ such that

$$(\nabla f_3)(\mathbf{f}^{-1}(\mathbf{q})) = \mathbf{0}.$$

(c) Show that one of the points $\mathbf{p}$ found in part (a) corresponds to a local maximum of f_1, one corresponds to a local minimum, and that the other two are neither (they are so-called "saddle points").

Which of the points $\mathbf{q}$ found in part (b) correspond to maxima or minima?

(d) Let λ be an irrational number, and define $\mathbf{g}(t) = \mathbf{f}(t, \lambda t)$. Prove that $\mathbf{g}$ is a 1-1 mapping of R^1 onto a dense subset of K. Prove that

$$|\mathbf{g}'(t)|^2 = a^2 + \lambda^2(b + a\cos t)^2.$$

Solution. The range K is a torus obtained by moving a circle of radius a with center on a circle of radius b, always keeping the planes of the two circles perpendicular and each plane passing through the center of the other circle. This can be seen by observing that in cylindrical coordinates the parametric equations say $r = b + a\cos s$, $z = a\sin x$, i.e., $(r-b)^2 + z^2 = a^2$, which, together with the equation $\theta = \text{const}$, gives the equation of a circle with center at $(b, 0)$ and radius a in the half-plane $\theta = \text{const}$.

(*a*) The equation $(\nabla f_1)(s,t) = \mathbf{0}$ says $-a\sin s\cos t = 0$ and $-(b + a\cos s)\sin t = 0$. This second equation requires $t = k\pi$, and since these functions have period 2π in both s and t, we may as well assume $t = 0$ or $t = \pi$. In that case the first equation implies $s = 0$ or $s = \pi$. Hence the only points $\mathbf{p}$ satisfying this equation are the images of the points $(0,0)$, $(0,\pi)$, $(\pi,0)$, and (π,π), i.e., the points $(b+a,0,0)$, $(b-a,0,0)$, $(-b+a,0,0)$, and $(-b-a,0,0)$.

(*b*) The equation $(\nabla f_3)(s,t) = \mathbf{0}$ says only that $a\cos s = 0$, i.e., $s = \frac{\pi}{2}$ or $s = \frac{3\pi}{2}$. The image of these two conditions consists of the two circles of radius b about the z-axis in the planes $z = \pm a$.

(*c*) The point $(a+b,0,0)$ is the maximum possible value of $f_1(s,t)$, and occurs only when $\cos s = 1$ and $\cos t = 1$. Likewise the point $(-a-b,0,0)$ is the minimum possible value, and occurs only when $\cos s = 1$ and $\cos t = -1$. The other two points, which occur when $s = 0, t = \pi$ and when $s = \pi, t = 0$, lie near points of both larger and smaller values of $f_1(s,t)$. For example, when $s = 0$, the point $t = \pi$ is a minimum for the function $\varphi(t) = f_1(0,t) = b\cos t$; but when $t = \pi$, the point $s = 0$ is a maximum of $\psi(s) = f_1(s,\pi) = -(b + a\cos s)$. Hence the point $(0,\pi)$ is neither a maximum nor a minimum for $f_1(s,t)$.

The points with $z = +a$ are obviously absolute maxima of $f_3(s,t)$, while those with $z = -a$ are the absolute minima.

(*d*) Suppose $\mathbf{g}(t_1) = \mathbf{g}(t_2)$. Then because $a\sin t_1 = a\sin t_2$, and

$$\sqrt{(f_1(t_1,\lambda t_1))^2 + (f_2(t_1,\lambda t_1))^2} = \sqrt{(f_1(t_2,\lambda t_2))^2 + (f_2(t_2,\lambda t_2))^2}$$

(that is, $b + a\cos t_1 = b + a\cos t_2$), we have $\sin t_1 = \sin t_2$ and $\cos t_1 = \cos t_2$. Therefore $\sin(t_1 - t_2) = 0$, which means $t_2 = t_1 + k\pi$ for some integer k. Because $\sin t_1 = \sin t_2$, it follows that k is an even integer, say $k = 2m$. It then follows, since $f_i(t_1,\lambda t_1) = f_i(t_2,\lambda t_2)$, $\lambda = 1,2$, that $\cos\lambda t_1 = \cos\lambda t_2$ and $\sin\lambda t_1 = \sin\lambda t_2$. This in turn implies that $\lambda t_2 = \lambda t_2 + 2r\pi$ for some integer r. Combining these two results, we find that $m\lambda = r$. Since λ is irrational, this means that $m = 0 = r$, i.e., $t_2 = t_1$. Thus $\mathbf{g}(t)$ is one-to-one.

To show that the range is dense in K, we need only show that the numbers $2\pi n\lambda$, $n = 0 \pm 1, \pm 2, \ldots$, are dense "modulo 2π," meaning that for any real number θ and any $\varepsilon > 0$ there is are integers m and n such that $|2\pi n\lambda - 2\pi m - \theta| < \varepsilon$. A proposition easily seen to be equivalent is that for any $\eta > 0$ and any real number c there exist integers m and n such that $|n\lambda - m - c| < \eta$. (This statement is obvious ($m = n = 0$) if $c = 0$.) To prove that, fix an integer r larger than $\frac{1}{\eta}$, and consider the numbers $0, \lambda - [\lambda], 2\lambda - [2\lambda], \ldots, r\lambda - [r\lambda]$. There are $r+1$ such numbers, all lying in the interval $[0,1)$. Hence two of them must be closer than $\frac{1}{r}$ to each other, say $0 < s\lambda - [s\lambda] - t\lambda + [t\lambda] < \frac{1}{r}$. In particular, the number $(s-t)\lambda$ lies within $\frac{1}{r}$ of an integer (namely $[s\lambda] - [t\lambda]$. Thus we have, say $(s-t)\lambda = k + \delta$, where $0 < \delta < \frac{1}{r}$. Let p be the unique integer such that $p\delta \le c < (p+1)\delta$. We then have $p(s-t)\lambda = pk + p\delta$, and hence, taking $n = p(s-t)$ and $m = pk$, we find $|n\lambda - m - c| = |p\delta - c| < \delta < \frac{1}{r} < \eta$.

This being established, consider any point in K, say the point $\mathbf{p} = ((b + a\cos s_0)\cos t_0, (b + a\cos s_0)\sin t_0, a\sin s_0)$, and let $\varepsilon > 0$ be given. According to what was just established, there are integers m, n such that $|2\pi m\lambda - 2\pi n - (t_0 - s_0\lambda)| < \frac{\varepsilon}{3a+3b}$. It then follows that

$$\begin{aligned} |\cos((s_0 + 2\pi m)\lambda) - \cos t_0| &= |\cos((s_0 + 2\pi m)\lambda - 2\pi n) - \cos t_0| \\ &\le \frac{\varepsilon}{3a+3b}, \end{aligned}$$

where we have used the inequality $|\cos u - \cos v| \le |u - v|$, with $u = (s_0 + 2\pi m)\lambda - 2\pi n$ and $v = t_0$. A similar inequality applies with sin in place of cos. It then follows that $|\mathbf{g}(s_0 + 2\pi m) - \mathbf{p}| \le \frac{2\varepsilon}{3} < \varepsilon$. Therefore the range of $\mathbf{g}$ is dense in K.

The equation

$$|\mathbf{g}'(t)|^2 = a^2 + \lambda^2(b + a\cos t)^2$$

is a routine, though tedious, computation.

Exercise 9.13 Suppose $\mathbf{f}$ is a differentiable mapping of R^1 into R^3 such that $|\mathbf{f}(t)| = 1$ for every t. Prove that $\mathbf{f}'(t) \cdot \mathbf{f}(t) = 0$.

Interpret this result geometrically.

Solution. This result is obtained by merely differentiating the relation $\mathbf{f}(t)\cdot\mathbf{f}(t) = 1$. Geometrically it asserts that the velocity vector of a point moving over a sphere is tangent to the sphere (perpendicular to the radius vector from the center of the sphere to the point).

Exercise 9.14 Define $f(0,0) = 0$ and

$$f(x,y) = \frac{x^3}{x^2 + y^2} \quad \text{if } (x,y) \ne (0,0).$$

(*a*) Prove that $D_1 f$ and $D_2 f$ are bounded functions in R^2. (Hence f is continuous.)
(*b*) Let $\mathbf{u}$ be any unit vector in R^2. Show that the directional derivative $(D_{\mathbf{u}} f)(0,0)$ exists, and that its absolute value is at most 1.
(*c*) Let γ be a differentiable mapping of R^1 into R^2 (in other words, γ is a differentiable curve in R^2), with $\gamma(0) = (0,0)$ and $|\gamma'(0)| > 0$. Put $g(t) = f(\gamma(t))$ and prove that g is differentiable for every $t \in R^1$.

If $\gamma \in C'$, prove that $g \in C'$.

(*d*) In spite of this, prove that f is not differentiable at $(0,0)$.
Hint: Formula (40) fails.

Solution. (*a*) For $(x,y) \neq (0,0)$ we have

$$D_1 f(x,y) = \frac{x^2(x^2+3y^2)}{(x^2+y^2)^2}, \quad D_2 f(x,y) = -\frac{2x^3 y}{(x^2+y^2)^2}.$$

It follows that

$$0 \leq D_1 f(x,y) \leq \frac{3x^2}{x^2+y^2} \leq 3$$

and

$$|D_2 f(x,y)| \leq \frac{x^2}{x^2+y^2} \leq 1.$$

Also $D_1 f(0,0) = \lim_{x\to 0} \frac{f(x,0)-f(0,0)}{x} = \lim_{x\to 0} \frac{x-0}{x} = 1$, and $D_2 f(0,0) = \lim_{y\to 0} \frac{f(0,y)-f(0,0)}{y} = \lim_{y\to 0} \frac{0}{y} = 0$. Hence, as asserted, $f(x,y)$ is continuous.

(*b*) Let $\mathbf{u} = (\cos\theta, \sin\theta)$. Then $D_{\mathbf{u}} f(0,0) = \lim_{t\to 0} \frac{f(t\cos\theta, t\sin\theta) - f(0,0)}{t} = \cos^3\theta$.

(*c*) Suppose $u(t)$ and $v(t)$ satisfy $u(0) = 0 = v(0)$, $u'(t)$ and $v'(t)$ exist for every t, and $u'(t)$ and $v'(t)$ do not both vanish at the same value of t. Setting $g(t) = f(u(t), v(t))$, we find that $g(t)$ is obviously differentiable at any value of t where $u(t)$ and $v(t)$ are not both zero. Now suppose $u(t_0) = v(t_0) = 0$. Then, since one of $u(t)$ and $v(t)$ is one-to-one on a neighborhood of t_0, it follows that, for small non-zero values of $t - t_0$ we have $(u(t))^2 + (v(t))^2 > 0$, and then

$$\begin{aligned} \frac{g(t)-g(t_0)}{t-t_0} &= \frac{f(u(t),v(t)) - f(u(t_0),v(t_0))}{t-t_0} \\ &= \frac{\left(\frac{u(t)-u(t_0)}{t-t_0}\right)^3}{\left(\frac{u(t)-u(t_0)}{t-t_0}\right)^2 + \left(\frac{v(t)-v(t_0)}{t-t_0}\right)^2}, \end{aligned}$$

so that

$$g'(t_0) = \lim_{t\to t_0} \frac{g(t)-g(t_0)}{t-t_0} = \frac{(u'(t_0))^3}{(u'(t_0))^2 + (v'(t_0))^2}.$$

Thus $g(t)$ is differentiable. Observe that if $\gamma(t) \neq (0,0)$, then

$$g'(t) = \frac{(u(t))^4 u'(t) + 3(u(t)v(t))^2 u'(t) - 2(u(t))^3 v(t) v'(t)}{((u(t))^2 + (v(t))^2)^2}.$$

The same argument used above to prove that $g'(t_0)$ exists shows that

$$\lim_{t\to t_0} g'(t) = \frac{(u'(t_0))^5 + (u'(t_0))^3 (v'(t_0))^2}{((u'(t_0))^2 + (v'(t_0))^2)^2} = \frac{(u'(t_0))^3}{(u'(t_0))^2 + (v'(t_0))^2} = g'(t_0),$$

so that g' is continuous at t_0 if u' and v' are. Continuity of g' at other points follows from the chain rule.

If f is differentiable at $(0,0)$, we necessarily have

$$f(x,y) = f(0,0) + \left[xD_1f(0,0) + yD_2f(0,0)\right] + \varepsilon(x,y),$$

where

$$\lim_{(x,y)\to(0,0)} \frac{\varepsilon(x,y)}{\sqrt{x^2+y^2}} = 0.$$

Since $D_1f(0,0) = 1$ and $D_2f(0,0) = 0$, it follows that

$$\varepsilon(x,y) = \frac{-xy^2}{x^2+y^2},$$

and so we must have

$$\lim_{(x,y)\to(0,0)} \frac{-xy^2}{(x^2+y^2)^{3/2}} = 0.$$

But this is clearly not the case, as we see by taking $x = y$. (The limit is then $-2^{-3/2}$.)

Exercise 9.15 Define $f(0,0) = 0$, and put

$$f(x,y) = x^2 + y^2 - 2x^2y - \frac{4x^6y^2}{(x^4+y^2)^2}$$

if $(x,y) \neq (0,0)$.

(a) Prove, for all $(x,y) \in R^2$, that

$$4x^4y^2 \le (x^4+y^2)^2.$$

Conclude that f is continuous.

(b) For $0 \le \theta \le 2\pi$, $-\infty < t < \infty$, define

$$g_\theta(t) = f(t\cos\theta, t\sin\theta).$$

Show that $g_\theta(0) = 0$, $g_\theta'(0) = 0$, $g_\theta''(0) = 2$. Each g_θ has therefore a strict local minimum at $t = 0$.

In other words, the restriction of f to each line through $(0,0)$ has a strict local minimum at $(0,0)$.

(c) Show that $(0,0)$ is nevertheless not a local minimum for f, since $f(x,x^2) = -x^4$.

Solution. (a) This inequality follows by squaring the inequality $2x^2|y| \le x^4+y^2$, which in turn is equivalent to the inequality $(x^2-|y|)^2 \ge 0$. Then, since $f(x,y)$

is obviously continuous except at $(0,0)$, the continuity at the remaining point follows from the inequality

$$|f(x,y)-f(0,0)| \leq 2x^2+y^2+2x^2|y|,$$

which is easily derived from the inequality just proved and the definition of $f(x,y)$.

(b) We observe that for $t \neq 0$ we have

$$g_\theta(t)=t^2-2t^3\cos^2\theta\sin\theta-4t^4\frac{\cos^6\theta\sin^2\theta}{(t^2\cos^4\theta+\sin^2\theta)^2},$$

from which it is routine computation to show that $g_\theta(0)=0=g'_\theta(0)$ and $g''_\theta(0)=2$.

(c) The assertion that $f(x,x^2)=-x^4$ is routine computation. It implies that $f(x,y)$ assumes negative values in any neighborhood of $(0,0)$, and hence that the $f(x,y)$ does not have a local minimum at $(0,0)$.

Exercise 9.16 Show that the continuity of $\mathbf{f}'$ at the point $\mathbf{a}$ is needed in the inverse function theorem, even in the case $n=1$: If

$$f(t)=t+2t^2\sin\left(\frac{1}{t}\right)$$

for $t \neq 0$, and $f(0)=0$, then $f'(0)=1$, f' is bounded in $(-1,1)$ but f is not one-to-one in any neighborhood of 0.

Solution. The assertion that $f'(0)=1$ is proved by direct computation: $\dfrac{f(t)}{t}=1+2t\sin\left(\frac{1}{t}\right)\to 1$ as $t\to 0$. Since $f'(t)=1+4t\sin\left(\frac{1}{t}\right)-2\cos\left(\frac{1}{t}\right)$ for $t\neq 0$, it follows that $|f'(t)|\leq 7$ for all $t\in(-1,1)$. To show that f is not one-to-one in any neighborhood of 0, we observe that $f'(\frac{1}{k\pi})=1+2(-1)^k$, so that $f(t)$ is decreasing at $t=\frac{1}{k\pi}$ if k is odd and increasing if k is even. It follows that the minimum value of $f(t)$ on the interval $\left[\frac{1}{(2k+1)\pi},\frac{1}{2k\pi}\right]$ is assumed at an interior point, so that $f(t)$ cannot be one-to-one on this interval.

Exercise 9.17 Let $\mathbf{f}=(f_1,f_2)$ be the mapping of R^2 into R^2 given by

$$f_1(x,y)=e^x\cos y, \quad f_2(x,y)=e^x\sin y.$$

(a) What is the range of $\mathbf{f}$?

(b) Show that the Jacobian of $\mathbf{f}$ is not zero at any point of R^2. Thus every point of R^2 has a neighborhood in which $\mathbf{f}$ is one-to-one. Nevertheless, $\mathbf{f}$ is not one-to-one on R^2.

(c) Put $\mathbf{a}=(0,\pi/3)$, $\mathbf{b}=f(\mathbf{a})$, let $\mathbf{g}$ be the continuous inverse of $\mathbf{f}$, defined in a neighborhood of $\mathbf{b}$, such that $\mathbf{g}(\mathbf{b})=\mathbf{a}$. Find an explicit formula for $\mathbf{g}$, compute $\mathbf{f}'(\mathbf{a})$ and $\mathbf{g}'(\mathbf{b})$, and verify the formula (52).

(*d*) What are the images under $\mathbf{f}$ of lines parallel to the coordinate axes?

Solution. (*a*) The range of $\mathbf{f}$ is all of R^2 except the point $(0,0)$. Indeed if $(u,v) \neq (0,0)$, choose y so that

$$\cos y = \frac{u}{\sqrt{u^2+v^2}}, \quad \sin y = \frac{v}{\sqrt{u^2+v^2}},$$

and let $x = \ln\sqrt{u^2+v^2}$, so that $e^x = \sqrt{u^2+v^2}$. It is then obvious from the equations defining y and x that $u = e^x \cos y$ and $v = e^x \sin y$. Hence every point except $(0,0)$ is in the range of $\mathbf{f}$. The point $(0,0)$ is not in the range, since $u^2+v^2 = e^{2x} > 0$ for any point $(u,v) = \mathbf{f}(x,y)$.

(*b*) The Jacobian of $\mathbf{f}(x,y)$ is e^{2x}, which is never zero. However, since $\mathbf{f}(x,y+2\pi) = \mathbf{f}(x,y)$, it follows that $\mathbf{f}$ is not one-to-one.

(*c*) By our definition $\mathbf{b} = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$. We can therefore take $y = \arctan\left(\frac{v}{u}\right)$ for (u,v) near $\mathbf{b}$, the arctangent being between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$. Thus we have $\mathbf{g}(u,v) = \left(\ln\sqrt{u^2+v^2}, \arctan\left(\frac{v}{u}\right)\right)$. We then have

$$\mathbf{f}'(x,y) = \begin{pmatrix} e^x\cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}, \quad \mathbf{g}'(u,v) = \begin{pmatrix} \frac{u}{u^2+v^2} & \frac{v}{u^2+v^2} \\ \frac{-v}{u^2+v^2} & \frac{u}{u^2+v^2} \end{pmatrix}.$$

When we take $u = e^x \cos y$ and $v = e^x \sin y$, we find that

$$\mathbf{g}'(\mathbf{f}(x,y)) = \begin{pmatrix} e^{-x}\cos y & e^{-x}\sin y \\ -e^{-x}\sin y & e^{-x}\cos y \end{pmatrix}.$$

It is then a routine computation to verify that $\mathbf{g}'(\mathbf{f}(x,y))\mathbf{f}'(x,y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Likewise we find

$$\mathbf{f}'(\mathbf{g}(u,v)) = \begin{pmatrix} u & -v \\ v & u \end{pmatrix},$$

and a routine computation shows that $\mathbf{f}'(\mathbf{g}(u,v))\mathbf{g}'(u,v) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

(*d*) The family of lines $x = c$ maps to the family of concentric circles $u^2+v^2 = e^{2c}$. The lines $y = c$ map to half-lines $v = Ku$, $u \geq 0$, where $K = \tan y$. (If y is an odd multiple of $\frac{\pi}{2}$, the half-line is either the positive or negative u-axis.

Exercise 9.18 Answer analogous questions for the mapping defined by

$$u = x^2 - y^2, \quad v = 2xy.$$

Solution. (*a*) the range of the mapping $\mathbf{f}(x,y) = (x^2-y^2, 2xy)$ is the entire plane R^2. Indeed, every point (u,v) except $(0,0)$ has two distinct preimages, one of which is

$$x = \sqrt{\frac{\sqrt{u^2+v^2}+u}{2}}, \quad y = (\operatorname{sgn} v)\sqrt{\frac{\sqrt{u^2+v^2}-u}{2}}.$$

(The other preimage is $-x$, $-y$, with this x and this y.)
(*b*) The Jacobian of $\mathbf{f}$ vanishes only at $x = y = 0$. Indeed,

$$\mathbf{f}'(x,y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}.$$

Hence the Jacobian is $4(x^2+y^2)$.
(*c*) Taking $\mathbf{a} = (3,4)$, so that $\mathbf{b} = (-7,24)$, we can take, locally

$$g(u,v) = \left(\sqrt{\frac{\sqrt{u^2+v^2}+u}{2}}, \sqrt{\frac{\sqrt{u^2+v^2}-u}{2}}\right).$$

We then have

$$\mathbf{g}'(u,v) = \begin{pmatrix} \frac{1}{4}\sqrt{\frac{2}{\sqrt{u^2+v^2}+u}}\left(1+\frac{u}{\sqrt{u^2+v^2}}\right) & \frac{1}{4}\sqrt{\frac{2}{\sqrt{u^2+v^2}+u}}\left(\frac{v}{\sqrt{u^2+v^2}}\right) \\ \frac{1}{4}\sqrt{\frac{2}{\sqrt{u^2+v^2}-u}}\left(-1+\frac{u}{\sqrt{u^2+v^2}}\right) & \frac{1}{4}\sqrt{\frac{2}{\sqrt{u^2+v^2}-u}}\left(\frac{v}{\sqrt{u^2+v^2}}\right) \end{pmatrix}$$

Noting that the defining relations imply $u^2+v^2 = (x^2+y^2)^2$, we see that

$$\mathbf{g}'(\mathbf{f}(x,y)) = \begin{pmatrix} \frac{x}{2(x^2+y^2)} & \frac{y}{2(x^2+y^2)} \\ \frac{-y}{2(x^2+y^2)} & \frac{x}{2(x^2+y^2)} \end{pmatrix},$$

from which we see easily that $\mathbf{g}'(\mathbf{f}(x,y))\mathbf{f}'(x,y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. The corresponding equality with $\mathbf{g}$ and $\mathbf{f}$ interchanged is likewise simple, though more cumbersome to write out.

Exercise 9.19 Show that the system of equations

$$\begin{aligned} 3x+y-z+u^2 &= 0 \\ x-y+2z+u &= 0 \\ 2x+2y-3z+2u &= 0 \end{aligned}$$

can be solved for x, y, u in terms of z; for x, z, u in terms of y; for y, z, u in terms of x; but not for x, y, z in terms of u.

Solution. Adding the last two equations and subtracting the first yields $3u - u^2 = 0$, whence either $u = 0$ or $u = 3$. Hence unless u has one of these two values, there are no solutions at all. Therefore the system cannot generally be solved for x, y, z in terms of u. If one of these two equations holds, we can solve just the last two equations for any two of the variables x, y, z in terms of the third. The remaining equation will then automatically be satisfied. For example,

$$x = -\frac{z}{4},\ y = \frac{7z}{4},\ u = 0; \quad x = -\frac{9+z}{4},\ y = \frac{3+7z}{4},\ u = 3.$$

We could also have

$$x = -\frac{y}{7}, \quad z = \frac{4y}{7}, \quad u = 0; \quad x = \frac{60+4y}{7}, \quad z = \frac{4y-3}{7}, \quad u = 3.$$

Finally, we could also have

$$y = -7x, \quad z = -4x, \quad u = 0; \quad y = \frac{7x-60}{4}, \quad z = 9 - 4x, \quad u = 3.$$

Note that the matrix of the derivative of the transformation $\mathbf{f}(x, y, z, u) = (3x + y - z + u^2, x - y + 2z + u, 2x + 2y - 3z + 2u)$ is

$$\mathbf{f}'(x, y, z, u) = \begin{pmatrix} 3 & 1 & -1 & 2u \\ 1 & -1 & 2 & 1 \\ 2 & 2 & -3 & 2 \end{pmatrix}$$

and any 3×3 submatrix containing the last column is invertible when $u = 0$ or $u = 3$. However, the first three columns of this matrix does not form an invertible matrix.

Exercise 9.20 Take $n = m = 1$ in the implicit function theorem, and interpret the theorem (as well as its proof) graphically.

Solution. The theorem asserts that if $f(x, y)$ is continuously differentiable in a neighborhood of (x_0, y_0), $f(x_0, y_0) = 0$, and $D_2 f(x_0, y_0) \neq 0$, then there exist 1) an interval $I = (x_0 - \delta, x_0 + \delta)$, 2) an interval $J = (y_0 - \eta, y_0 + \eta)$, and 3) a continuously differentiable function $\varphi : I \longrightarrow J$ such that for all $(x, y) \in I \times J$ the equation $f(x, y) = 0$ holds if and only if $y = \varphi(x)$.

The proof amounts to the argument that, since $D_2 f(x_0, y_0) \neq 0$ and f is continuously differentiable, it must be that $D_2 f(x, y) \neq 0$ for all (x, y) near (x_0, y_0). Hence the function $g(y) = f(x_0, y)$ is strictly monotonic near $y = y_0$. Therefore, since $g(y_0) = 0$, there is a small interval $[y_0 - \eta, y_0 + \eta]$ such that $g(y_0 - \eta)$ and $g(y_0 + \eta)$ have opposite signs. By the continuity of $f(x, y)$, it follows that $f(x, y_0 - \eta)$ has the same sign as $f(x_0, y_0 - \eta)$ if x is near x_0, and similarly $f(x, y_0 + \eta)$ has the same sign as $f(x_0, y_0 + \eta)$ for x near x_0. That is, $f(x, y_0 - \eta)$ and $f(x, y_0 + \eta)$ have opposite signs if x is near x_0. It follows that there is a point $\varphi(x) \in (y_0 - \eta, y_0 + \eta)$ such that $f(x, \varphi(x)) = 0$. By restricting the neighborhood so that $D_2 f(x, y)$ is of constant sign, we assure that $g_x(y) = f(x, y)$ is monotonic on $[y_0 - \eta, y_0 + \eta]$ for each x near x_0. It then follows that there can be at most one value of y in $(y_0 - \eta, y_0 + \eta)$ satisfying the equation $f(x, y) = 0$. That is, the function $\varphi(x)$ is unique. This proves all but the differentiability of φ.

The graphical interpretation is that, near a point on a smooth curve $f(x, y) = 0$ where the tangent is not vertical ($D_2 f(x_0, y_0) \neq 0$) the curve intersects each vertical line exactly once.

Exercise 9.21 Define f in R^2 by

$$f(x,y) = 2x^3 - 3x^2 + 2y^3 + 3y^2.$$

(a) Find the four points in R^2 at which the gradient of f is zero. Show that f has exactly one local maximum and one local minimum in R^2.
(b) Let X be the set of all $(x,y) \in R^2$ at which $f(x,y) = 0$. Find those points of S that have no neighborhoods in which the equation $f(x,y) = 0$ can be solved for y in terms of x (or for x in terms of y). Describe S as precisely as you can.

Solution. (a) We have $\nabla f(x,y) = 6(x^2 - x)\mathbf{i} + 6(y^2 + y)\mathbf{j}$. Hence $\nabla f(x,y) = \mathbf{0}$ precisely at the four points $(0,0)$, $(1,0)$, $(0,-1)$, $(1,-1)$. Since the Hessian matrix of f is

$$\begin{pmatrix} 12x-2 & 0 \\ 0 & 12y+2 \end{pmatrix}$$

this matrix has a positive determinant when $x > \frac{1}{6}$ and $y > \frac{-1}{6}$ or when $x < \frac{1}{6}$ and $y < \frac{-1}{6}$. Thus $(1,0)$ and $(0,-1)$ are possible extrema. Since $12x - 2 > 0$ at $(1,0)$, that point is a minimum. Likewise $(0,-1)$ is a maximum.

(b) Since $f(x,y) = (x+y)[2x^2 - 2xy + 2y^2 - 3x + 3y]$, the equation $f(x,y) = 0$ has the real solution $y = -x$ for every real value of x. In addition, if $-\frac{1}{2} \le x \le \frac{3}{2}$, it has the real solutions

$$y = \frac{2x - 3 + \sqrt{9 + 12x - 12x^2}}{4}, \quad y = \frac{2x - 3 - \sqrt{9 + 12x - 12x^2}}{4}.$$

According to the implicit function theorem, the only possible points near which there might not be a unique solution are for y in terms of x are those where $y = 0$ or $y = -1$. The corresponding values of x are $x = 0$ and $x = \frac{3}{2}$ for $y = 0$ and $x = 1$ and $x = -\frac{1}{2}$ for $y = -1$.

We observe that both solutions $y = -x$ and $y = \dfrac{2x - 3 + \sqrt{9 + 12x - 12x^2}}{4}$ tend to 0 as $x \to 0$. Hence there is no unique solution for y near $(0,0)$. As $x \uparrow \frac{3}{2}$, the quantity under the radical sign tends to zero, and hence these two solutions converge toward the common value $y = 0$. Hence the point $\left(\frac{3}{2}, 0\right)$, is another point around which the solution for y is not unique. The two radicals also tend to zero as $x \downarrow -\frac{1}{2}$, causing the two values of y both to tend toward -1, so that $\left(-\frac{1}{2}, -1\right)$ is not a point of unique solvability. Finally, as $x \to 1$, the three y values tend toward -1, $\frac{1}{2}$, and -1. Since two of these values are identical, there is no unique solution around the point $(1,-1)$.

Finally, the three x-values corresponding to any y are

$$x = -y, \quad x = \frac{2y + 3 \pm \sqrt{9 - 12y - 12y^2}}{4},$$

where the quantity under the radical is nonnegative in the range $-\frac{3}{2} \le y \le \frac{1}{2}$. The values where $D_1 f(x,y) = 0$ are $x = 0$ and $x = 1$, and the four points near which a solution for x might not be unique are $(0,0)$, $\left(0, -\frac{3}{2}\right)$, $(1,-1)$, and

$(1, \frac{1}{2})$. As y tends to zero, two of these tend to zero. Hence $(0,0)$ is not a point of unique solvability for x in terms of y. As y tends to -1, two of the x-values tend to 1, so that $(1,-1)$ is not a point of unique solvability for x. Finally, as y tends to $-\frac{3}{2}$ or $\frac{1}{2}$, the radical disappears, and so once again two of the x values tend to the same value, namely 1 as $y \to \frac{1}{2}$ and 0 as $y \to -\frac{3}{2}$. Thus these four points are not points of unique solvability for x.

In sum, the points near which the equation $f(x,y) = 0$ does not define either y as a function of x or x as a function of y are $(0,0)$ and $(1,-1)$.

Exercise 9.22 Give a similar discussion for

$$f(x,y) = 2x^3 + 6xy^2 - 3x^2 + 3y^2.$$

Solution. The gradient is

$$\nabla 6(x^2 + y^2 - x)\mathbf{i} + 6(2xy + y)\mathbf{j}$$

As we see from solving the appropriate equations, this gradient vanishes at the points $(0,0)$ and $(1,0)$. The point $(0,0)$ is a saddle point, since $f(x,0)$ is negative for $x < 0$ and $f(0,y)$ is positive for y near zero. The Hessian determinant is positive at $(1,0)$, and the upper left-hand entry is also; hence $(1,0)$ is a minimum.

Because the equation $f(x,y) = 0$ can be written as

$$(6x + 3)y^2 = (3 - 2x)x^2,$$

there will be real solutions y if and only if $-\frac{1}{2} < x \le \frac{3}{2}$. (When $x = -\frac{1}{2}$, the equation does not contain y.) In this range there are two distinct values of y except for $x = 0$ and $x = \frac{3}{2}$. Hence the two points on the locus of $f(x,y) = 0$ at which the equation cannot be solved for y are $(0,0)$ and $(\frac{3}{2}, 0)$.

Since the equation is cubic in x, its solvability is more complicated from this point of view. Every value of y gives at least one value of x (but those x-values always lie between $-\frac{1}{2}$ and $\frac{3}{2}$). To find the points where two of the three (complex) x-roots coincide, we observe that at such points $D_1 f(x,y) = 0$, and hence also $3f(x,y) - xD_1 f(x,y) = 0$. This last equation says $x^2 - 4xy^2 + 3y^2 = 0$, i.e., $y^2 = \dfrac{x^2}{4x+3}$. Substituting this value of y^2 into $f(x,y) = 0$, we get either $x = 0$ and $y = 0$ or

$$x^2 = \frac{3}{4}.$$

Since we have to have $-\frac{1}{2} < x$, we must have $x = \frac{\sqrt{3}}{2}$, and this gives $y^2 = \frac{2\sqrt{3}-3}{4}$. Hence the points near which $f(x,y) = 0$ cannot be solved uniquely for x are $(0,0)$ and $(\frac{\sqrt{3}}{2}, \pm\frac{\sqrt{2\sqrt{3}-3}}{4})$.

Exercise 9.23 Define f in R^3 by

$$f(x, y_1, y_2) = x^2 y_1 + e^x + y_2.$$

Show that $f(0,1,-1) = 0$, $(D_1 f)(0,1,-1) \neq 0$, and that there exists therefore a differentiable function g in some neighborhood of $(1,-1)$ in R^2 such that $g(1,-1) = 0$ and

$$f(g(y_1, y_2), y_1, y_2) = 0.$$

Find $(D_1 g)(1,-1)$ and $(D_2 g)(1,-1)$.

Solution. The proof that $f(0,1,-1) = 0$ is a routine computation. We have $(D_1 f)(x, y_1, y_2) = 2xy_1 + e^x$, so that $(D_1 f)(0,1,-1) = 1 \neq 0$. To find the partial derivatives of g we use the chain rule. Let $\psi(y_1, y_2) = f(g(y_1, y_2), y_1, y_2) \equiv 0$. Then

$$0 = D_1\psi(y_1, y_2) = D_1 f(g(y_1, y_2), y_1, y_2) D_1 g(y_1, y_2) + D_2 f(g(y_1, y_2), y_1, y_2),$$

so that

$$0 = \left(2y_1 g(y_1, y_2) + e^{g(y_1, y_2)}\right) D_1 g(y_1, y_2) + \left(g(y_1, y_2)\right)^2.$$

Similarly, setting

$$0 = D_2\psi(y_1, y_2) = D_1 f(g(y_1, y_2), y_1, y_2) D_2 g(y_1, y_2) + D_3 f(g(y_1, y_2), y_1, y_2),$$

we find

$$0 = \left(2y_1 g(y_1, y_2) + e^{g(y_1, y_2)}\right) D_2 g(y_1, y_2) + 1.$$

Taking $y_1 = 1$, $y_2 = -1$, $g(y_1, y_2) = 0$, we get

$$D_1 g(1,-1) = 0, \quad D_2 g(1,-1) = -1.$$

Exercise 9.24 For $(x, y) \neq (0,0)$, define $\mathbf{f} = (f_1, f_2)$ by

$$f_1(x, y) = \frac{x^2 - y^2}{x^2 + y^2}, \quad f_2(x, y) = \frac{xy}{x^2 + y^2}.$$

Compute the rank of $\mathbf{f}'(x, y)$, and find the range of $\mathbf{f}$.

Solution. The matrix of $\mathbf{f}'(x, y)$ is

$$\begin{pmatrix} \dfrac{4xy^2}{(x^2+y^2)^2} & \dfrac{-4x^2y}{(x^2+y^2)^2} \\ \dfrac{y(y^2-x^2)}{(x^2+y^2)^2} & \dfrac{x(x^2-y^2)}{(x^2+y^2)^2} \end{pmatrix}.$$

Its determinant is 0 at every point. Hence its rank is either 0 or 1 at every point. Since the point $(0,0)$ is excluded from the domain, the rank is 1 at every point. The range must therefore be 1-dimensional, i.e., there is some non-trivial relation connecting f_1 and f_2. Indeed, it is easy to verify that if $u = f_1(x, y)$ and $v = f_2(x, y)$, then

$$u^2 + 4v^2 = 1.$$

Thus the range of $\mathbf{f}$ is a subset of this ellipse. In fact, it is all of this ellipse. The point $(1, 0)$ is its own image, and the point $(-1, 0)$ is the image of $(0, 1)$. For any other point (u, v) on this ellipse we have $-1 < u < 1$ and $v = \pm\frac{1}{2}\sqrt{1 - u^2}$. The point (u, v) is the image of the point $\left(1, \pm\sqrt{\frac{1-u}{1+u}}\right)$ (and, of course, many other points as well).

Exercise 9.25 Suppose $A \in L(R^n, R^m)$, let r be the rank of A.
(a) Define S as in the proof of Theorem 9.32. Show that SA is a projection in R^n whose nullspace is $\mathcal{N}(A)$ and whose range is $\mathcal{R}(S)$. *Hint:* By (68), $SASA = SA$.
(b) Use (a) to show that

$$\dim \mathcal{N}(A) + \dim \mathcal{R}(A) = n.$$

Solution. We recall that S is defined by first choosing a basis for the range of A, say $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$, then choosing vectors $\{\mathbf{z}_1, \ldots, \mathbf{z}_r\}$ such that $A\mathbf{z}_i = \mathbf{y}_i$ for $i = 1, 2, \ldots, r$. We then define $S\mathbf{y}_i = \mathbf{z}_i$ on the vectors $\mathbf{y}_i$ (and S arbitrary on any set of vectors $\mathbf{y}_{r+1}, \ldots, \mathbf{y}_m$ that can be adjoined to $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$ so as to make a basis of R^m). Thus S is a left inverse of the restriction of A to the subspace spanned by $\mathbf{z}_1, \ldots, \mathbf{z}_r$. Since $A\mathbf{x}$ belongs to the range of A, it follows, as in (68), that $ASA\mathbf{x} = A\mathbf{x}$, from which we conclude that $SASA\mathbf{x} = SA\mathbf{x}$, i.e., SA is a projection. Then every vector $\mathbf{x}$ has the unique decomposition $\mathbf{x} = SA\mathbf{x} + (\mathbf{x} - SA\mathbf{x})$, where the first vector on the right belongs to the range of SA and the second to the nullspace of this projection. The two subspaces have only the zero vector in common. Since S is an isomorphism of the range of A, the range of SA has the same dimension as the range of A. Since $A = ASA$, the nullspace of SA is the same as the nullspace of A. Thus $n = \dim \mathcal{N}(SA) + \dim \mathcal{R}(SA) = \dim \mathcal{N}(A) + \dim \mathcal{R}(A)$.

Exercise 9.26 Show that the existence (and even the continuity) of $D_{12}f$ does not imply the existence of $D_1 f$. For example, let $f(x, y) = g(x)$, where g is nowhere differentiable.

Solution. The second sentence in the exercise *is* its solution. Since $D_2 f$ is identically zero, $D_{12}f$ is also identically zero, hence certainly continuous.

Exercise 9.27 Put $f(0, 0) = 0$, and

$$f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}$$

if $(x,y) \neq (0,0)$. Prove that
(*a*) f, $D_1 f$, and $D_2 f$ are continuous in R^2;
(*b*) $D_{12}f$ and $D_{21}f$ exist at every point of R^2, and are continuous except at $(0,0)$;
(*c*) $(D_{12}f)(0,0) = 1$, and $(D_{21}f)(0,0) = -1$.

Solution. (*a*) The continuity of f is obvious at every point except $(0,0)$; at $(0,0)$ it follows from the inequality $|f(x,y)| \le \frac{1}{2}(x^2+y^2)$. It is also clear that $D_1 f(0,0) = 0 = D_2 f(0,0)$. For $(x,y) \neq (0,0)$ we have $D_1 f(x,y) = \dfrac{x^4y + 4x^2y^3 - y^5}{(x^2+y^2)^2}$ and $D_2 f(x,y) = \dfrac{x^5 - 4x^3y^2 - xy^4}{(x^2+y^2)^2}$. The continuity of the partial derivatives at every point except $(0,0)$ is obvious. It is easy to see that these derivatives satisfy the inequalities $|D_1 f(x,y)| \le 2|y|$ and $|D_2 f(x,y)| \le 2|x|$, so that $D_1 f$ and $D_2 f$ are also continuous at $(0,0)$.

(*b*) Since $f(x,y)$ is a rational function with non-zero denominator for $(x,y) \neq (0,0)$, it has continuous partial derivatives of all orders on this set.

(*c*) Since $D_1 f(0,y) = -y$ and $D_2 f(x,0) = x$, it follows that $D_{21}f(0,y) = -1$ for all y and $D_{12}f(x,0) = 1$ for all x.

Exercise 9.28 For $t \ge 0$ put

$$\varphi(x,t) = \begin{cases} x & (0 \le x \le \sqrt{t}) \\ -x + 2\sqrt{t} & (\sqrt{t} \le x \le 2\sqrt{t}) \\ 0 & (\text{otherwise}), \end{cases}$$

and put $\varphi(x,t) = -\varphi(x,|t|)$ if $t < 0$.

Show that φ is continuous on R^2, and

$$(D_2\varphi)(x,0) = 0$$

for all x. Define

$$f(t) = \int_{-1}^{1} \varphi(x,t)\, dx.$$

Show that $f(t) = t$ if $|t| < \frac{1}{4}$. Hence

$$f'(0) \neq \int_{-1}^{1} (D_2\varphi)(x,0)\, dx.$$

Solution. This function is zero in the (closed) left half-plane of the xt-plane and on the positive x-axis. Since the functions by which it is defined are continuous, we need only verify that they agree on the boundary curves $x = \sqrt{t}$ and $x = 2\sqrt{t}$ in the first quadrant that separate the three different regions of definition. This is a routine computation.

Likewise the computation showing that $(D_2\varphi)(x,0) = 0$ is routine, since for each $x > 0$ $\varphi(x,t) = 0$ for $0 \le t \le \frac{1}{4}x^2$, while $\varphi(x,t) = 0$ for all t if $x \le 0$.

If $0 < t < \frac{1}{4}$, then

$$\begin{aligned} f(t) &= \int_0^{\sqrt{t}} x\,dx + \int_{\sqrt{t}}^{2\sqrt{t}} -x + 2\sqrt{t}\,dt \\ &= \frac{1}{2}t - \frac{1}{2}(4t - t) + 2\sqrt{t}(2\sqrt{t} - \sqrt{t}) \\ &= \frac{t}{2} - \frac{3t}{2} + 4t - 2t = t. \end{aligned}$$

Obviously $f(0) = 0$, and if $t < 0$, then $f(t) = -f(-t) = t$. Therefore $f'(0) = 1$. However

$$\int_{-1}^{1} (D_2\varphi)(x, 0)\,dx = 0.$$

Note: This result is possible only because $D_2\varphi(x,t)$ is not bounded on $[-1, 1] \times [-a, a]$ for any $a > 0$. Also note that having -1 as the lower limit of the integral was a needless complication. The problem would have been more effective it the lower limit had been 0.

Exercise 9.29 Let E be an open set in R^n. The classes $\mathcal{C}'(E)$ and $\mathcal{C}''(E)$ are defined in the text. By induction $\mathcal{C}^{(k)}(E)$ can be defined as follows for all positive integers k: To say that $f \in \mathcal{C}^{(k)}(E)$ means that the partial derivatives $D_1 f, \ldots, D_n f$ belong to $\mathcal{C}^{(k-1)}(E)$.

Assume $f \in \mathcal{C}^{(k)}(E)$, and show (by repeated application of Theorem 9.41) that the kth-order derivative

$$D_{i_1 i_2 \ldots i_k} f = D_{i_1} D_{i_2} \ldots D_{i_k} f$$

is unchanged if the subscripts $i_1, \ldots, i_k$ are permuted.

For instance, if $n \geq 3$, then

$$D_{1213} f = D_{3112} f$$

for every $f \in \mathcal{C}^{(4)}$.

Solution. If the permutation leaves i_k fixed, this follows from the result for $k-1$ applied to $D_{i_k} f$. To get the general result, we observe that by the case $k = 2$ we have $D_{i_{k-1} i_k} f = D_{i_k i_{k-1}} f$. Hence the result holds for any permutation that maps i_{k-1} to i_k. But any permutation that maps i_j to i_k ($j \neq k, k-1$) can be written as the composition of a permutation that maps i_j to i_{k-1}, leaving i_k fixed, followed by the interchange of i_{k-1} and i_k, followed by a second permutation that leaves i_k fixed. Therefore the result applies to all permutations whatsoever.

Exercise 9.30 Let $f \in \mathcal{C}^{(m)}(E)$, where E is an open subset of R^n. Fix $\mathbf{a} \in E$, and suppose $\mathbf{x} \in R^n$ is so close to $\mathbf{0}$ that the points

$$\mathbf{p}(t) = \mathbf{a} + t\mathbf{x}$$

lie in E whenever $0 \leq t \leq 1$. Define

$$h(t) = f(\mathbf{p}(t))$$

for all $t \in R^1$ for which $\mathbf{p}(t) \in E$.
(a) For $1 \leq k \leq m$, show (by repeated application of the chain rule) that

$$h^{(k)}(t) = \sum (D_{i_1 \ldots, i_k} f)(\mathbf{p}(t)) x_{i_1} \ldots x_{i_k}.$$

The sum extends over all ordered k-tuples $(i_1, \ldots, i_k)$ in which each i_j is one of the integers $1, \ldots, n$.
(b) By Taylor's theorem (5.15)

$$h(1) = \sum_{k=0}^{m-1} \frac{h^{(k)}(0)}{k!} + \frac{h^{(m)}(t)}{m!}$$

for some $t \in (0, 1)$. Use this to prove Taylor's theorem in n variables by showing that the formula

$$f(\mathbf{a} + \mathbf{x}) = \sum_{k=0}^{m-1} \frac{1}{k!} \sum (D_{i_1 \ldots i_k} f)(\mathbf{a}) x_{i_1} \ldots x_{i_k} + r(\mathbf{x})$$

represents $f(\mathbf{a} + \mathbf{x})$ as the sum of its so-called "Taylor polynomial of degree $m - 1$," plus a remainder that satisfies

$$\lim_{\mathbf{x} \to \mathbf{0}} \frac{r(\mathbf{x})}{|\mathbf{x}|^{m-1}} = 0.$$

Each of the inner sums extends over all ordered k-tuples $(i_1, \ldots, i_k)$, as in part (a); as usual, the zero-order derivative of f is simply f, so that the constant term of the Taylor polynomial of f at $\mathbf{a}$ is $f(\mathbf{a})$.
(c) Exercise 29 shows that repetition occurs in the Taylor polynomial as written in part (b). For instance D_{113} occurs three times, as D_{113}, D_{131}, D_{311}. The sum of the corresponding three terms can be written in the form

$$3(D_1^2 D_3 f)(\mathbf{a}) x_1^2 x_3.$$

Prove (by calculating how often each derivative occurs) that the Taylor polynomial in (b) can be written in the form

$$\sum \frac{(D_1^{s_1} \cdots D_n^{s_n} f)(\mathbf{a})}{s_1! \cdots s_n!} x_1^{s_1} \cdots x_n^{s_n}.$$

Here the summation extends over all ordered n-tuples $(s_1, \ldots, s_n)$ such that each s_i is a nonnegative integer, and $s_1 + \cdots + s_n \le m - 1$.

Solution. (a) This formula is a simple application of the chain rule together with the fact that $D\mathbf{p}^{(i)}(t) = x_i$. The proof proceeds by induction on k.

(b) The formula is an immediate application of the fact that $\mathbf{p}(1) = \mathbf{a} + \mathbf{x}$, so that $h(1) = f(\mathbf{a} + \mathbf{x})$. The right-hand side is then an immediate application of the fact that $\mathbf{p}(0) = \mathbf{a}$. The only assertion that requires verification is that on the order of the remainder. The one-variable Taylor's theorem gives $r(\mathbf{x}) = \sum (D_{i_1,\ldots,i_m} f(\mathbf{p}(t)) x_{i_1} \cdots x_{i_m}$ for some $t \in (0,1)$, so that $|r(\mathbf{x})| \le K|\mathbf{x}|^m$ for some constant K. The assertion as to the order of r follows from this fact.

(c) If $s_1+\cdots+s_n \le m-1$, the number of terms having the derivative combination $D_1^{s_1} \cdots D_n^{s_n} f$ is $\binom{s_1 + \cdots + s_n}{s_1, \ldots, s_n} = \dfrac{(s_1 + \cdots + s_n)!}{s_1! \cdots s_n!}$. Thus the $k!$ that occurs in the one-variable Taylor's theorem is $(s_1 + \cdots + s_n)!$; and when the terms are consolidated, this factor cancels the numerator of the multinomial symbol, effectively being replaced by $s_1! \cdots s_n!$.

Exercise 9.31 Suppose $f \in C^{(3)}$ in some neighborhood of a point $\mathbf{a} \in R^2$, the gradient of f is $\mathbf{0}$ at $\mathbf{a}$, but not all second-order derivatives of f are 0 at $\mathbf{a}$. Show how one can then determine from the Taylor polynomial of f at $\mathbf{a}$ (of degree 2) whether f has a local maximum or a local minimum, or neither, at the point $\mathbf{a}$.

Extend this to R^n in place of R^2.

Solution. Let us simply do R^n in the first place and save the trouble of doing R^2. According to Taylor's theorem

$$f(\mathbf{a}+\mathbf{x}) - f(\mathbf{a}) = \frac{1}{2} \sum_{i_1,i_2} (D_{i_1 i_2} f)(\mathbf{a}) x_{i_1} x_{i_2} + r(\mathbf{x}),$$

where $|\mathbf{x}|^{-2} r(\mathbf{x}) \to 0$ as $\mathbf{x} \to \mathbf{0}$. Note that the Taylor polynomial can be concisely written as $\frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle$, where A is the $n \times n$ Hessian matrix whose i,j entry is $D_{ij} f(\mathbf{a})$ and the angle brackets denote the inner product. If A is positive-definite, i.e., if $\langle A\mathbf{x}, \mathbf{x}\rangle > 0$ when $\mathbf{x} \ne \mathbf{0}$, there is a positive constant c such that $\langle A\mathbf{x}, \mathbf{x}\rangle \ge c|\mathbf{x}|^2$. (The constant c is the minimum value of $\langle A\mathbf{x}, \mathbf{x}\rangle$ on the unit sphere $|\mathbf{x}| = 1$.) Hence if $\delta > 0$ is chosen so that $|r(\mathbf{x})| < c|\mathbf{x}|^2$ when $0 < |\mathbf{x}| < \delta$, we see that $f(\mathbf{a}+\mathbf{x}) - f(\mathbf{a}) > 0$ if $0 < |\mathbf{x}| < \delta$, i.e., $\mathbf{a}$ is a local minimum of f. Likewise if A is negative-definite, then $\mathbf{a}$ is a local maximum of f.

It is well-known from linear algebra that a necessary and sufficient condition for positive-definiteness of the matrix A is that the principal minors be positive, i.e., the $k \times k$-submatrix consisting of the elements in the first k rows and columns of A has a positive determinant. For negative-definiteness the corresponding criterion is that this minor have the same sign as $(-1)^k$.

There are no other resonably regular cases that guarantee a maximum or minimum. A nonnegative-definite or nonpositive-definite matrix may well fail to guarantee a maximum or minimum, even in R^1. If the quadratic form $\langle A\mathbf{x}, \mathbf{x}\rangle$ assumes both signs, then the point $\mathbf{a}$ is definitely not either a maximum or a minimum. (If $\langle A\mathbf{x}, \mathbf{x}\rangle > 0$, then $f(\mathbf{a}+t\mathbf{x}) - f(\mathbf{a}) > 0$ for small values of t, while if $\langle A\mathbf{x}, \mathbf{x}\rangle < 0$, then $f(\mathbf{a}+t\mathbf{x}) - f(\mathbf{a}) < 0$ for small values of t.)