Exercise 5.1. For each of the following equations determine at which points one cannot find a function $y = f(x)$ which describes the graph in this neighbourhood. Sketch the graphs.

(a)

$$\frac{1}{3}y^3 - 2y + x = 1$$

(b)

$$x^2\left(\frac{\cos^2\phi}{a^2} + \frac{\sin^2\phi}{b^2}\right) - xy\left(\frac{1}{a^2} - \frac{1}{b^2}\right)\sin(2\phi) + y^2\left(\frac{\sin^2\phi}{a^2} + \frac{\cos^2\phi}{b^2}\right) = 1,$$

where $a > 0$, $b > 0$, $0 \leq \phi \leq \pi/2$ are fixed parameters. Note the cases $a = b$, $\phi = 0$, $\phi = \pi/2$.

Solution: (a) Let

$$F(x, y) = \frac{1}{3}y^3 - 2y + x - 1.$$

The solutions of the equation satisfy $F(x, y) = 0$. To employ the Implicit Function Theorem, we need to identify the solutions (x, y) of $F(x, y) = 0$ such that $\frac{\partial}{\partial y}F(x, y) \neq 0$. Solving the equation $\frac{\partial}{\partial y}F(x, y) = 0$ gives $y = \pm\sqrt{2}$. Substituting $y = +\sqrt{2}$ in $F(x, y) = 0$ we get $x = 1 - \frac{4}{3}\sqrt{2}$, and substituting $y = -\sqrt{2}$ in $F(x, y) = 0$ we get $x = 1 + \frac{4}{3}\sqrt{2}$. Thus, the theorem does not apply at the points

$$(1 - \frac{4}{3}\sqrt{2}, \sqrt{2}), \qquad (1 + \frac{4}{3}\sqrt{2}, -\sqrt{2}).$$

Now by the Implicit Function Theorem, for every (x, y) in $\mathbb{R}^2$, except the above two points, the solution of the equation $F(x, y) = 0$ near (x, y) is the graph of a function. That is, given (x, y), there are open sets A containing x and B containing y, and a function $g : A \to B$ such that $(x', y') \in A \times B$ is a solution of the equation $F(x', y') = 0$ if and only if $y' = g(x')$.

To see what is happening at the two exceptional points, we may rewriting the equation in the form

$$x = -\frac{1}{3}y^3 + 2y + 1.$$

We note that the first derivative $\frac{d}{dy}x = 0$ and the second derivative $\frac{d^2}{dy^2}x \neq 0$ at any of the two exceptional points. Thus, those points are either a maximum or minimum for the graph of the function which gives the solution in terms of y. Thus, $y = g(x)$ does not exist in any neighbourhood.

(b) As in the previous part, we may write the equation in the form $F(x, y) = 0$, for a suitable function F. The candidate points where the Implicit Function Theorem cannot be applied are the solutions of the equation $\frac{\partial}{\partial y}F(x, y) = 0$. That gives us

$$y = x\frac{b^2 - a^2}{b^2\sin^2\phi + a^2\cos^2\phi}\frac{\sin(2\phi)}{2}.$$

Please send any corrections to d.cheraghi@imperial.ac.uk
Questions marked with $*$ are optional

If we substitute the above relation in the equation $F(x,y)=0$, we obtain 2 points on the graph (one for the pluses signs and one for the minuses signs):

$$x = \pm\sqrt{b^2\sin^2\phi + a^2\cos^2\phi}, \qquad y = \pm\frac{b^2-a^2}{\sqrt{b^2\sin^2\phi + a^2\cos^2\phi}}\frac{\sin(2\phi)}{2}.$$

Note that the solution of the equation $F(x,y)=0$ is in fact the ellipse

$$\frac{x'^2}{a^2} + \frac{y'^2}{b^2} = 1$$

rotated by the angle ϕ, using the transformation

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Thus, at the two points we have identified, the solution cannot be written as the graph of a function. Indeed, for $a=b$, this problem reduces to the one we considered in the lectures.

Exercise 5.2. Consider the equation

$$2x^2 + 4xy + y^2 = 3x + 4y$$

a) Show that this system of equations (implicitly) defines a function $y=f(x)$ with $f(1)=1$.

Solution: We consider the function $F:\mathbb{R}^2\to\mathbb{R}$ defined as

$$F(x,y) = (2x^2+4xy+y^2) - (3x+4y).$$

We note that $F(1,1)=0$, that is, $(x_0,y_0)=(1,1)$ is a solution of the equation $F(x,y)=0$. We aim to employ the Implicit Function Theorem.

We have

$$D_2F(x,y) = 4x+2y-4,$$

which shows that D_2F is a continuous function. Moreover, $D_2F(1,1)=2\neq 0$.

By the (simple version of the) Implicit Function Theorem, there exists a neighbourhood $U\subset\mathbb{R}$ of $x_0=1$ and a continuously differentiable function $f:U\to\mathbb{R}$ satisfying $f(1)=f(x_0)=y_0=1$ such that

$$F(x,f(x)) = 0 \text{ for all } x\in U.$$

b) Compute $f'(1)$ without knowing f explicitly.

Solution: Let us consider the map $g(x)=F(x,f(x))$, for $x\in U$. We may write this map as the composition of the maps $h(x)=(x,f(x))$ followed by the map $F(x,y)$. That is, $g(x)=F\circ h(x)$. By the chain rule, we have

$$\begin{aligned} Dg(x) = DF(h(x))\circ Dh(x) &= \begin{pmatrix} D_1F(x,f(x)) & D_2F(x,f(x)) \end{pmatrix}\begin{pmatrix} 1 \\ f'(x) \end{pmatrix} \\ &= D_1F(x,f(x)) + D_2F(x,f(x))f'(x). \end{aligned}$$

From the definition of the function F, we have

$$D_1F(x,y) = 4x + 4y - 3,$$

and hence $D_1F(1,1) = 5$. On the other hand, since $g \equiv 0$ on U, we have $g'(1) = 0$. Therefore, the above equation at $x = 1$ gives us

$$0 = 5 + 2f'(1),$$

which implies $f'(1) = -5/2$.

c) Find an explicit formula for f and check your result from b).

Solution: To identify f explicitly, we must solve the equation $F(x,y) = 0$ for y, which is possible here since F is a quadratic equation. That gives us

$$y = 2 - 2x \pm \sqrt{2x^2 - 5x + 4}.$$

Since $f(1) = 1 > 0$ we must choose the positive sign in the above equation, which becomes

$$f(x) = 2 - 2x + \sqrt{2x^2 - 5x + 4}.$$

It follows that

$$f'(x) = -2 + \frac{4x - 5}{2\sqrt{2x^2 - 5x + 4}},$$

and hence $f'(1) = -2 - 1/2 = -5/2$.

Exercise 5.3. Let $X = \mathbb{R}^n$ and define the function $\mathrm{d}_\infty : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ as

$$\mathrm{d}_\infty(x,y) = \max\{|x^1 - y^1|, \ldots, |x^n - y^n|\}.$$

Show that d_∞ is a metric on $\mathbb{R}^n$.

Solution: We must verify the three properties M1-M3.

M1: By the properties of the modulus function, for all $x \in \mathbb{R}$, $|x| \geq 0$. This implies that $\mathrm{d}_\infty(x,y) \geq 0$. Moreover, for every $x \in \mathbb{R}$, $|x| = 0$ iff $x = 0$. Therefore, $\mathrm{d}_\infty(x,y) = 0$ iff $x^i = y^i$ for all $i = 1, 2, \ldots, n$ iff $x = y$.

M2: Since $|-x| = |x|$, we have

$$\mathrm{d}_\infty(x,y) = \max\{|x^1 - y^1|, \ldots, |x^n - y^n|\} = \max\{|y^1 - x^1|, \ldots, |y^n - x^n|\} = \mathrm{d}_\infty(y,x).$$

M3: Let

$$x = (x^1, x^2, \ldots, x^n), \quad y = (y^1, y^2, \ldots, y^n), \quad z = (z^1, z^2, \ldots, z^n)$$

be arbitrary elements in $\mathbb{R}^n$. By the triangle inequality for the modulus, for every $i = 1, 2, \ldots, n$, we have

$$|x^i - z^i| \leq |x^i - y^i| + |y^i - z^i|.$$

For every $k \in \{1, 2, \ldots n\}$ we have

$$\begin{aligned} |x^k - z^k| &\leq |x^k - y^k| + |y^k - z^k| \\ &\leq \max\{|x^1 - y^1|, \ldots, |x^n - y^n|\} + \max\{|y^1 - z^1|, \ldots, |y^n - z^n|\} \\ &= \mathrm{d}_\infty(x,y) + \mathrm{d}_\infty(y,z). \end{aligned}$$

This implies that

$$d_\infty(x, z) = \max\{|x^1 - z^1|, \dots, |x^n - z^n|\} \leq d_\infty(x, y) + d_\infty(y, z).$$

Alternatively, the last step for the proof of property M3, can be given as follows. First note that if A and B are finite sets of real numbers, we have

$$\max(A + B) \leq \max A + \max B.$$

Therefore,

$$\begin{aligned}
d_\infty(x, z) &= \max\{|x^1 - z^1|, \dots, |x^n - z^n|\} \\
&= \max\{|x^1 - y^1 + y^1 - z^1|, |x^2 - y^2 + y^2 - z^2|, \dots, |x^n - y^n + y^n - z^n| \\
&\leq \max\{|x^1 - y^1| + |y^1 - z^1|, |x^2 - y^2| + |y^2 - z^2|, \dots, |x^n - y^n| + |y^n - z^n|\} \\
&\leq \max\{|x^1 - y^1|, \dots, |x^n - y^n|\} + \max\{|y^1 - z^1|, \dots, |y^n - z^n|\} \\
&= d_\infty(x, y) + d_\infty(y, z).
\end{aligned}$$

Exercise 5.4. Show that each of the following functions is a metric on $\mathbb{R}$:

(i) $d(x, y) = |x^3 - y^3|$, (here x^3 means x raised to power 3)

(ii) $d(x, y) = |e^x - e^y|$,

(iii) $d(x, y) = |\tan^{-1}(x) - \tan^{-1}(y)|$.

Which property of the maps $x \mapsto x^3$, $x \mapsto e^x$, and $x \mapsto \tan^{-1}(x)$ makes these functions a metric.

Solution: Let $f(x)$ stand for any of the functions $x \mapsto x^3$, $x \mapsto e^x$, and $x \mapsto \tan^{-1}(x)$. By the properties of the modulus function, we immediately obtain $d(x, y) \geq 0$, and $d(x, y) = d(y, x)$. Also, by the inequalities

$$d(x, y) = |f(x) - f(y)| \leq |f(x) - f(z)| + |f(z) - f(y)| = d(x, z) + d(z, y),$$

we obtain the triangle inequality for the functions d in each case.

There remains to see that $d(x, y) = 0$ iff $x = y$. Clearly, if $x = y$, $d(x, x) = 0$. The opposite implication follows from the fact that f is injective in all the three cases. That is, if $f(x) = f(y)$ for some x and y in $\mathbb{R}$, we must have $x = y$.

This exercise shows that there are many metrics on $\mathbb{R}$, as there are many injective maps from $\mathbb{R}$ to $\mathbb{R}$. Note that the continuity of f is not required here.

Exercise 5.5. Assume that $a < b$ are real numbers, and $h : (a, b) \to (0, \infty)$ is a continuous function. For x and y in (a, b), we define

$$d_h(x, y) = \int_{\min\{x,y\}}^{\max\{x,y\}} h(t)\, dt.$$

Show that d_h is a metric on (a, b).

Solution: M2: Since $\{x, y\} = \{y, x\}$ as sets, by the definition of d_h, we immediately see that $\mathrm{d}_h(x, y) = \mathrm{d}_h(y, x)$. Therefore, without loss of generality, below we assume that $x \leq y$.

M1: For real numbers $x \leq y$ and a function $h \geq 0$, the Riemann integral satisfies $\int_x^y h(t)dt \geq 0$. Moreover, by the definition of integral, if $x = y$ we have $\int_x^y h(t)dt = 0$. On the other hand, by a lemma proved in the typed lectures, if $h > 0$ and $x < y$, we must have $\int_x^y h(t)dt > 0$. Since $h > 0$, this implies that if $\int_x^y h(t)dt = 0$ we must have $x = y$. Therefore, $\mathrm{d}_h(x, y) = 0$ iff $x = y$.

M3: Let x, y and z be arbitrary real numbers. Without loss of generality, assume that $x \leq y$. Recall from the properties of the Riemann integral that if $x \leq z \leq y$, we have

$$\int_x^y h(t)dt = \int_x^z h(t)dt + \int_z^y h(t)dt.$$

This implies that for all real numbers $x \leq z \leq y$, we have

$$\mathrm{d}_h(x, y) \leq \mathrm{d}_h(x, z) + \mathrm{d}_h(z, y).$$

If $z \notin [x, y]$, we must either have $z \leq x$ or $z \geq y$. In the first case, we have $[x, y] \subset [z, y]$, and hence

$$\int_x^y h(t)dt \leq \int_z^y h(t)dt = \int_z^x h(t)dt + \int_x^y h(t)dt$$

and in the second case we have $[x, y] \subseteq [x, z]$, and hence

$$\int_x^y h(t)dt \leq \int_x^z h(t)dt = \int_x^y h(t)dt + \int_y^z h(t)dt.$$

Each of these inequalities imply that

$$\mathrm{d}_h(x, y) \leq \mathrm{d}_h(x, z) + \mathrm{d}_h(z, y).$$

Exercise 5.6. Consider the function $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined as

$$g(x, y) = |x - y|^2.$$

Show that g is not a metric on $\mathbb{R}$.

Solution: It is sufficient to show that one of the properties of the metric does not hold. Consider the three points $2, 3, 4$. Then,

$$g(2, 4) = 4 \nleq 1 + 1 = g(2, 3) + g(3, 4).$$

This shows the triangle inequality does not hold for the three points $2, 3, 4$.

Another counter example is given by the three points $0, 10, 20$ in $\mathbb{R}$, as

$$g(0, 20) = 400 \nleq 200 = 100 + 100 = g(0, 10) + g(10, 20).$$

Exercise 5.7. Let $X = \mathbb{R}^2$, and define $\mathrm{d}_{\text{rail}} : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ as

$$\mathrm{d}_{\text{rail}}(x, y) = \begin{cases} \|x - y\| & \text{if } x = ky \text{ for some } k \in \mathbb{R} \\ \|x\| + \|y\| & \text{otherwise} \end{cases}$$

Show that d_{rail} is a metric on $\mathbb{R}^2$.

This is called the British rail metric. The intuition behind this metric is that if two towns are on the same rail line, then we travel between them, but if the towns are on distinct lines, we travel via London (represented as the origin in $\mathbb{R}^2$).

Solution: The properties $\mathrm{d_{rail}}(x,y) \geq 0$, $\mathrm{d_{rail}}(x,y) = 0$ iff $x = y$, and $\mathrm{d_{rail}}(x,y) = \mathrm{d_{rail}}(y,x)$ easily follow from the properties of the norm. We need to show the triangle inequality for $\mathrm{d_{rail}}$. Let x, y, and z be arbitrary points in $\mathbb{R}^2$. We consider few cases below.

1) Assume that there is $k \in \mathbb{R}$ such that $x = ky$.

$$\mathrm{d_{rail}}(x,y) = \|x-y\| \leq \|x-z\| + \|z-y\| \leq \mathrm{d_{rail}}(x,z) + \mathrm{d_{rail}}(z,y),$$

since we always have $\|a-b\| \leq \|a\| + \|b\|$.

2) Assume that for all $k \in \mathbb{R}$ we have $x \neq ky$. In particular, $x \neq 0$. There are several cases to look at in this case.

(i) $z = 0$. We have

$$\mathrm{d_{rail}}(x,y) = \|x\| + \|y\| = \mathrm{d_{rail}}(x,0) + \mathrm{d_{rail}}(0,y)$$

(ii) There is $m \in \mathbb{R} \setminus \{0\}$ such that $z = my$. Then, for all $s \in \mathbb{R}$, we have $z \neq sx$, otherwise $my = sx$, and (as $s \neq 0$) $x = (m/s)y$, which is a contradiction. In particular,

$$\mathrm{d_{rail}}(x,y) = \|x\| + \|y\| \leq \|x\| + \|z-y\| + \|z\| = \mathrm{d_{rail}}(x,z) + \mathrm{d_{rail}}(z,y).$$

(iii) There is $l \in \mathbb{R} \setminus \{0\}$ such that $z = lx$. This is similar to case (ii), as one may switch x and y in that proof.

(iv) For all $m \in \mathbb{R}$, we have $z \neq my$ and $z \neq mx$. Then,

$$\mathrm{d_{rail}}(x,y) = \|x\| + \|y\| \leq \|x\| + \|z\| + \|z\| + \|y\| = \mathrm{d_{rail}}(x,z) + \mathrm{d_{rail}}(z,y).$$