M2AA1 Differential Equations
Exercise sheet 3 answers

1. We have, $d(\mathbf{x},\mathbf{y}) = d(\mathbf{y},\mathbf{x})$, $d(\mathbf{x},\mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$ and the triangle inequality $d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z}) \geq d(\mathbf{x},\mathbf{z})$. From the latter it follows that $d(\mathbf{y},\mathbf{z}) = \frac{1}{2}(d(\mathbf{z},\mathbf{y}) + d(\mathbf{y},\mathbf{z})) \geq \frac{1}{2}d(\mathbf{z},\mathbf{z}) = 0$ for all $\mathbf{y},\mathbf{z}$.

2. A metric space is complete if every Cauchy sequence in it converges. For the mentioned examples, we use the (natural) distance function $d(x,y) = |x-y|$.

   - $\mathbb{Q}$: is not a complete metric space, since Cauchy sequences in $\mathbb{Q}$ need not converge to have a limit in $\mathbb{Q}$. For example the Fibonacci sequence $\{a_n\}_{n\in\mathbb{N}}$ with $a_0 = 1$ and $a_n = 1/(1+a_{n-1})$ converges to $(\sqrt{5}-1)/2 \notin \mathbb{Q}$.
   - $\mathbb{R}$: the real line is in fact defined as the *completion* of $\mathbb{Q}$, i.e. $\mathbb{R}$ is the smallest complete metric space containing $\mathbb{Q}$.
   - $\mathbb{Z}$: the tail of every Cauchy sequence in $\mathbb{Z}$ must be constant, with integer value. Hence, the limit of any such sequence is in $\mathbb{Z}$ so $(\mathbb{Z}, d)$ is a complete metric space.
   - $[0,1]$: If for a Cauchy sequence in $\mathbb{R}$ every element lies inside $[0,1]$ then so lies the limit point, so $[0,1]$ is a complete metric space.
   - $[0,1)$: If the interval is half open, we can have a Cauchy sequence in $[0,1)$ that converges to 1. For instance, the sequence $\{a_n\}_{n\in\mathbb{N}}$ with $a_n = 1-1/n$ is a Cauchy sequence in $[0,1)$ but it does converge to 1 which lies outside $[0,1)$. Hence $[0,1)$ is not a complete metric space.

3. First we note that from the condition it immediately follows that if there is a fixed point of F then it must be unique. Namely, if $x \neq y$, then if x and y would be fixed points of F it follows that $d(x,y) < d(x,y)$ which is a contradiction.

   Define $x_n = F^n(x)$ and let $a_n = d(x_{n+1}, x_n)$. Because F is shrinking, a_n is decreasing. Since a_n is bounded from below (by 0), a_n converges to some $c \geq 0$ as $n \to \infty$: $\lim_{n\to\infty} d(x_{n+1}, x_n) = c \geq 0$. By compactness of X, the sequence $\{x_n\}_{n\in\mathbb{N}}$ has a converging subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ (with n_k a strictly increasing function of k). We now consider the sequences $\{y_k\}_{k\in\mathbb{N}}$ and $\{z_k\}_{k\in\mathbb{N}}$ with $y_k := x_{n_k}$ and $z_k := x_{n_k+1}$. As $\{x_{n_k}\}_{k\in\mathbb{N}}$ is by construction converging to some point $x \in X$, we have $\lim_{k\to\infty} y_k = x$ and $\lim_{k\to\infty} z_k = F(x)$ with $d(x, F(x)) = c$. Then $d(F(x), F^2(x)) = c$, as if $d(F(x), F^2(x)) < c$ we would have $\lim_{n\to\infty} a_n < c$ as well, by continuity of F and d. But this implies that in case $x \neq F(x)$ we obtain a contradiction: $c = d(F(x), F^2(x)) < d(x, F(x)) = c$. Hence $\lim_{n\to\infty} a_n = d(x, F(x)) = c = 0$ so that $F(x) = x$ and the point of convergence x is a fixed point (and the unique fixed point) of F. Hence $\lim_{n\to\infty} F^n(z) = x$ for all $z \in X$ where x is the unique fixed point of F in X.

   An example that satisfies the theorem, but where convergence is not exponential is $F : [0,1] \to [0,1]$ with $F(x) = x(1-x/2)$, so that if $(x,y) \neq (0,0)$ $d(F(x),F(y)) = |x - \frac{1}{2}x^2 - y + \frac{1}{2}y^2| = |(x-y)(1-(x+y)/2)| < |x-y|$ as $1-(x+y)/2 < 1$. Assume that K were a contraction constant for F. But taking $x = 0$, we would have for all $1 \geq y > 0$: $y(1-y/2) = |y - (y^2)/2| \leq Ky$. But for y sufficiently close to $x = 0$ we have $1 - y/2 > K$, contradicting the fact that K is a contraction constant.

4. As in course notes; but also $F \circ F^{-1}(y) = y$ implies $\frac{d}{dy}(F \circ F^{-1})(y) = 1$ and by the chain rule $F'(F^{-1}(y)) \cdot (F^{-1})'(y) = 1$ implying $(F^{-1})'(y) = 1/F'(x)$ since $F(x) = y$.

5. We first establish the remark in the hint. By definition, since $\mathbf{x}(t)$ with initial condition $\mathbf{x}(0) = \mathbf{y}$ satisfies $\mathbf{x}(t) = \Phi^t(\mathbf{y})$, we have
$$\frac{d}{dt}\Phi^t(\mathbf{y}) = f(\Phi^t(\mathbf{y})).$$
   We note in this expression that $\mathbf{y}$ is a constant initial condition (independent of t). Still, of course, the expression holds for all possible initial conditions $\mathbf{y} \in \mathbb{R}^m$. Thus we also have
$$\frac{d}{dt}\Phi^t = f \circ \Phi^t,$$

where $\circ$ denotes composition.

We differentiate both sides with respect to $\mathbf{y}$ in the equilibrium point $\mathbf{x}_0$:

$$\frac{d}{dt}D\Phi^t(\mathbf{x}_0) = Df(\Phi^t(x_0))D\Phi^t(\mathbf{x}_0) = Df(x_0)D\Phi^t(\mathbf{x}_0).$$

From this linear (matrix valued) ODE it follows that

$$D\Phi^t(x_0) = \exp(Df(\mathbf{x}_0)t),$$

where we used the fact that $\Phi^0(\mathbf{y}) = D\Phi^0(\mathbf{y}) = \mathbf{y}$ for all $\mathbf{y}$.

If all eigenvalues λ_i of $Df(\mathbf{x}_0)$ have negative real part then all eigenvalues of $D\Phi^t(\mathbf{x}_0)$ (with $t > 0$) have eigenvalues $\exp(\lambda_i)$, the modulus of which consquently is bounded from above by a number $K < 1$. We can apply the derivative test with C being equal to a sphere with radius ρ around $\mathbf{x}_0$, with ρ suffuciently close to zero.

6. See attachment.

7. See atachments.

8. (a) If F and G are linear maps, then the surfaces are linear subspace of $\mathbb{R}^m$. Let us call the intersection point of these surfaces (without loss of generality) 0 so that $F(0) = G(0) = 0$. Let $H(x, y) = F(x) - G(y)$. It is useful to note that if H is linear we can represent it in matrix form as

$$H = (F, -G) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k, \quad \text{so that} \quad H\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = (F, -G)\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = F\mathbf{x} - G\mathbf{y},$$

where F is a $k \times n$ matrix and G a $k \times m$ matrix. Of course, we have $H(0) = 0$. The matrix representation of H (that is equal to DH since H is linear) consists of column vectors $\frac{\partial F}{\partial s_i}(0)$ (with $i = 1, \ldots n$) and $\frac{\partial G}{\partial s_i}(0)$ (with $i = n + 1, \ldots n + m$). Let us assume that these column vectors span $\mathbb{R}^k$, so that there are at least k linearly independent column vectors. Let A be the matrix obtained from DH by taking k column such linearly independent column vectors and write $DH = A(\lambda)$ where $\lambda = \mathbb{R}^{m+n-k}$ are parameters representing the remaining column vectors. By construction A is invertible so that, by the Implicit Function Theorem, the zeros of H can be continued uniquely as function of $\lambda \in \mathbb{R}^{n+m-k}$. We thus have $\mathbf{x}(\lambda)$ and we can see this solution set as a graph over λ so that locally the dimension of this solution set is equal to the dimension of λ, i.e. $n + m - k$.

   (b) If the functions F and G are nonlinear, the same argument can be carried through where we observe that the columns of DH are given by the tangent vectors $\frac{\partial F}{\partial s_i}(0)$ (with $i = 1, \ldots n$) to the surface given by F, and tangent vectors $\frac{\partial G}{\partial s_i}(0)$ (with $i = n + 1, \ldots n + m$) to the surface given by G. One often refers to the span of these respective sets of tangent vectors as the *tangent space* at 0. If the *sum* of the tangent spaces of the two surfaces in the intersection point is equal to the ambient space $\mathbb{R}^m$ we can carry through the argument used above to obtain the solution set $\mathbf{x}(\lambda)$ as a graph over $\mathbb{R}^{m+n-k}$ so that the dimension of the intersection set is indeed $m + n - k$.

   (c) If we add additional p parameters, we can include them in λ so that $\lambda \in \mathbb{R}^{m+n-k+p}$. By application of the same argument (noting that the transversality is a property that persists under small perturbations (in the C^1 sense, i.e. small deviations in the value of map <u>and</u> it's derivative)) we obtain a $(m+n-k+p)$-dimensional solution set near 0 which we can also view as a continuation in p directions of the $(m+n-k)$-dimensional solution set obtained in (b) above.

In the particular case of the intersection of two two-dimensional surfaces in $\mathbb{R}^3$ we thus find that transversality implies that (locally) the intersection consists of a curve. Transversality implies for the intersection of a curve and a two-dimensional surface in $\mathbb{R}^3$ that the intersection is isolated (i.e. there is a small ball around the intersection point in which no other intersection point lies).

9. See atachments.

Consider the polar coordinate transformation

$$\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \quad \text{where} \quad \varphi(r,\theta) \ = \ (r\,\cos\theta, r\,\sin\theta)\,.$$

By the inverse function theorem, the transformation φ is invertible *near a point* $\alpha = (r,\theta)$ if its derivative $\varphi'(\alpha)$ is also invertible.

Now

$$\varphi'(r,\theta) \, : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

is defined by

$$\begin{bmatrix} r \\ \theta \end{bmatrix} \ \mapsto \ \begin{bmatrix} \frac{\partial\varphi_1}{\partial r}(r,\theta) & \frac{\partial\varphi_1}{\partial\theta}(r,\theta) \\ \frac{\partial\varphi_2}{\partial r}(r,\theta) & \frac{\partial\varphi_2}{\partial\theta}(r,\theta) \end{bmatrix} \cdot \begin{bmatrix} r \\ \theta \end{bmatrix}$$

where

$$\begin{aligned} \varphi_1(r,\theta) \ &= \ r\,\cos\theta \\ \varphi_2(r,\theta) \ &= \ r\,\sin\theta. \end{aligned}$$

Therefore, since

$$\begin{aligned} \frac{\partial\varphi_1}{\partial r} \ &= \ \cos\theta & \qquad \frac{\partial\varphi_1}{\partial\theta} \ &= \ -r\,\sin\theta \\ \frac{\partial\varphi_2}{\partial r} \ &= \ \sin\theta & \qquad \frac{\partial\varphi_2}{\partial\theta} \ &= \ r\,\cos\theta\,, \end{aligned}$$

we conclude that

$$\varphi'(r,\theta)\begin{bmatrix} r \\ \theta \end{bmatrix} \ = \ \begin{bmatrix} \cos\theta & -r\,\sin\theta \\ \sin\theta & r\,\cos\theta \end{bmatrix} \cdot \begin{bmatrix} r \\ \theta \end{bmatrix}.$$

Because

$$\det\begin{bmatrix} \cos\theta & -r\,\sin\theta \\ \sin\theta & r\,\cos\theta \end{bmatrix} \ = \ r\,(\cos^2\theta + \sin^2\theta) \ = \ r\,,$$

the matrix is invertible for all points $\alpha = (r,\theta)$ provided $r \neq 0$.

By the inverse function theorem, therefore, the polar coordinate transformation φ is invertible in a neighbourhood of each of those points $(r,\theta) \in \mathbb{R}^2$ for which $r \neq 0$.

Consider the spherical coordinate transformation

$$\Sigma : \mathbb{R}^3 \longrightarrow \mathbb{R}^3, \qquad \Sigma(\rho,\theta,\phi) = \rho(\sin\phi\ \cos\theta, \sin\phi\ \sin\theta, \cos\phi).$$

By the inverse function theorem, the transformation Σ is invertible *near a point* $\alpha = (\rho,\theta,\phi)$ if its derivative $\Sigma'(\alpha)$ is also invertible.

Now

$$\Sigma'(\rho,\theta,\phi) : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

is represented by the matrix

$$\begin{bmatrix} \frac{\partial \Sigma_1}{\partial \rho}(\rho,\theta,\phi) & \frac{\partial \Sigma_1}{\partial \theta}(\rho,\theta,\phi) & \frac{\partial \Sigma_1}{\partial \phi}(\rho,\theta,\phi) \\ \frac{\partial \Sigma_2}{\partial \rho}(\rho,\theta,\phi) & \frac{\partial \Sigma_2}{\partial \theta}(\rho,\theta,\phi) & \frac{\partial \Sigma_2}{\partial \phi}(\rho,\theta,\phi) \\ \frac{\partial \Sigma_3}{\partial \rho}(\rho,\theta,\phi) & \frac{\partial \Sigma_3}{\partial \theta}(\rho,\theta,\phi) & \frac{\partial \Sigma_3}{\partial \phi}(\rho,\theta,\phi) \end{bmatrix}$$

where

$$\begin{aligned} \Sigma_1(\rho,\theta,\phi) &= \rho\sin\phi\ \cos\theta \\ \Sigma_2(\rho,\theta,\phi) &= \rho\sin\phi\ \sin\theta \\ \Sigma_3(\rho,\theta,\phi) &= \rho\cos\phi. \end{aligned}$$

We conclude that

$$\Sigma'(\rho,\theta,\phi) = \begin{bmatrix} \sin\phi\ \cos\theta & -\rho\sin\phi\ \sin\theta & \rho\cos\phi\ \cos\theta \\ \sin\phi\ \sin\theta & \rho\sin\phi\ \cos\theta & \rho\cos\phi\ \sin\theta \\ \cos\phi & 0 & -\rho\sin\phi \end{bmatrix}.$$

Since

$$\det \begin{bmatrix} \sin\phi\ \cos\theta & -\rho\sin\phi\ \sin\theta & \rho\cos\phi\ \cos\theta \\ \sin\phi\ \sin\theta & \rho\sin\phi\ \cos\theta & \rho\cos\phi\ \sin\theta \\ \cos\phi & 0 & -\rho\sin\phi \end{bmatrix} = -\rho^2\sin\phi\,,$$

$\Sigma'(\rho,\theta,\phi)$ is a linear isomorphism for all points $\alpha = (\rho,\theta,\phi)$ provided $\rho \neq 0$ and $\phi \neq \pm k\pi$ for all $k \in \mathbb{N}$.

By the inverse function theorem, therefore, the spherical coordinate transformation Σ is invertible in a neighbourhood of each of those points $(\rho,\theta,\phi) \in \mathbb{R}^3$ for which $\rho \neq 0$ and $\phi \neq \pm k\pi$ for all $k \in \mathbb{N}$.

Exercise 2.

Determine whether the function

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 , \qquad f(x, y) = (\cos x, xy)$$

has a differentiable inverse in a neighbourhood of the point $\alpha = (\pi, -1)$.

If so, find the derivative of its inverse function at $f(\pi, -1)$.

Solution We first let

$$\begin{aligned} \eta := f(\pi, -1) &= (\cos \pi, \pi(-1)) = (-1, -\pi) \\ f_1(x, y) &:= \cos x \\ f_2(x, y) &:= xy . \end{aligned}$$

By the inverse function theorem, f is invertible in a neighbourhood of $\alpha = (\pi, -1)$ if its derivative function $f'(\pi, -1)$ is onto. Now the linear map $f'(x, y)$ is represented by the matrix

$$f'(x, y) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{bmatrix} = \begin{bmatrix} -\sin x & 0 \\ y & x \end{bmatrix} .$$

Therefore, $f'(\pi, -1)$ is represented by the matrix

$$f'(\pi, -1) = \begin{bmatrix} 0 & 0 \\ -1 & \pi \end{bmatrix} ,$$

and is defined as the map

$$f'(\pi, -1) : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 , \quad (x, y) \mapsto \begin{bmatrix} 0 & 0 \\ -1 & \pi \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} .$$

Now $f'(\pi, -1)$ is *not* onto as the rows of the matrix $f'(\pi, -1)$ are not linearly independent (row rank $= 1 \neq 2 = \dim \mathbb{R}^2$). Therefore, f does *not* have a differentiable inverse in a neighbourhood of $\alpha = (\pi, -1)$.

Answer No, the function

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 , \qquad f(x, y) = (\cos y, xy) ,$$

does not have a differentiable inverse in a neighbourhood of the point $\alpha = (\pi, -1)$.

Exercise 3.

Determine whether the function

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3, \qquad f(x, y, z) = (3x + y, z - 3y, x + z),$$

has a differentiable inverse in a neighbourhood of the point $\alpha = (2, -3, 5)$.

If so, find the derivative of its inverse function at $f(2, -3, 5)$.

Solution Guided by the inverse function theorem, we compute

$$f'(x, y, z) = \begin{bmatrix} 3 & 1 & 0 \\ 0 & -3 & 1 \\ 1 & 0 & 1 \end{bmatrix} = f'(2, -3, 5).$$

Thus f has a differentiable inverse near $(2, -3, 5)$ if and only if $\det(f'(2, -3, 5)) \neq 0$. We find

$$\det A = \det \begin{bmatrix} 3 & 1 & 0 \\ 0 & -3 & 1 \\ 1 & 0 & 1 \end{bmatrix} = -8 \neq 0$$

and conclude that f has a differentiable inverse near the point $(2, -3, 5)$. Its derivative at

$$\eta := f(2, -3, 5) = (-1, 14, 7)$$

is

$$(f^{-1})'(\eta) = \begin{bmatrix} 3 & 1 & 0 \\ 0 & -3 & 1 \\ 1 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{3}{8} & \frac{1}{8} & \frac{-1}{8} \\ \frac{-1}{8} & \frac{-3}{8} & \frac{3}{8} \\ \frac{-3}{8} & \frac{-1}{8} & \frac{9}{8} \end{bmatrix}.$$

Thus far the conclusions we can draw by depending solely upon the inverse function theorem. In this particular case, we can obtain additional insight by observing that f itself is a linear map whose matrix is

$$A := \begin{bmatrix} 3 & 1 & 0 \\ 0 & -3 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

is also its derivative at each point of $\mathbb{R}^3$: $f'(x, y, z) = A$. Now the linear function f is invertible exactly when $\det(A) \neq 0$. We saw above that $\det(A) = -8$. Therefore, we conclude that f is an isomorphism of vector spaces whose inverse is given by the matrix

$$A^{-1} = \begin{bmatrix} \frac{3}{8} & \frac{1}{8} & \frac{-1}{8} \\ \frac{-1}{8} & \frac{-3}{8} & \frac{3}{8} \\ \frac{-3}{8} & \frac{-1}{8} & \frac{9}{8} \end{bmatrix}.$$

Answer Yes, the function

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3, \qquad f(x, y, z) = (3x + y, z - 3y, x + z),$$

is linear and is invertible as a whole. Its inverse

$$f^{-1} = \begin{bmatrix} \frac{3}{8} & \frac{1}{8} & \frac{-1}{8} \\ \frac{-1}{8} & \frac{-3}{8} & \frac{3}{8} \\ \frac{-3}{8} & \frac{-1}{8} & \frac{9}{8} \end{bmatrix}$$

again linear, hence satisfies $(f^{-1})'(\eta) = f^{-1}$ for each $\eta \in \mathbb{R}^3$.

Exercise 4.

Determine whether the function

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 , \qquad f(x,y,z) \;=\; (xy, yz, xz) ,$$

has a differntiable inverse in a neighbourhood of the point $\alpha = (1, 0, -1)$.

If so, find the derivative of its inverse function at $f(1, 0, -1)$.

Solution We first let

$$\begin{aligned} \eta := \; f(1,0,-1) \; &= \; (0,0,-1) \\ f_1(x,y,z) :=& \; xy \\ f_2(x,y,z) :=& \; yz \\ f_3(x,y,z) :=& \; xz \, . \end{aligned}$$

By the inverse function theorem, f is invertible in a neighbourhood of $\alpha = (1, 0, -1)$ if its derivative function $f'(1, 0, -1)$ is onto. Now the linear map $f'(x, y, z)$ is represented by the matrix

$$f'(x,y,z) \;=\; \begin{bmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} & \dfrac{\partial f_1}{\partial z} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} & \dfrac{\partial f_2}{\partial z} \\ \dfrac{\partial f_3}{\partial x} & \dfrac{\partial f_3}{\partial y} & \dfrac{\partial f_3}{\partial z} \end{bmatrix} \;=\; \begin{bmatrix} y & x & 0 \\ 0 & z & y \\ z & 0 & x \end{bmatrix}$$

Therefore,

$$f'(1,0,-1) \;=\; \begin{bmatrix} 0 & 1 & 0 \\ 0 & -1 & 0 \\ -1 & 0 & 1 \end{bmatrix} .$$

As row 1 and 2 are multiples of one another, the row rank of $f'(1, 0, -1) = 2 \neq 3 = \dim \mathbb{R}^3$. This implies that $f'(1, 0, -1)$ is not onto, and, hence, f is not invertible in a neighbourhood of the point $(1, 0, -1)$.

Answer The function

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 , \qquad f(x,y,z) \;=\; (xy, yz, xz) ,$$

does not have a differentiable inverse in a neighbourhood of the point $\alpha = (1, 0, -1)$.

Exercise 5.

Determine whether the function

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 , \qquad f(x,y,z) \;=\; (xe^y, xyz, \ln|z|) ,$$

has a differentiable inverse in a neighbourhood of the point $\alpha = (2,0,1)$.

If so, find the derivative of its inverse function at $f(2,0,1)$.

Solution We first let

$$\begin{aligned} \eta := f(2,0,1) &= (2,0,0) \\ f_1(x,y,z) &:= xe^y \\ f_2(x,y,z) &:= xyz \\ f_3(x,y,z) &:= \ln|z| . \end{aligned}$$

By the inverse function theorem, f has a differentiable inverse in a neighbourhood of $\alpha = (2,0,1)$ if its derivative function $f'(2,0,1)$ is onto. Now the linear map $f'(x,y,z)$ is represented by the matrix

$$f'(x,y,z) \;=\; \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} \end{bmatrix} \;=\; \begin{bmatrix} e^y & xe^y & 0 \\ yz & xz & xy \\ 0 & 0 & \frac{1}{z} \end{bmatrix} .$$

Therefore,

$$f'(2,0,1) \;=\; \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

As $f'(2,0,1)$ is in echelon form, we conclude that $f'(2,0,1)$ is of full rank and is onto. Hence, f is invertible in a neighbourhood of the point $(2,0,1)$, and the derivative function of its inverse at $f(2,0,1)$ is

$$(f|^{-1})'(2,0,0) \;=\; [f'(2,0,1)]^{-1} \;=\; \begin{bmatrix} 1 & -1 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

Answer The function

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 , \qquad f(x,y,z) \;=\; (xe^y, xyz, \ln|z|) ,$$

is invertible in a neighbourhood of the point $\alpha = (2,0,1)$, and the derivative function of its inverse at $f(2,0,1)$ is

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

Now we adapt this argument to $\mathbb{R}^n$:

Theorem 9.2.2 (Inverse-Function Theorem) *Suppose $O \subset \mathbb{R}^m$ is open, $f\colon O \to \mathbb{R}^m$ is differentiable, and Df is invertible at a point $a \in O$ and continuous at a. Then there exist neighborhoods $U \subset O$ of a and V of $b := f(a) \in \mathbb{R}^m$ such that f is a bijection from U to V [that is, f is one-to-one on U and $f(U) = V$]. The inverse $g\colon V \to U$ of f is differentiable with $Dg(y) = (Df(g(y)))^{-1}$. Furthermore, if f is C^r (that is, all partial derivatives of f up to order r exist and are continuous) on U, then so is its inverse.*

Proof The proof is actually the same as before. We only need to replace various numbers by linear maps and some absolute values by norms.

The space. The contraction acts in $\mathbb{R}^m$.

The map. For any given $y \in \mathbb{R}^m$, consider the map

$$\varphi_y(x) := x + Df(a)^{-1}(y - f(x))$$

on O. Notice that $\varphi_y(x) = x$ if and only if $f(x) = y$, so we try to find a unique fixed point for φ_y. We need a set W on which it is a contraction.

The contraction property. Let $A := Df(a)$, $\alpha < \|A^{-1}\|^{-1}/2$, and, using continuity of Df at a, take $\epsilon > 0$ such that $\|Df(x) - A\| < \alpha$ for x in the closure of $W := B(a, \epsilon)$. To see that φ_y is a contraction, note that

$$\|D\varphi_y(x)\| = \|\operatorname{Id} - A^{-1}Df(x)\| = \|A^{-1}(A - Df(x))\| < \|A^{-1}\|\alpha =: \lambda < 1/2$$

for $x \in W$ and apply Corollary 2.2.15 to get $\|\varphi_y(x) - \varphi_y(x')\| \le \lambda\|x - x'\|$ for $x, x' \in W$. Therefore, by Proposition 2.2.20 there is a neighborhood V of b such that φ_y is a contraction of $\overline{W}$ for all $y \in V$ and has a unique fixed point $g(y) \in W$ (which depends continuously on y). $U := g(V) = W \cap f^{-1}(V)$ is open.

The determinant of $Df(x)$ depends continuously on Df and hence is continuous at a as a function of x. Thus, by taking V (and hence U) smaller, if necessary, we may assume $\det Df \neq 0$ on U and therefore that $Df(x)$ is invertible on U.

For $y = f(x) \in V$ we want to show that $Dg(y)$ exists and is the inverse of $B := Df(g(y))$. Take $y + k = f(x + h) \in V$. Then

$$\text{(9.2.1)} \qquad \begin{aligned}\frac{\|h\|}{2} &\ge \|\varphi_y(x+h) - \varphi_y(x)\| = \|h + A^{-1}(f(x) - f(x+h))\| \\ &= \|h - A^{-1}k\| \ge \|h\| - \|A^{-1}\|\|k\|,\end{aligned}$$

so

$$\frac{\|k\|}{\alpha} > \|A^{-1}\|\|k\| \ge \frac{\|h\|}{2} \qquad \text{and} \qquad \frac{1}{\|k\|} < \frac{2}{\alpha\|h\|}.$$

Since $g(y + k) - g(y) - B^{-1}k = h - B^{-1}k = -B^{-1}(f(x + h) - f(x) - Bh)$, we get

$$\frac{\|g(y+k) - g(y) - B^{-1}k\|}{\|k\|} < \frac{\|B^{-1}\|}{\alpha/2}\frac{\|f(x+h) - f(x) - Bh\|}{\|h\|} \xrightarrow[\|h\| \le 2\|k\|/\alpha \to 0]{} 0,$$

which proves $Dg(y) = B^{-1}$.

Finally, suppose $f \in C^r$ and $g \in C^k$ for some $k < r$. Then $Df(g(y)) \in C^k$ and so is its inverse Dg by using the formula for matrix inverses (the entries of A^{-1} are polynomials in those of A divided by $\det A \neq 0$). Thus, $g \in C^{k+1}$. □