Problem 22. Let $I : \mathscr{C}([0,1];\mathbb{R}) \to \mathbb{R}$ be defined by

$$I(f) = \int_0^1 f(x)^2 dx\,.$$

Show that I is differentiable at every "point" $f \in \mathscr{C}([0,1];\mathbb{R})$.
Hint: Figure out what $(DI)(f)$ is by computing $I(f+h) - I(f)$, where $h \in \mathscr{C}([0,1];\mathbb{R})$ is a "small" continuous function.
Remark. A map from a space of functions such as $\mathscr{C}([0,1];\mathbb{R})$ to a scalar field such as $\mathbb{R}$ or $\mathbb{C}$ is usually called a ***functional***. The derivative of a functional I is usually denoted by δI instead of DI.

Proof. For each $f \in \mathscr{C}([0,1];\mathbb{R})$, define $L_f(h) = 2\int_0^1 f(x)h(x)dx$.
claim: $L_f \in \mathscr{B}(\mathscr{C}([0,1];\mathbb{R}),\mathbb{R})$.
Proof of claim: It is trivial that $L_f \in \mathscr{L}(\mathscr{C}([0,1];\mathbb{R}),\mathbb{R})$. Let $h \in \mathscr{C}([0,1];\mathbb{R})$. Then

$$\big|L_f(h)\big| \leqslant 2\int_0^1 \big|f(x)\big|\big|h(x)\big|dx \leqslant 2\|f\|_\infty \|h\|_\infty\,;$$

thus

$$\|L_f\|_{\mathscr{B}(\mathscr{C}([0,1];\mathbb{R}),\mathbb{R})} = \sup_{\|h\|_\infty = 1} \big|L_f(h)\big| \leqslant 2\|f\|_\infty < \infty\,.$$

Claim: $\displaystyle\lim_{\|h\|_\infty \to 0} \frac{\big|I(f+h) - I(f) - L_f(h)\big|}{\|h\|_\infty} = 0.$
Proof of claim: Since

$$\begin{aligned}\big|I(f+h) - I(f) - L_f(h)\big| &= \Big|\int_0^1 \Big[\big(f(x)+h(x)\big)^2 - f(x)^2 - 2f(x)h(x)\Big]dx\Big| \\ &= \Big|\int_0^1 h(x)^2 dx\Big| \leqslant \|h\|_\infty^2\,,\end{aligned}$$

by the sandwich lemma we conclude that

$$0 \leqslant \lim_{\|h\|_\infty \to 0} \frac{\big|I(f+h) - I(f) - L_f(h)\big|}{\|h\|_\infty} \leqslant \lim_{\|h\|_\infty \to 0} \frac{\|h\|_\infty^2}{\|h\|_\infty} = 0\,.$$

Therefore, I is differentiable at f, and $(DI)(f)(h) = L_f(h)$. □

Problem 30. Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable. Assume that for all $x \in \mathbb{R}$, $0 \leqslant f'(x) \leqslant f(x)$. Show that $g(x) = e^{-x}f(x)$ is decreasing. If f vanishes at some point, conclude that f is zero.

Proof. To see that g is decreasing, we compute the derivative of g and find that

$$g'(x) = -e^{-x}f(x) + e^{-x}f'(x) = e^{-x}\big(f'(x) - f(x)\big) \leqslant 0\,;$$

thus g is a decreasing function. Now suppose that $f(c) = 0$ for some $c \in \mathbb{R}$.

1. Since g is decreasing, $g(x) \leqslant g(c) = 0$ for all $x \geqslant c$; thus $f(x) = e^x g(x) = 0$ for all $x \geqslant c$.

2. Since $f'(x) \geqslant 0$, f is an increasing function, thus $f(x) \leqslant f(c) = 0$ for all $x \leqslant c$. Since f is assumed to be non-negative, we must have $f(x) = 0$ for all $x \leqslant c$.

Combining 1 and 2, we conclude that $f(x) = 0$ for all $x \in \mathbb{R}$. □

Problem 32. 1. If $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $g : B \subseteq \mathbb{R}^m \to \mathbb{R}^\ell$ are twice differentiable and $f(A) \subseteq B$, then for $x_0 \in A$, $u, v \in \mathbb{R}^n$, show that

$$\begin{aligned} &D^2(g \circ f)(x_0)(u,v) \\ &\quad = (D^2 g)(f(x_0))\big((Df)(x_0)(u), Df(x_0)(v)\big) + (Dg)(f(x_0))\big((D^2 f)(x_0)(u,v)\big)\,. \end{aligned}$$

2. If $p : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map plus some constant; that is, $p(x) = Lx + c$ for some $L \in \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$, and $f : A \subseteq \mathbb{R}^m \to \mathbb{R}^s$ is k-times differentiable, prove that

$$D^k(f \circ p)(x_0)(u^{(1)}, \cdots, u^{(k)}) = (D^k f)\big(p(x_0)\big)\big((Dp)(x_0)(u^{(1)}), \cdots, (Dp)(x_0)(u^{(k)}\big)\,.$$

Proof. 1. First of all, we show that $g \circ f$ is twice differentiable. Since g and f are both differentiable, the chain rule implies that $g \circ f$ is differentiable and

$$D(g \circ f)(x) = (Dg)(f(x))(Df)(x) = \Big(\big((Dg) \circ f\big)(Df)\Big)(x)\,.$$

Since g and f are twice differentiable, Dg and Df are differentiable. By the chain rule again, $(Dg) \circ f$ is differentiable; thus the product rule implies that $\big((Dg) \circ f\big)(Df)$ is differentiable. Therefore, $g \circ f$ is twice differentiable.

Now by Proposition 6.69 in 共筆, we have

$$D^2(g \circ f)(x_0)(u,v) = \sum_{i,j=1}^{n} \frac{\partial^2 (g \circ f)}{\partial x_j \partial x_i}(x_0) u_i v_j\,.$$

By the chain rule,

$$\begin{aligned} \frac{\partial^2 (g \circ f)}{\partial x_j \partial x_i}(x_0) &= \frac{\partial}{\partial x_j}\Big|_{x=x_0} \frac{\partial (g \circ f)}{\partial x_i}(x) = \frac{\partial}{\partial x_j}\Big|_{x=x_0} \sum_{k=1}^{m} \Big[\frac{\partial g}{\partial y_k}(f(x)) \frac{\partial f_k}{\partial x_i}(x)\Big] \\ &= \sum_{k=1}^{m} \frac{\partial}{\partial x_j}\Big|_{x=x_0} \Big[\frac{\partial g}{\partial y_k}(f(x)) \frac{\partial f_k}{\partial x_i}(x)\Big] \\ &= \sum_{k=1}^{m} \sum_{\ell=1}^{m} \frac{\partial^2 g}{\partial y_\ell \partial y_k}(f(x_0)) \frac{\partial f_\ell}{\partial x_j}(x_0) \frac{\partial f_k}{\partial x_i}(x_0) + \sum_{k=1}^{m} \frac{\partial g}{\partial y_k}(f(x_0)) \frac{\partial^2 f_k}{\partial x_j \partial x_i}(x_0)\,; \end{aligned}$$

thus

$$
\begin{aligned}
&D^2(g\circ f)(x_0)(u,v)\\
&\quad=\sum_{i,j=1}^{n}\Big[\sum_{k,\ell=1}^{m}\frac{\partial^2 g}{\partial y_\ell \partial y_k}(f(x_0))\frac{\partial f_\ell}{\partial x_j}(x_0)\frac{\partial f_k}{\partial x_i}(x_0)+\sum_{k=1}^{m}\frac{\partial g}{\partial y_k}(f(x_0))\frac{\partial^2 f_k}{\partial x_j\partial x_i}(x_0)\Big]u_i v_j\\
&\quad=\sum_{k,\ell=1}^{m}\frac{\partial^2 g}{\partial y_\ell \partial y_k}(f(x_0))\Big(\sum_{j=1}^{n}\frac{\partial f_\ell}{\partial x_j}(x_0)v_j\Big)\Big(\sum_{i=1}^{n}\frac{\partial f_k}{\partial x_i}(x_0)u_i\Big)\\
&\qquad+\sum_{k=1}^{m}\frac{\partial g}{\partial y_k}(f(x_0))\Big(\sum_{i,j=1}^{n}\frac{\partial^2 f_k}{\partial x_j\partial x_i}(x_0)u_i v_j\Big).
\end{aligned}
$$

Letting $\big((Df)(x_0)(w)\big)_r$ denote the r-th component of $(Df)(x_0)(w)$, we obtain that

$$
\begin{aligned}
&D^2(g\circ f)(x_0)(u,v)\\
&\quad=\sum_{k,\ell=1}^{m}\frac{\partial^2 g}{\partial y_\ell \partial y_k}(f(x_0))\big((Df)(x_0)(v)\big)_\ell\big((Df)(x_0)(u)\big)_k\\
&\qquad+\sum_{k=1}^{m}\frac{\partial g}{\partial y_k}(f(x_0))\big(D^2 f)(x_0)(u,v)\big)_k\\
&\quad=(D^2 g)(f(x_0))\big((Df)(x_0)u,(Df)(x_0)v\big)+(Dg)(f(x_0))\big((D^2 f)(x_0)(u,v)\big).
\end{aligned}
$$

2. The validity of the desired equality for the case $k=1$ is the chain rule. Suppose that the desired holds for $k=K$. Then for $k=K+1$, by Corollary 6.70 in 共筆 we obtain that

$$
\begin{aligned}
D^{K+1}(f\circ p)(x_0)(u^{(1)},\cdots,u^{(K+1)})&=\sum_{j=1}^{n}u_j^{(K+1)}\frac{\partial}{\partial x_j}\Big|_{x=x_0}(D^K(f\circ p))(x)(u^{(1)},\cdots,u^{(K)})\\
&=\sum_{j=1}^{n}u_j^{(K+1)}\frac{\partial}{\partial x_j}\Big|_{x=x_0}(D^k f)\big(p(x)\big)\big((Dp)(x)(u^{(1)}),\cdots,(Dp)(x)(u^{(k)}\big).
\end{aligned}
$$

Noting that $(Dp)(x)(u^r)=Lu^{(r)}$ (which is independent of x), by Proposition 6.69 in 共筆 we find that

$$
\begin{aligned}
&(D^K f)\big(p(x)\big)\big((Dp)(x)(u^{(1)}),\cdots,(Dp)(x)(u^{(K)}\big)\\
&\quad=\sum_{j_1,\cdots,j_K=1}^{n}\frac{\partial^K f}{\partial y_{j_K}\cdots\partial y_{j_1}}(p(x))(Lu^{(1)})_{j_1}\cdots(Lu^{(K)})_{j_K},
\end{aligned}
$$

where $(Lu^r)_s$ denotes the s-th component of the vector $Lu^{(r)}$. As a consequence,

$$
\begin{aligned}
&\sum_{j=1}^n u_j^{(K+1)} \frac{\partial}{\partial x_j}\Big|_{x=x_0} (D^K f)\big(p(x)\big)\big((Dp)(x)(u^{(1)}),\cdots,(Dp)(x)(u^{(K)})\big) \\
&= \sum_{j=1}^n u_j^{(K+1)} \frac{\partial}{\partial x_j}\Big|_{x=x_0} \sum_{j_1,\cdots,j_K=1}^m \frac{\partial^K f}{\partial y_{j_K}\cdots\partial y_{j_1}}(p(x))(Lu^{(1)})_{j_1}\cdots(Lu^{(K)})_{j_K} \\
&= \sum_{j=1}^n u_j^{(K+1)} \sum_{j_1,\cdots,j_K,j_{K+1}=1}^m \frac{\partial^{K+1} f}{\partial y_{j_{K+1}}\cdots\partial y_{j_1}}(p(x_0))\frac{\partial p_{j_{K+1}}}{\partial x_j}(x_0)(Lu^{(1)})_{j_1}\cdots(Lu^{(K)})_{j_K} \\
&= \sum_{j_1,\cdots,j_K,j_{K+1}=1}^m \frac{\partial^{K+1} f}{\partial y_{j_{K+1}}\cdots\partial y_{j_1}}(p(x_0))\Big(\sum_{j=1}^n u_j^{(K+1)}\frac{\partial p_{j_{K+1}}}{\partial x_j}(x_0)\Big)(Lu^{(1)})_{j_1}\cdots(Lu^{(K)})_{j_K} \\
&= \sum_{j_1,\cdots,j_K,j_{K+1}=1}^m \frac{\partial^{K+1} f}{\partial y_{j_{K+1}}\cdots\partial y_{j_1}}(p(x_0))\big((Dp)(x_0)u^{(K+1)}\big)_{j_{K+1}}(Lu^{(1)})_{j_1}\cdots(Lu^{(K)})_{j_K} \\
&= \sum_{j_1,\cdots,j_K,j_{K+1}=1}^m \frac{\partial^{K+1} f}{\partial y_{j_{K+1}}\cdots\partial y_{j_1}}(p(x_0))(Lu^{(1)})_{j_1}\cdots(Lu^{(K)})_{j_K}(Lu^{(K+1)})_{j_{K+1}} \\
&= (D^{K+1} f)\big(p(x_0)\big)\big((Dp)(x_0)(u^{(1)}),\cdots,(Dp)(x_0)(u^{(K+1)}\big)
\end{aligned}
$$

which shows the validity of the desired equality for the case $k = K+1$. □

Problem 34. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable, and Df is a constant map in $\mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$; that is, $(Df)(x_1)(u) = (Df)(x_2)(u)$ for all $x_1, x_2 \in \mathbb{R}^n$ and $u \in \mathbb{R}^n$. Show that f is a linear term plus a constant and that the linear part of f is the constant value of Df.

Proof. Since Df is a constant map, Df is continuous; thus $f \in \mathscr{C}^1$. Therefore, the Taylor Theorem implies that

$$f(x) = f(0) + (Df)(c)(x-0)$$

for some c on the line segment joining x and 0. Let $L = (Df)(x)$. Then

$$f(x) = f(0) + L(x-0) = Lx + f(0)\,. \qquad \square$$

Problem 38. Prove Corollary 7.5; that is, show that if $\mathcal{U} \subseteq \mathbb{R}^n$ is open, $f : \mathcal{U} \to \mathbb{R}^n$ is of class $\mathscr{C}^1$, and $(Df)(x)$ is invertible for all $x \in \mathcal{U}$, then $f(\mathcal{W})$ is open for every open set $\mathcal{W} \subseteq \mathcal{U}$.

Proof. Let $\mathcal{W} \subseteq \mathcal{U}$ be an open set. For each $x \in \mathcal{W}$, $(Df)(x)$ is invertible; thus the inverse function theorem implies that there exists $\delta_x > 0$ such that

(a) $D(x,\delta_x) \subseteq \mathcal{W}$; (b) $f\big(D(x,\delta_x)\big)$ is open; (c) $f : D(x,\delta) \to f\big(D(x,\delta_x)\big)$ is one-to-one and onto.

Since $\mathcal{W} = \bigcup_{x\in\mathcal{W}} D(x,\delta_x)$,

$$f(\mathcal{W}) = \bigcup_{x\in\mathcal{W}} f\big(D(x,\delta_x)\big)$$

is the union of infinitely many open sets; thus $f(\mathcal{W})$ is open. □

Problem 40. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be of class $\mathscr{C}^1$, and for some $(a,b) \in \mathbb{R}^2$, $f(a,b) = 0$ and $f_y(a,b) \neq 0$. Show that there exist open neighborhoods $\mathcal{U}$ of a and $\mathcal{V}$ of b such that every $y \in \mathcal{V}$ corresponds to a unique $x \in \mathcal{U}$ such that $f(x,y) = 0$. In other words, there exists a function $y = y(x)$ such that $y(a) = b$ and $f(x, y(x)) = 0$ for all $x \in \mathcal{U}$.

Proof. Let $z = (x,y)$ and $w = (u,v)$, where $x,y,u,v \in \mathbb{R}$. Define $w = F(z)$, where F is given by $F(x,y) = \big(x, f(x,y)\big)$. Then $F : \mathcal{D} \to \mathbb{R}^2$, and

$$\big[(DF)(z)\big] = \begin{bmatrix} 1 & 0 \\ f_x(x,y) & f_y(x,y) \end{bmatrix}.$$

We note that the Jacobian of F at (a,b) is $f_y(a,b) \neq 0$, so the inverse function theorem implies that there exists open neighborhoods $\mathcal{O} \subseteq \mathbb{R}^2$ of (a,b) and $\mathcal{W} \subseteq \mathbb{R}^2$ of $\big(a, f(a,b)\big) = (a,0)$ such that

(a) $F : \mathcal{O} \to \mathcal{W}$ is one-to-one and onto;

(b) the inverse function $F^{-1} : \mathcal{W} \to \mathcal{O}$ is of class $\mathscr{C}^r$;

(c) $(DF^{-1})\big(x, f(x,y)\big) = \big((DF)(x,y)\big)^{-1}$.

W.L.O.G. we can assume that $\mathcal{O} = \mathcal{U} \times \mathcal{V}$, where $\mathcal{U} \subseteq \mathbb{R}$ and $\mathcal{V} \subseteq \mathbb{R}$ are open, and $a \in \mathcal{U}$, $b \in \mathcal{V}$.

Write $F^{-1}(u,v) = \big(\varphi(u,v), \psi(u,v)\big)$, where $\varphi : \mathcal{W} \to \mathcal{U}$ and $\psi : \mathcal{W} \to \mathcal{V}$. Then

$$(u,v) = F\big(\varphi(u,v), \psi(u,v)\big) = \big(\varphi(u,v), f(u, \psi(u,v))\big)$$

which implies that $\varphi(u,v) = u$ and $v = f(u, \psi(u,v))$. Let $y(x) = \psi(x,0)$. Then $\big(u, f(u)\big) \in \mathcal{U} \times \mathcal{V}$ is the unique point satisfying $f\big(u, y(u)\big) = 0$ if $u \in \mathcal{U}$. Therefore, $f : \mathcal{U} \to \mathcal{V}$, and

$$f\big(x, y(x)\big) = 0 \qquad \forall\, x \in \mathcal{U}\,.$$

Since $G(a,b) = (a,0) = G\big(a, f(a)\big)$, $(a,b), \big(a, f(a)\big) \in \mathcal{O}$, and $G : \mathcal{O} \to \mathcal{W}$ is one-to-one, we must have $b = f(a)$. $\square$