Homework #5 Solutions

Thayer Anderson, Davis Lazowski, Handong Park, Rohil Prasad
Eric Peterson

1 For submission to Thayer Anderson

Problem 1.1. Let P be a partition of $[a, b]$ into subintervals S_j, and for each S_j choose a point $s_j \in S_j$. We will call this data a *marked partition*, and we say that the maximum volume of any rectangle in the partition is the *mesh size* of P.

For an integrable function $f\colon [a, b] \to \mathbb{R}$ and any $\varepsilon > 0$, show that there must exist a corresponding $\delta > 0$ such that if the mesh size of a marked partition P is less than δ then

$$\sum_{S_j} \left| f(s_j)\operatorname{vol}(S_j) - \int_{S_j} f(x)\,\mathrm{d}x \right| < \varepsilon.$$

(Hint: Have a look at the proof of our characterization of the "integrability" condition for clues.)

Solution. Take Q to be a partition so that the associated Darboux sums satisfy

$$\mathcal{U}_f(Q) - \mathcal{L}_f(Q) < \varepsilon/2,$$

and set $\delta = \operatorname{mesh}(Q)$, and let n be the number of points in Q, and let $|f| < M$ be a global bound on f. Suppose, then, that P is some auxiliary partition satisfying $\operatorname{mesh}(P) < \varepsilon/(2Mn)$; then any subrectangle of P is either contained in a subrectangle of Q or it meets exactly one point of Q. In the first case, the upper and lower summands associated to that subrectangle (and hence to the volume of that subrectangle times any marked point on it) are bounded by the upper and lower summands associated to the parent subrectangle in Q. In the second case, the value of the rectangle differs from the Darboux summands by at most $2M \cdot \varepsilon/2 \cdot 1/(2Mn) = \varepsilon/(2n)$. The total difference is thus at most $\varepsilon/2$, and since the Darboux upper- and lower-sums themselves differ by at most $\varepsilon/2$, the marked sum differs from the integral by at most $\varepsilon/2 + \varepsilon/2 = \varepsilon$.[1] (ECP)

Problem 1.2. Let $\{f_n\}$ be a *nondecreasing* sequence of integrable functions defined on an interval $[a, b]$, and suppose that for each $x \in [a, b]$ the limit $\lim_{n\to\infty} f_n(x)$ exists. We use this to define a new function $f(x) = \lim_{n\to\infty} f_n(x)$.

Suppose that this new function is itself Riemann integrable. We will conclude the limit interchange law

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} \left(\int_a^b f_n(x)\,\mathrm{d}x \right).$$

0. We first need to establish some vocabulary; there's nothing to prove in this bullet point. Consider the sequence of errors $g_n = f - f_n$, and write $\eta = \varepsilon/(b - a + 1)$ for an auxiliary error usefully satisfying $\eta \operatorname{vol}[a, b] < \varepsilon$. For each n, use the previous Problem to select a δ_n such that if the mesh size of a marked partition P is beneath δ_n, then

$$\sum_{S_j \in P} \left| g_n(s_j)\operatorname{vol}(S_j) - \int_{S_j} g_n(x)\,\mathrm{d}x \right| < 2^{-n}\eta.$$

[1] This is not super well-stated. Sorry.

1. For each x, let $N(x)$ denote the first integer such that $n \geq N$ forces $g_n(x) < \eta$, and then let $\delta(x) = \delta_{N(x)}$. Show that there exists a marked partition satisfying $\operatorname{vol}(S_j) < \delta(s_j)$.
2. Write P_n for the collection of those rectangles S_j in the partition (guaranteed by the previous part) which satisfy $N(s_j) = n$; since there are finitely many rectangles in the partition, we let N denote the largest n so appearing. Justify all of the steps in the following sequence:

$$0 \leq \int_a^b g_N(x)\,\mathrm{d}x = \sum_{S_j \in P} \int_{S_j} g_N(x)\,\mathrm{d}x = \sum_{n=1}^{N} \sum_{S_j \in P_n} \int_{S_j} g_N(x)\,\mathrm{d}x \leq \sum_{n=1}^{N} \sum_{S_j \in P_n} \int_{S_j} g_n(x)\,\mathrm{d}x$$
$$\leq \sum_{n=1}^{N} \left(\eta 2^{-n} + \sum_{S_j \in P_n} g_n(s_j) \operatorname{vol}(S_j) \right) < \sum_{n=1}^{N} \left(\eta 2^{-n} + \sum_{S_j \in P_n} \eta \operatorname{vol}(S_j) \right) < \eta(b - a + 1) = \varepsilon.$$

3. Use this to finish the argument and conclude the interchange law.

Solution. 1. Our proof of this mimicks the proof of Heine–Borel, the claim that all closed and bounded intervals are compact subsets of $\mathbb{R}$. Let $S \subseteq [a, b]$ be the set of endpoints of initial segments for which the claim is true:

$$L = \{l \in [a, b] \mid [a, l] \text{ admits a marked partition } S \text{ with } \operatorname{vol}(S_j) < \delta(s_j)\}.$$

This set L is nonempty, since $\operatorname{vol}[a, a] = 0$ means that $a \in L$ holds. It therefore has a supremum: set $l = \sup L$, and suppose $l < b$. In this case, consider $0 < \delta(l)$, evaluated at this supremum: first choosing a marked partition of $[a, l - \delta(l)/4]$, we can introduce a subrectangle of radius $\delta(l)/3$ centered at l so that the resulting marked partition satisfies the desired property out through $l + \delta(l)/3$, contradicting maximality. It follows, then, that $l = \sup L = b$. Similarly, by evaluating $\delta(b)$ at b, we see that we can add a subrectangle containing b and reaching down into L, so that b is not just a supremum but an achieved maximum. This was the desired claim all along.

2. We work step-by-step.

$$0 \leq \int_a^b g_N(x)\,\mathrm{d}x$$

This is certainly true: each g_N is a nonnegative function, so each lower sum is nonnegative, so each integral is nonnegative.

$$\int_a^b g_N(x)\,\mathrm{d}x = \sum_{S_j \in P} \int_{S_j} g_N(x)\,\mathrm{d}x$$

By subdividing $[a, b]$ into disjoint regions, we can write the integral of the disjoint union of rectangular regions into the sum of the integrals.

$$\sum_{S_j \in P} \int_{S_j} g_N(x)\,\mathrm{d}x = \sum_{n=1}^{N} \sum_{S_j \in P_n} \int_{S_j} g_N(x)\,\mathrm{d}x$$

By collecting the rectangular regions into disjoint collections P_n, we can again write the sum over the entire collection P as a sum of smaller sums over the subcollections P_n.

$$\sum_{n=1}^{N} \sum_{S_j \in P_n} \int_{S_j} g_N(x)\,\mathrm{d}x \leq \sum_{n=1}^{N} \sum_{S_j \in P_n} \int_{S_j} g_n(x)\,\mathrm{d}x$$

By monotonicity, it is always true that $g_N \leq g_n$.

$$\sum_{n=1}^{N} \sum_{S_j \in P_n} \int_{S_j} g_n(x)\,\mathrm{d}x \leq \sum_{n=1}^{N} \left(\eta 2^{-n} + \sum_{S_j \in P_n} g_n(s_j)\,\mathrm{vol}(S_j) \right)$$

This is exactly Part 0, rearranged.

$$\sum_{n=1}^{N} \left(\eta 2^{-n} + \sum_{S_j \in P_n} g_n(s_j)\,\mathrm{vol}(S_j) \right) < \sum_{n=1}^{N} \left(\eta 2^{-n} + \sum_{S_j \in P_n} \eta\,\mathrm{vol}(S_j) \right)$$

Our partition was chosen so that the marked points satisfy $g_n(s_j) < \eta$.

$$\sum_{n=1}^{N} \left(\eta 2^{-n} + \sum_{S_j \in P_n} \eta\,\mathrm{vol}(S_j) \right) < \eta(b - a + 1) = \varepsilon.$$

The total volume of S_j is $b - a$, and the sum $\sum_{n=1}^{N} 2^{-n}$ is bounded by 1. (ECP)

Problem 1.3. We now use Problem 1.2 to describe a new kind of integrability for *unbounded* functions. Given an unbounded function $f \colon [a, b] \to \mathbb{R}_{\geq 0} \cup \{\infty\}$, we set its N^{th} truncation to be $f^{\leq N}(x) = \min\{f(x), N\}$. We say that f is *improperly integrable* when the limit

$$\lim_{N \to \infty} \int_a^b f^{\leq N}(x)\,\mathrm{d}x =: \int_a^b f(x)\,\mathrm{d}x$$

exists.

1. If f is integrable to begin with, show that $f^{\leq N}$ is integrable. Conclude also that if f is integrable to begin with, then its improper integral agrees with its integral.
2. Show that if $f \colon [a, b] \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ is improperly integrable but not integrable, then the set of points where it takes the value ∞ is measure zero.
3. Show that this problem is compatible with the above one: for a sequence $f_n \colon [a, b] \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ of integrable functions converging pointwise to an integrable function $f \colon [a, b] \to \mathbb{R}_{\geq 0} \cup \{\infty\}$, demonstrate the interchange identity
$$\lim_{n \to \infty} \int f_n(x)\,\mathrm{d}x = \int f(x)\,\mathrm{d}x,$$
where the integrals are interpreted as improper integrals.

Solution. 1. If f is integrable, then it is also bounded, and hence $f^{\leq N} = f$ for large N. In turn, the limit is a limit of a constant sequence, so the new integral agrees with the old one.

2. Let $B = f^{-1}(\infty)$ denote the set of bad points, and suppose that B is not of measure zero, so that any open cover of it has volume at least ε for some $\varepsilon > 0$. Using the fact that $f^{\leq N}|_B = N$ becomes the constant function, this points a lower bound on the lower sums associated to f:
$$\mathcal{L}_{f^{\leq N}}(P) \geq \varepsilon \cdot N.$$
As N tends to ∞, these lower sums tend to ∞, and hence f is *not* improperly integrable.

3. This, too, happens in steps:

$$\int f = \int \lim_{n\to\infty} f_n = \lim_{N\to\infty} \int \left(\lim_{n\to\infty} f_n\right)^{\leq N}$$

So far, we have just expanded definitions: the definition of f, and then the definition of the improper integral of f.

$$\lim_{N\to\infty} \int \left(\lim_{n\to\infty} f_n\right)^{\leq N} = \lim_{N\to\infty} \int \lim_{n\to\infty} (f_n)^{\leq N}$$

Because the sequence of functions is everywhere convergent, then trimming the result by N agrees with trimming the sequence itself by N.

$$\lim_{N\to\infty} \int \lim_{n\to\infty} (f_n)^{\leq N} = \lim_{N\to\infty} \lim_{n\to\infty} \int (f_n)^{\leq N}$$

Now we use the result of the previous problem, since $(f_n)^{\leq N}$ is a sequence of bounded functions for any fixed N.

$$\lim_{N\to\infty} \lim_{n\to\infty} \int (f_n)^{\leq N} = \lim_{n\to\infty} \lim_{N\to\infty} \int (f_n)^{\leq N}$$

This is actually delicate: limits in general don't commute, but in this case it is simultaneously true that in the grid of values $(\int (f_n)^{\leq N}$ indexed by n and N, it is *simultaneously* true that increasing n or increasing N causes the values to go up. This is enough for the limits to commute.

$$\lim_{n\to\infty} \lim_{N\to\infty} \int (f_n)^{\leq N} = \lim_{n\to\infty} \int f_n.$$

Finally, we use the definition of the improper integral again. (ECP)

2 For submission to Davis Lazowski

Our goal in this section is to make the computation

$$\int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x = \sqrt{\pi},$$

which is a formula that comes up often when studying normal distributions (or "bell curves"). We will do this in many parts — I suggest doing them in order.

Problem 2.1. Define $f\colon \{r \mid r > 0\} \times (0, 2\pi) \to \mathbb{R}^2$ by $f(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$. Show that f is injective, compute $D_{(r,\theta)}f$, and show that $\det D_{(r,\theta)}f \neq 0$ for all (r,θ). Show that the image of f is the following set A from an old homework:

$$A = \mathbb{R} \setminus ([0,\infty) \times \{0\}) = \{(x,y) \in \mathbb{R}^2 \mid \text{if } x \geq 0, \text{ then } y \neq 0\}.$$

Solution.

- f **is injective**: Observe that $||f(r,\theta)|| = r$. So $f(r,\theta)$ has a unique norm dependent only on its r. Therefore if $f(r,\theta) = f(r',\theta')$, then $r = r'$, otherwise the norms would be different. So we can divide out by r and it's enough to show that the function

$$\theta \to (\cos\theta, \sin\theta)$$

is injective over $(0, 2\pi)$. Suppose $\cos\theta' = \cos\theta$. Then either $\theta' = \theta$ or $\theta' = 2\pi - \theta$. But $\sin(2\pi - \theta) = -\sin\theta$. So unless $\sin\theta = 0$, therefore $\theta' = \theta$. If $\sin\theta = 0$, then $\theta = \pi$. So $2\pi - \theta = \theta$.

- **Compute the Jacobian and show its determinant is nonzero**: Directly, the derivative is

$$D_{(r,\theta)}f = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

  Then

$$\det D_{(r,\theta)}f = r(\cos^2\theta + \sin^2\theta) = r$$

  And since $r \neq 0$ in our domain, the determinant is nonzero.

- **Show that the image is the set** A: If $x \geq 0$, and $y \neq 0$, then we can represent the point (x, y) uniquely as $r(x_c, y_c)$, where $r = \sqrt{x^2+y^2}$ and (x_c, y_c) is a point on the unit circle. There is a one-to-one map between points on the unit circle and angles θ. Our set of angles is $(0, 2\pi)$, so includes all angles except $\theta = 0$, so the point on the unit circle $(1, 0)$. So the only points discluded are of the form $(r, 0)$, for all $r \geq 0$. But these are exactly the points left out in A. (DL)

Problem 2.2. We write $P = f^{-1}$ for the inverse function, so named after the *polar coordinate system*. Show that the components $P(x, y) = (r(x, y), \theta(x, y))$ take the form

$$r(x,y) = \sqrt{x^2+y^2}, \qquad \theta(x,y) = \begin{cases} \tan^{-1} y/x & \text{if } x > 0,\ y > 0, \\ \pi + \tan^{-1} y/x & \text{if } x < 0, \\ 2\pi + \tan^{-1} y/x & \text{if } x > 0,\ y < 0, \\ \pi/2 & \text{if } x = 0,\ y > 0, \\ 3\pi/2 & \text{if } x = 0,\ y < 0. \end{cases}$$

Calculate $D_{(x,y)}P$.

Solution. Since the inverse exists and is unique, it's enough to just show $f(r(x, y), \theta(x, y)) = (x, y)$. First, check the $\theta(x, y)$ part.

$$\begin{pmatrix} \cos\theta(x,y)) \\ \sin\theta(x,y) \end{pmatrix} = \begin{cases} \begin{pmatrix} \cos\tan^{-1} y/x \\ \sin\tan^{-1} y/x \end{pmatrix} = \begin{pmatrix} \frac{|x|}{\sqrt{x^2+y^2}} \\ \frac{|y|}{\sqrt{x^2+y^2}} \end{pmatrix} & \text{if } x > 0,\ y > 0, \\ \begin{pmatrix} \cos(\pi + \tan^{-1} y/x) \\ \sin(\pi + \tan^{-1} y/x) \end{pmatrix} = \begin{pmatrix} \frac{-|x|}{\sqrt{x^2+y^2}} \\ \frac{|y|}{\sqrt{x^2+y^2}} \end{pmatrix} & \text{if } x < 0, \\ \begin{pmatrix} \cos(2\pi - \tan^{-1} y/x) \\ \sin(2\pi - \tan^{-1} y/x) \end{pmatrix} = \begin{pmatrix} \frac{|x|}{\sqrt{x^2+y^2}} \\ \frac{-|y|}{\sqrt{x^2+y^2}} \end{pmatrix} & \text{if } x > 0,\ y < 0, \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \text{if } x = 0,\ y > 0, \\ \begin{pmatrix} -1 \\ 0 \end{pmatrix} & \text{if } x = 0,\ y < 0. \end{cases}$$

Multiplying through by $\sqrt{x^2+y^2}$, and observing that the absolute value signs work out to preserve the signs of x and y in each region, we get back (x, y). (DL)

Problem 2.3. 1. Fix values r_1, r_2, θ_1, and θ_2. Let $C \subseteq A$ be the region between the circles of radii r_1 and r_2 and the half-lines through 0 which make angles of θ_1 and θ_2 with the x–axis. If $h\colon C \to \mathbb{R}$ is integrable and $h(x, y) = g(r(x, y), \theta(x, y))$, show that

$$\int_C h = \int_{r_1}^{r_2} \int_{\theta_1}^{\theta_2} g(r,\theta) r \,\mathrm{d}\theta \,\mathrm{d}r.$$

2. Fix a value ρ, and let $B = \{(x, y) \mid x^2 + y^2 \le \rho^2\}$. Show that

$$\int_B h = \int_0^\rho \int_0^{2\pi} g(r, \theta) r \, \mathrm{d}\theta \, \mathrm{d}r.$$

Solution. 1. $f(r, \theta)$, as proved in our previous problems in this section, is a coordinate transform. So

$$\begin{aligned}\int_C h(x, y) dx dy &= \int_C g(r, \theta) \det D_{(r,\theta)} f dr d\theta \\ &= \int_C g(r, \theta) r dr d\theta\end{aligned}$$

We can now explicitly parametrise C in terms of r and θ as given in the problem:

$$= \int_{r_1}^{r_2} \int_{\theta_1}^{\theta_2} g(r, \theta) r dr d\theta$$

2. This result appears as the double limit of the previous result. Let $r_2 = \rho$, and take the limit as $r_1 \to 0$. If g is integrable over this region, then

$$\lim_{r_1 \to 0} \int_{r_1}^{\rho} \int_{\theta_1}^{\theta_2} g(r, \theta) r dr d\theta = \int_0^\rho \int_{\theta_1}^{\theta_2} g(r, \theta) r dr d\theta$$

Then take the limits $\theta_1 \to 0$, $\theta_2 \to 2\pi$.

$$\lim_{\theta_2 \to 2\pi} \lim_{\theta_1 \to 0} \int_0^\rho \int_0^{2\pi} g(r, \theta) r dr d\theta = \int_0^\rho \int_0^{2\pi} g(r, \theta) r dr d\theta$$

Where we can swap $\int_0^\rho$ and $\int_{\theta_1}^{\theta_2}$ when we take the limit by Fubini's theorem. (DL)

Problem 2.4. Again fix a value ρ, and set $D = [-\rho, \rho] \times [-\rho, \rho]$. Show both of the equalities

$$\int_B e^{-(x^2+y^2)} \, \mathrm{d}x \, \mathrm{d}y = \pi(1 - e^{-\rho^2}), \qquad \int_D e^{-(x^2+y^2)} \, \mathrm{d}x \, \mathrm{d}y = \left(\int_{-\rho}^{\rho} e^{-z^2} \, \mathrm{d}z \right)^2.$$

(Hint: It may be useful to solve **??** first.)

Solution.

- For the first equality, by 2.3.2

$$\int_B e^{-(x^2+y^2)} dx dy = \int_0^\rho \int_0^{2\pi} r e^{-r} dr d\theta$$

By 4.1

$$\int_0^\rho \int_0^{2\pi} r e^{-r^2} dr d\theta = \int_0^{2\pi} d\theta \int_0^\rho r e^{-r^2} dr = (2\pi) \left(\frac{1}{2} - \frac{e^{-\rho^2}}{2} \right) = \pi(1 - e^{-\rho^2})$$

- For the second equality, by 4.1

$$\int_D e^{-(x^2+y^2)} dx dy = \int_D e^{-x^2} e^{-y^2} dx dy = \int_{-\rho}^{\rho} e^{-x^2} dx \int_{-\rho}^{\rho} e^{-y^2} dy = \left(\int_{-\rho}^{\rho} e^{-z^2} dz \right)^2$$

Where the last step comes because dx, dy are just dummy variables. (DL)

Problem 2.5. Prove the limit equality

$$\lim_{\rho\to\infty}\int_B e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y = \lim_{\rho\to\infty}\int_D e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y,$$

and hence conclude

$$\int_{-\infty}^{\infty} e^{-x^2} = \sqrt{\pi}.$$

Solution. By the first equality in 2.4,

$$\lim_{\rho\to\infty}\int_B e^{-(x^2+y^2)} dxdy = \lim_{\rho\to\infty}\pi(1-e^{-\rho^2}) = \pi$$

Also, consider $D\setminus B$. This region has area $4\rho^2-\pi\rho^2=(4-\pi)\rho^2$. Also, over this region, $e^{-(x^2+y^2)}\le e^{-\rho^2}$. So

$$\lim_{\rho\to\infty}\int_{D\setminus B} e^{-(x^2+y^2)} dxdy \le \lim_{\rho\to\infty}(4-\pi)\rho^2 e^{-\rho^2} = 0$$

So therefore

$$\lim_{\rho\to\infty}\int_B e^{-(x^2+y^2)} dxdy = \lim_{\rho\to\infty}\int_D e^{-(x^2+y^2)} dxdy = \pi.$$

By the second equality in 2.4,

$$\begin{aligned}\pi = \lim_{\rho\to\infty}\int_D e^{-(x^2+y^2)} dxdy = \lim_{\rho\to\infty}\left(\int_{-\rho}^{\rho} e^{-x^2}dx\right)^2 = \left(\lim_{\rho\to\infty}\int_{-\rho}^{\rho} e^{-x^2}dx\right)^2 \\ \implies \sqrt{\pi} = \lim_{\rho\to\infty}\int_{-\rho}^{\rho} e^{-x^2}dx = \int_{-\infty}^{\infty} e^{-x^2}dx. \qquad \text{(DL)}\end{aligned}$$

3 For submission to Handong Park

Problem 3.1. Let $g_1, g_2\colon \mathbb{R}^2\to\mathbb{R}$ be continuously differentiable, and as in a homework problem from a few weeks ago assemble these into a function f according to

$$f(x,y) = \int_0^x g_1(t,0)\,\mathrm{d}t + \int_0^y g_2(x,t)\,\mathrm{d}t.$$

Back then we complained about how f can be poorly behaved if $\frac{\partial g_1}{\partial y}$ does not agree with $\frac{\partial g_2}{\partial x}$, so now we instead assume that they are equal. Show that under this assumption we have $\frac{\partial f}{\partial x}(x,y) = g_1(x,y)$.

Solution. This amounts to a Fubini juggle, with several applications of both versions of the single-variable fundamental theorem of calculus.

$$\begin{aligned}\frac{\partial}{\partial x}f(x,y) &= \frac{\partial}{\partial x}\int_0^x g_1(t,0)\,\mathrm{d}t + \frac{\partial}{\partial x}\int_0^y g_2(x,t)\,\mathrm{d}t \\ &= g_1(x,0) + \frac{\partial}{\partial x}\int_0^y\int_0^x \frac{\partial}{\partial s}g_2(s,t)\,\mathrm{d}s\,\mathrm{d}t \\ &= g_1(x,0) + \frac{\partial}{\partial x}\int_0^x\int_0^y \frac{\partial}{\partial s}g_2(s,t)\,\mathrm{d}t\,\mathrm{d}s \\ &= g_1(x,0) + \frac{\partial}{\partial x}\int_0^x\int_0^y \frac{\partial}{\partial t}g_1(s,t)\,\mathrm{d}t\,\mathrm{d}s \\ &= g_1(x,0) + \frac{\partial}{\partial x}\int_0^x (g_1(s,y)-g_1(s,0))\,\mathrm{d}s \\ &= g_1(x,0) + g_1(x,y) - g_1(x,0) = g_1(x,y). \qquad \text{(ECP)}\end{aligned}$$

Problem 3.2. Our goal in this Problem is to show that if U is a rectangle and g is a linear transformation, then the volume of $g(U)$ is the volume of U scaled up by $|\det g|$.

1. Let $g\colon \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation of one of the following types:
   (a) Fixing an index n and a scalar k, g takes the form $g(e_j) = \begin{cases} ke_n & \text{if } j = n, \\ e_j & \text{otherwise.} \end{cases}$
   (b) Fixing indices m and n, g takes the form $g(e_j) = \begin{cases} e_n + e_m & \text{if } j = n, \\ e_j & \text{otherwise.} \end{cases}$
   (c) Fixing indices m and n, g takes the form $g(e_j) = \begin{cases} e_m & \text{if } j = n, \\ e_n & \text{if } j = m, \\ e_j & \text{otherwise.} \end{cases}$
2. Conclude from this the case of a general linear transformation g.

Solution. 1. This problem is reasonably straightforward if you believe that integrals of geometric figures compute their volumes. In particular, the volume of a rectangle is given by the product of its side lengths, and the volume of a parallelopiped is computed by the product of its altitude lengths.
   (a) Scaling the n^{th} axis by k scales the volume of U by k, and the determinant of g is also k.
   (b) The effect of g is to shear the rectangle to a parallelogram-prism, without modifying the lengths of its altitudes. The volume of $g(U)$ is therefore the same as that of U, and also $\det g = 1$.
   (c) Swapping two of the coordinate axes does not change the product of side lengths, so the volume of $g(U)$ agrees with the volume of U. We also have $\det g = -1$, so that the absolute value of the determinant is 1, as required.
2. Using Gaussian elimination, a generic matrix can be written as a product of matrices of the above form. In class, we showed that if the change-of-coordinates theorem holds for two composable functions, then it also holds for their composite. By induction, we conclude the result for all matrices. (ECP)

Problem 3.3. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function. Use Fubini's theorem to show that if $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y \partial x}$ are both continuous, then they are equal. (Hint: If their difference is nonzero at any point, then by continuity it is nonzero on an open rectangle containing that point.)

Solution. As in the hint, consider $D = \frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x}$. Suppose that there is some a for which $D(a) \neq 0$, and without loss of generality suppose furthermore that $D(a) > 0$. By continuous differentiability, then there is an open rectangle $U = [r, s] \times [t, u]$ containing a on which D is positive, from which it follows that $\int_U D > 0$ is positive as well.

We can use Fubini's theorem to make a separate calculation of $\int_U D$ as follows.

$$\begin{aligned} \int_U D &= \int_U \left(\frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} \right) = \int_U \left(\frac{\partial^2 f}{\partial x \partial y} \right) - \int_U \left(\frac{\partial^2 f}{\partial y \partial x} \right) \\ &= \int_t^u \int_r^s \left(\frac{\partial^2 f}{\partial x \partial y} \right) \mathrm{d}x\, \mathrm{d}y - \int_r^s \int_t^u \left(\frac{\partial^2 f}{\partial y \partial x} \right) \mathrm{d}y\, \mathrm{d}x. \end{aligned}$$

Now apply the single-variable Fundamental Theorem of Calculus to each term, twice:

$$\begin{aligned} &= \int_t^u \left(\left.\frac{\partial f}{\partial y}\right|_{r,y} - \left.\frac{\partial f}{\partial y}\right|_{s,y} \right) \mathrm{d}y - \int_r^s \left(\left.\frac{\partial f}{\partial x}\right|_{x,t} - \left.\frac{\partial f}{\partial x}\right|_{x,u} \right) \mathrm{d}x \\ &= f(r,u) - f(r,t) - (f(s,u) - f(t,u)) - (f(s,t) - f(r,t) - (f(s,u) - f(r,u))) = 0. \end{aligned}$$

(ECP)

Problem 3.4. Let $f\colon [a,b]\times[c,d]\to\mathbb{R}$ be continuous, and suppose that $\frac{\partial f}{\partial y}$ exists and is continuous. Define $F(y) = \int_a^b f(x,y)\,\mathrm{d}x$. Using the classical fundamental theorem of calculus, prove the following identity, known as *Leibniz's law*:

$$F'(y) = \int_a^b \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x.$$

That is: differentiation and integration commute. (Hint: $F(y) = \int_a^b f(x,y)\,\mathrm{d}x = \int_a^b (\int_c^y \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}y + f(x,c))\,\mathrm{d}x$.)[2]

Solution. As instructed, we investigate the hint. The hint is stated for $F(y)$, but we are interested in $F'(y)$, so we modify it appropriately:

$$\begin{aligned} F'(y) &= \frac{\partial}{\partial y}\int_a^b f(x,y)\,\mathrm{d}x = \frac{\partial}{\partial y}\int_a^b \left(\int_c^y \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}y + f(x,c)\right)\,\mathrm{d}x \\ &= \frac{\partial}{\partial y}\left(\int_a^b\int_c^y \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}y\,\mathrm{d}x + \int_a^b f(x,c)\,\mathrm{d}x\right) \end{aligned}$$

Fubini's theorem lets us change the order of the integrals, and the second term is constant as y changes, so partial differentiation erases it:

$$= \frac{\partial}{\partial y}\int_c^y\int_a^b \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

Finally, we apply the Fundamental Theorem of Calculus once more to eliminate the differentiation-integration pair:

$$= \int_a^b \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x. \tag{ECP}$$

4 For submission to Rohil Prasad

Problem 4.1. 1. Suppose $f, g : [a,b]\times[c,d]\to\mathbb{R}$ are integrable functions on a common domain D, and suppose that f is independent of the second variable and that g is independent of the first. Conclude the formula

$$\int_c^d\int_a^b f\cdot g dxdy = \left(\int_a^b f(x,0)dx\right)\cdot\left(\int_c^d g(0,y)dy\right).$$

2. Use this to calculate

$$\int_0^1\int_0^1 x^{n-1}y^{n-1}dxdy = \frac{1}{n^2}.$$

Solution. 1. Since g is independent of x, it follows that $g(x,y) = g(0,y)$ for every $x\in[a,b]$. Similarly, $f(x,y) = f(x,0)$ for every $y\in[c,d]$.

Therefore, we can write the function $\int_a^b f\cdot g dx$ as $g(0,y)\cdot\int_a^b f(x,y)dx = g(0,y)\cdot\int_a^b f(x,0)dx$. Integrating this with respect to y over $[c,d]$, we obtain the formula

$$\int_c^d g(0,y)\cdot\left(\int_a^b f(x,0)dx\right)dy = \left(\int_a^b f(x,0)dx\right)\cdot\left(\int_c^d g(0,y)dy\right).$$

2. This is equal to

$$\left(\int_0^1 x^{n-1}dx\right)\cdot\left(\int_0^1 y^{n-1}dy\right) = \frac{1}{n}\cdot\frac{1}{n} = \frac{1}{n^2}. \tag{RP}$$

[2] You'll discover that you can get away with much less than continuity of $\partial f/\partial y$.

Problem 4.2. Sum the above identity over n, then use Problem 1.2 and Problem 1.3 to conclude that

$$\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2} = \int_0^1 \int_0^1 \frac{1}{1-xy} dxdy.$$

Solution. By definition,

$$\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2} = \sum_{n=1}^{\infty} \int_0^1 \int_0^1 x^{n-1}y^{n-1} dxdy.$$

Our goal is now to show

$$\sum_{n=1}^{\infty} \int_0^1 \int_0^1 x^{n-1}y^{n-1} dxdy = \int_0^1 \int_0^1 \sum_{n=1}^{\infty} x^{n-1}y^{n-1} dxdy = \int_0^1 \int_0^1 \frac{1}{1-xy} dxdy.$$

The combined results of $1.2, 1.3$ generalize quite easily to integration over a rectangle in $\mathbb{R}^n$.
The functions $f_n = \sum_{k=1}^n x^{k-1}y^{k-1}$ are increasing, integrable on $[0,1]^2$ and converge pointwise to $\frac{1}{1-xy}$.
If we can show that $\frac{1}{1-xy}$ is improperly integrable, then by Problem 1.2 and Problem 1.3

$$\lim_{n\to\infty} \int_{[0,1]^2} \sum_{k=1}^{n} x^{k-1}y^{k-1} = \int_{[0,1]^2} \frac{1}{1-xy}$$

as desired.

This is clear given the computations in the next problems. (RP)

Problem 4.3. There are several ways to compute this integral.

1. Demonstrate

$$\int_0^1 \int_0^1 \left(\frac{1}{1-xy} - \frac{1}{1+xy} \right) dxdy = \frac{1}{2}\zeta(2)$$

by using the alternative coordinate system

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x^2 \\ y^2 \end{pmatrix}.$$

2. Demonstrate also

$$\int_0^1 \int_0^1 \left(\frac{1}{1-xy} + \frac{1}{1+xy} \right) dxdy = 2\int_0^1 \int_0^1 \frac{1}{1-x^2y^2} dxdy.$$

(Note the sign change.)

3. Conclude

$$\zeta(2) = \frac{4}{3} \int_0^1 \int_0^1 \frac{1}{1-x^2y^2} dxdy.$$

Solution. 1. We perform change of coordinates using the map $\varphi : (u,v) \mapsto (\sqrt{u}, \sqrt{v}) = (x,y)$.
The Jacobian is

$$\begin{pmatrix} \frac{1}{2\sqrt{u}} & 0 \\ 0 & \frac{1}{2\sqrt{v}} \end{pmatrix}$$

with Jacobian determinant $\frac{1}{4\sqrt{uv}}$.

By change of coordinates, we get

$$\int_0^1 \int_0^1 \frac{2xy}{1-x^2y^2} dxdy = \int_0^1 \int_0^1 \frac{2\sqrt{uv}}{1-uv} \cdot \frac{1}{4\sqrt{uv}} dudv = \frac{1}{2} \int_0^1 \int_0^1 \frac{1}{1-uv} dudv.$$

This is exactly equal to $\frac{1}{2}\zeta(2)$.

2. This follows by adding up the integrand:

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

3. Add up the integrals in the previous two parts to get

$$\int_0^1\int_0^1\frac{2}{1-xy}=2\int_0^1\int_0^1\frac{1}{1-x^2y^2}dxdy+\frac{1}{2}\zeta(2).$$

The left-hand side is equal to $2\zeta(2)$, so after simplification we can conclude

$$\zeta(2)=\frac{4}{3}\int_0^1\int_0^1\frac{1}{1-x^2y^2}dxdy. \tag{RP}$$

Problem 4.4. Finally, use the second coordinate change

$$\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}\sin r/\cos s\\ \sin s/\cos r\end{pmatrix}$$

to compute the last integral in the preceding step.

1. Calculate the image of the unit square in xy-coordinates through the coordinate transform to the triangle with vertices at $(0,0)$, $(\pi/2,0)$, and $(0,\pi/2)$ in rs-coordinates.
2. Calculate the Jacobian of the transformation.
3. Use the change of coordinates theorem, and evaluate the resulting integral.

Solution. Let the change of coordinates map be $\varphi:(r,s)\mapsto(\sin r/\cos s,\sin s/\cos r)=(x,y)$ from $\mathbb{R}^2\to\mathbb{R}^2$.

1. We will show $\varphi^{-1}([0,1]^2)$ is the set of points (r,s) such that $r+s\le\pi/2$ and $r,s\ge 0$, which is our desired triangle.

First, pick (r,s) such that $r+s\le\pi/2$, $r,s\ge 0$. Then $0\le s\le\pi/2-r$. Since the cosine function is decreasing on $[0,\pi/2]$, it follows that $\cos s\le\cos(\pi/2-r)=\sin r$, so $\sin r/\cos s\in[0,1]$.

By a symmetric argument, $\sin s/\cos r\in[0,1]$ and so $\varphi(r,s)\in[0,1]^2$.

Next, we show that the map is surjective. Pick $(x,y)\in[0,1]$. Then, we have $\sin r=x\cos s$, $\sin s=y\cos r$. It follows that $r=\arcsin(x\cos s)$.

We evaluate $\cos r=\sqrt{1-(x\cos s)^2}$, so it follows that $\sin s=y\sqrt{1-(x\cos s)^2}$. Squaring this, we get $1-\cos^2 s=y^2-y^2x^2\cos^2 s$, so $(y^2x^2-1)\cos^2 s=y^2-1$.

From here, we obtain $s=\arccos(\sqrt{\frac{y^2-1}{y^2x^2-1}})$ and $r=\arcsin(\sqrt{\frac{x(y^2-1)}{y^2x^2-1}})$. The skeptical reader can check that this is indeed in the triangle.

2. The Jacobian of the transformation is

$$\begin{pmatrix}\cos r/\cos s & \frac{\sin r\sin s}{(\cos s)^2}\\ \frac{\sin r\sin s}{(\cos r)^2} & \cos s/\cos r\end{pmatrix}.$$

We can then calculate the Jacobian determinant to be

$$1-\frac{\sin^2 r\sin^2 s}{\cos^2 r\cos^2 s}.$$

3. Finally, we can apply change of coordinates. Pulling back by φ gives us

$$\begin{aligned}
\frac{4}{3}\int_0^1\int_0^1 \frac{1}{1-x^2y^2}dxdy &= \frac{4}{3}\int_T \frac{1}{1-\frac{\sin^2 r \sin^2 s}{\cos^2 r\cos^2 s}} \cdot (1 - \frac{\sin^2 r \sin^2 s}{\cos^2 r \cos^2 s}) \\
&= \frac{4}{3}\int_T 1 \\
&= \frac{4}{3}\cdot \pi^2/8 \\
&= \frac{\pi^2}{6}. \qquad \text{(RP)}
\end{aligned}$$