Miscellaneous UC Berkeley Prelim Style Problems

A Few Opening Remarks: In preparation for my preliminary exam as a first year PhD student at UC Berkeley, I solved several hundred problems from old preliminary exams and other sources with similar problems. Below are my solutions to a handful of the problems I solved which I deemed, for whatever reason, important or interesting enough to write into a LaTeX document. These problems are well designed and should be of interest even to undergraduates who are not yet pursuing doctoral studies. If you are a first year PhD student at UC Berkeley preparing for this exam, you should probably make your own document of this form as preparation. While it is rare for the department to drop students for failing the prelim nowadays, I learned a lot from preparing this intensely, and I consider myself a better mathematician for it.

As a warning, there are probably some typos and mistakes here and there. Also, you might notice a heavy bias of problems in complex analysis and linear algebra. This is because I already felt good about my real analysis, so I only added real analysis problems to this list if I felt particularly strong about them. I am generally not a fan of abstract algebra, so there are fewer of those problems here as well.

Exercise 1. A continuous function $\phi : (a,b) \to \mathbb{R}$ is convex if and only if

$$\phi\left(\frac{x+y}{2}\right) \leq \frac{1}{2}\phi(x) + \frac{1}{2}\phi(y)$$

for all $x, y \in (a,b)$.

Proof. If ϕ is convex, then $\phi\left(\frac{x+y}{2}\right) \leq \frac{1}{2}\phi(x) + \frac{1}{2}\phi(y)$ is clear directly from the definition of convexity.

To prove the converse, we establish by induction on $m \geq 1$ that for any $x, y \in (a,b)$

$$\phi\left(\frac{n}{2^m}x + \left(1 - \frac{n}{2^m}\right)y\right) \leq \frac{n}{2^m}\phi(x) + \left(1 - \frac{n}{2^m}\right)\phi(y) \quad \text{for all } 0 \leq n \leq 2^m.$$

The $m = 1$ case is true by assumption. If the $m \geq 1$ case is true, then for $0 \leq n \leq 2^m$ one has

$$\begin{aligned}
\phi\left(\frac{\frac{n}{2^m}x + \left(1 - \frac{n}{2^m}\right)y + y}{2}\right) &\leq \frac{1}{2}\phi\left(\frac{n}{2^m}x + \left(1 - \frac{n}{2^m}\right)y\right) + \frac{1}{2}\phi(y) \\
&\leq \frac{1}{2}\left(\frac{n}{2^m}\phi(x) + \left(1 - \frac{n}{2^m}\right)\phi(y)\right) + \frac{1}{2}\phi(y) \\
&= \frac{n}{2^{m+1}}\phi(x) + \left(1 - \frac{n}{2^{m+1}}\right)\phi(y).
\end{aligned}$$

But $\phi\left(\frac{\frac{n}{2^m}x + \left(1 - \frac{n}{2^m}\right)y + y}{2}\right) = \phi\left(\frac{n}{2^{m+1}}x + \left(1 - \frac{n}{2^{m+1}}\right)y\right)$, hence

$$\phi\left(\frac{n}{2^{m+1}}x + \left(1 - \frac{n}{2^{m+1}}\right)y\right) \leq \frac{n}{2^{m+1}}\phi(x) + \left(1 - \frac{n}{2^{m+1}}\right)\phi(y)$$

for $0 \leq n \leq 2^m$. If $2^m < n \leq 2^{m+1}$ then $0 < n - 2^m \leq 2^m$, hence

$$\begin{aligned}
\phi\left(\frac{\frac{n-2^m}{2^m}x + \left(1 - \frac{n-2^m}{2^m}\right)y + x}{2}\right) &\leq \frac{1}{2}\phi\left(\frac{n-2^m}{2^m}x + \left(1 - \frac{n-2^m}{2^m}\right)y\right) + \frac{1}{2}\phi(x) \\
&\leq \frac{1}{2}\left(\frac{n-2^m}{2^m}\phi(x) + \left(1 - \frac{n-2^m}{2^m}\right)\phi(y)\right) + \frac{1}{2}\phi(x) \\
&= \frac{n}{2^{m+1}}\phi(x) + \left(1 - \frac{n}{2^{m+1}}\right)\phi(y)
\end{aligned}$$

But $\phi\left(\frac{\frac{n-2^m}{2^m}x + \left(1 - \frac{n-2^m}{2^m}\right)y + x}{2}\right) = \phi\left(\frac{n}{2^{m+1}}x + \left(1 - \frac{n}{2^{m+1}}\right)y\right)$, so the $m+1$ case is proved.

Denote by $\mathcal{D} = \{n2^{-m} : m \geq 1,\, 0 \leq n \leq 2^m\}$. We have just shown that if $t \in \mathcal{D}$, then the convexity condition $\phi(tx+(1-t)y) \leq t\phi(x)+(1-t)\phi(y)$ holds. Therefore, fix $t \in [0,1]$, and define the sequence $\{t_n\}$ by $t_n = \lfloor 2^n t \rfloor / 2^n$. Evidently, $t_n \in \mathcal{D}$ for each $n \geq 1$, and

$$|t - t_n| \leq \frac{1}{2^n},$$

which means $t_n \to t$ as $n \to \infty$. Then by continuity,

$$\phi(tx+(1-t)y) = \lim_{n\to\infty} \phi(t_n x + (1-t_n)y) \leq \lim_{n\to\infty} t_n\phi(x) + (1-t_n)\phi(y) = t\phi(x) + (1-t)\phi(y).$$

Since $t \in [0,1]$ and $x, y \in (a,b)$ were arbitrary, this proves that ϕ is convex. □

The same exact strategy employed here is enough to establish the following corollary:

Exercise 2 (Sard's Theorem in one dimension). Suppose that $f \in C^1(\mathbb{R})$, and let $E = \{x : f'(x) = 0\}$. Then $f(E)$ is Lebesgue measurable and $m(f(E)) = 0$.

Proof. Fix a compact set K and $\epsilon > 0$, and by the uniform continuity of f' on K fix a $\delta > 0$ small enough that

$$\left| \frac{f(x) - f(y)}{x - y} \right| < \epsilon$$

whenever x and y are in a δ-neighborhood of $E \cap K$. Furnish a sequence of intervals $\{(a_k, b_k)\}$ each of length at most δ covering $E \cap K$ with

$$\sum_{k=1}^{\infty} b_k - a_k < m(E \cap K) + \epsilon.$$

By discarding intervals which do not intersect $E \cap K$, we assume that every interval contains a point of $E \cap K$. By the intermediate value theorem we then have

$$m(f((a_k, b_k))) = \max_{x,y \in [a_k, b_k]} |f(y) - f(x)|.$$

But each $[a_k, b_k]$ contains a point of E, and since the length of $[a_k, b_k]$ is at most δ,

$$\max_{x,y\in[a_k,b_k]} |f(y) - f(x)| \leq \epsilon \max_{x,y\in[a_k,b_k]} \epsilon|y - x| = \epsilon(b_k - a_k).$$

Therefore

$$m\left(f(E \cap K)\right) \leq \sum_{k=1}^{\infty} m(f((a_k, b_k))) \leq \sum_{k=1}^{\infty} \epsilon(b_k - a_k) < \epsilon(m(E \cap K) + \epsilon).$$

Since $m(E \cap K) < \infty$ and ϵ was arbitrary, this shows $m\left(f(E \cap K)\right) = 0$. Since K was an arbitrary compact set, setting $K = [-n, n]$ for $n \geq 1$ and then taking $n \to \infty$ gives $m(f(E)) = 0$, as desired. □

Exercise 3. Suppose that $x_1, \ldots, x_n \geq 0$. Then

$$(1 + x_1) \cdots (1 + x_n) \geq \left(1 + (x_1 \cdots x_n)^{\frac{1}{n}}\right)^n,$$

with equality if and only if $x_1 = \cdots = x_n$.

Proof. By the inequality of arithmetic and geometric means, for each $k \geq 1$,

$$\sum_{1 \leq i_1 < \cdots < i_k \leq n} x_{i_1} \cdots x_{i_k} \geq \binom{n}{k} \left(\prod_{1 \leq i_1 < \cdots < i_k \leq n} x_{i_1} \cdots x_{i_k} \right)^{\frac{1}{\binom{n}{k}}}.$$

Next, we observe that

$$\prod_{1\le i_1<\cdots<i_k\le n} x_{i_1}\cdots x_{i_k} = (x_1\cdots x_n)^{\binom{n}{k}-\binom{n-1}{k}},$$

since each x_j appears in exactly $\binom{n}{k}-\binom{n-1}{k}$ many terms. Therefore, after making the obvious simplifications

$$\binom{n}{k}\left(\prod_{1\le i_1<\cdots<i_k\le n} x_{i_1}\cdots x_{i_k}\right)^{\frac{1}{\binom{n}{k}}} = \binom{n}{k}(x_1\cdots x_n)^{\frac{k}{n}}.$$

This yields

$$(1+x_1)\cdots(1+x_n) = 1+\sum_{k=1}^{n}\sum_{1\le i_1<\cdots<i_k\le n} x_{i_1}\cdots x_{i_k} \ge 1+\sum_{k=1}^{n}\binom{n}{k}(x_1\cdots x_n)^{\frac{k}{n}} = \left(1+(x_1\cdots x_n)^{\frac{1}{n}}\right)^n,$$

and our usage of the AM-GM inequality implies that equality holds if and only if $x_1=\cdots=x_n$. $\square$

Exercise 4. Suppose that W, X, and Y are vector spaces, and $\alpha : W\to X$ and $\beta : W\to Y$ are linear functions. If $\ker\alpha\subset\ker\beta$, then there is a linear map $f:\alpha(W)\to Y$ such that $\beta=f\circ\alpha$.

Proof. Define f by $f(y)=\beta(x)$ if $\alpha(x)=y$. The map f is well-defined since $\alpha(x)=\alpha(x')=y$ means $\alpha(x-x')=0$, hence $\beta(x-x')=0$, so $f(y)=\beta(x)=\beta(x')=f(y)$. Clearly f is linear and $\beta=f\circ\alpha$, so we are done. $\square$

Corollary 5. *If $T : X\to\mathbb{C}$ and $\Lambda_1,\dots,\Lambda_n : X\to\mathbb{C}$ are linear maps, and $T(x)=0$ for every x such that $\Lambda_1(x)=\cdots=\Lambda_n(x)=0$, then there exist constants $\alpha_1,\dots,\alpha_n$ such that $T=\alpha_1\Lambda_1+\cdots+\alpha_n\Lambda_n$.*

Proof. If $U:X\to\mathbb{C}^n$ is defined by $U=(\Lambda_1,\dots,\Lambda_n)$, then $\ker U\subset\ker T$, hence $T=f\circ U$ for some linear map $f:\mathbb{C}^n\to\mathbb{C}$. As $\mathbb{C}^n$ is finite dimensional, $f(x_1,\dots,x_n)=\alpha_1x_1+\cdots\alpha_nx_n$ for some $\alpha_1,\dots,\alpha_n$, hence $T=\alpha_1\Lambda_1+\cdots+\alpha_n\Lambda_n$, as desired. $\square$

Exercise 6 (*Berkeley Problems in Mathematics*, Exc. 1.8.3). Suppose that $f:\mathbb{R}\to\mathbb{R}$ is a continuous function satisfying

$$f(x)\le\frac{1}{2h}\int_{x-h}^{x+h} f(t)\,dt$$

for all $x\in\mathbb{R}$ and $h>0$. Then f is convex.

Proof. First assume that $f\in C^2$. Suppose that there exists $x_0\in\mathbb{R}$ such that $f''(x_0)<0$. Put $\eta=-\frac{1}{2}f''(x_0)>0$, and fix $\delta>0$ small enough that $f''(t)\le-\eta$ whenever $t\in[x_0-\delta,x_0+\delta]$. For such t, by Taylor's theorem we have

$$f(t)\le f(x_0)+f'(x_0)(t-x_0)-\frac{\eta}{2}(t-x_0)^2.$$

Integrating this inequality over $[x_0-\delta,x_0+\delta]$ yields

$$\frac{1}{2\delta}\int_{x_0-\delta}^{x_0+\delta} f(t)\,dt\le f(x_0)-\frac{\eta}{2}\frac{\delta^3}{3}<f(x_0),$$

which contradicts the assumptions on f. Therefore, $f''(x)\ge0$ for all $x\in\mathbb{R}$, which means f is convex.

Now suppose that f is an arbitrary continuous function. Let ϕ be a non-negative C^2 function supported in $[-1,1]$ with $\int\phi=1$. For $\epsilon>0$, let $\phi_\epsilon(x)=\epsilon^{-1}\phi(\epsilon^{-1}x)$ and $f_\epsilon=f*\phi_\epsilon$. For $h>0$, Fubini's theorem shows

$$\begin{aligned}\frac{1}{2h}\int_{x-h}^{x+h} f_\epsilon(t)\,dt &= \int \phi_\epsilon(y)\left[\frac{1}{2h}\int_{x-h}^{x+h} f(t-y)\,dt\right] dy\\ &= \int \phi_\epsilon(y)\left[\frac{1}{2h}\int_{x-y-h}^{x-y+h} f(t)\,dt\right] dy\\ &\geq \int \phi_\epsilon(y) f(x-y)\,dy\\ &= f_\epsilon(x)\end{aligned}$$

whenever $x \in \mathbb{R}$. Since $f_\epsilon \in C^2$, our previous work shows that f_ϵ is convex for all $\epsilon > 0$. Therefore,

$$\lim_{\epsilon \to 0} f_\epsilon = f$$

is convex, since pointwise limits of convex functions are convex. □

Exercise 7 (***Berkeley Problems in Mathematics*, Exc. 4.1.18**). Let $X \subset \mathbb{R}^n$ be a compact set and suppose that $f : X \to \mathbb{R}$ is a continuous function. Then for every $\epsilon > 0$, there is $M > 0$ such that

$$|f(x) - f(y)| \leq M|x-y| + \epsilon$$

for every $x, y \in X$.

Proof. Let $D = \operatorname{diam} f(X) < \infty$, and pick $\delta > 0$ small enough that $|f(x) - f(y)| < \epsilon$ if $|x-y| < \delta$. We choose $M > 0$ so large that $D \leq M\delta + \epsilon$. For this choice of M, if $|x-y| < \delta$, we have

$$|f(x) - f(y)| < \epsilon \leq M|x-y| + \epsilon,$$

and if $|x-y| \geq \delta$ we have

$$|f(x) - f(y)| \leq D \leq M\delta + \epsilon \leq M|x-y| + \epsilon.$$

Therefore, this choice of M satisfies the requirements of the problem. □

Exercise 8 (***Berkeley Problems in Mathematics* Exc. 1.8.2**). Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function such that $f(x) > 0$ for all $x \in \mathbb{R}$, and suppose that $e^{cx} f(x)$ is a convex function for every $c \in \mathbb{R}$. Then $\log f$ is a convex function.

We require the following simple lemma.

Lemma 9. *A continuous function $f : \mathbb{R} \to \mathbb{R}$ is convex if and only if*

$$\frac{f(x+h) + f(x-h)}{2} \geq f(x)$$

for all $x \in \mathbb{R}$ and $h \neq 0$.

Proof. If f is convex, then inequality holds by definition of convexity. If f is a C^2 function satisfying this inequality, then

$$\frac{f(x+h) + f(x-h) - 2f(x)}{2h^2} \geq 0,$$

and taking $h \to 0$ implies that $f''(x) \geq 0$ for all $x \in \mathbb{R}$. For the general case f is continuous case, one can reduce to the C^2 case using an approximate identity. □

Proof. By applying Lemma 9 to $\log f$, it suffices to show that

$$f(x-h)f(x+h) \geq f(x)^2$$

for all $x \in \mathbb{R}$ and $h \neq 0$. For a fixed $x \in \mathbb{R}$ and $h \neq 0$, define a quadratic polynomial p by

$$p(y) = y^2 f(x+h) - 2y f(x) + f(x-h).$$

Since $e^{cx}f(x)$ is convex, again by Lemma 9 it holds

$$e^{c(x+h)}f(x+h) - 2e^{cx}f(x) + e^{c(x-h)}f(x-h) \geq 0$$

for all $c \in \mathbb{R}$ and $x \in \mathbb{R}$ and $h \neq 0$. Letting $c = \frac{1}{h}\log y$ for $y > 0$ implies that

$$p(y) = y^2 f(x+h) - 2yf(x) + f(x-h) \geq 0$$

for $y > 0$. The vertex of p is at $y = f(x)/f(x+h) > 0$. Since the leading coefficient of p is positive, p attains a minimum value at this point, hence $p(y) \geq 0$ for all $y \in \mathbb{R}$. Therefore, the discriminant of p is negative, which is to say that $f(x)^2 \leq f(x-h)f(x+h)$, as desired. □

Exercise 10 (*Crux Mathematicorum* June 2024 Problem 4959). For $\alpha > 0$, evaluate the limit

$$\lim_{n\to\infty} \sum_{k=1}^{2n} (-1)^k \left(\frac{k}{2n}\right)^\alpha$$

Here is a cheap way to do this using the Stolz-Cesáro theorem.

Proof. We first show that

$$\lim_{n\to\infty} \frac{(2n+2)^\alpha - (2n)^\alpha}{(2n+2)^\alpha - (2n+1)^\alpha} = 2.$$

By the Mean-Value Theorem, there is $\xi_n \in (2n, 2n+1)$ such that $(2n+1)^\alpha - (2n)^\alpha = \alpha\xi_n^{\alpha-1}$. Therefore,

$$\frac{(2n+2)^\alpha - (2n)^\alpha}{(2n+2)^\alpha - (2n+1)^\alpha} = 1 + \left(\frac{\xi_n}{\xi_{n+1}}\right)^{\alpha-1}.$$

Since

$$\frac{2n}{2n+3} \leq \frac{\xi_n}{\xi_{n+1}} \leq \frac{2n+1}{2n+2},$$

as $n \to \infty$ we get

$$\lim_{n\to\infty} \frac{(2n+2)^\alpha - (2n)^\alpha}{(2n+2)^\alpha - (2n+1)^\alpha} = 1 + \left(\lim_{n\to\infty} \frac{\xi_n}{\xi_{n+1}}\right)^{\alpha-1} = 2$$

by the Squeeze Theorem. Therefore, by the Stolz-Cesaro Theorem, we have

$$\lim_{n\to\infty} \frac{(2n)^\alpha}{\sum_{k=1}^{2n}(-1)^k k^\alpha} = \lim_{n\to\infty}\lim_{n\to\infty} \frac{(2n+2)^\alpha - (2n)^\alpha}{(2n+2)^\alpha - (2n+1)^\alpha} = 2,$$

hence

$$\lim_{n\to\infty} \sum_{k=1}^{2n} (-1)^k \left(\frac{k}{2n}\right)^\alpha = \frac{1}{2},$$

as desired. □

But here is what is likely the intended way.

Proof. By summing over the even and odd indices, we obtain

$$\begin{aligned}\sum_{k=1}^{2n} (-1)^k \left(\frac{k}{2n}\right)^\alpha &= \sum_{k=1}^{n} \left(\frac{2k}{2n}\right)^\alpha - \sum_{k=1}^{n} \left(\frac{2k-1}{2n}\right)^\alpha \\ &= \frac{1}{n^\alpha} \sum_{k=1}^{n} k^\alpha - \left(k - \frac{1}{2}\right)^\alpha.\end{aligned}$$

By the Mean-Value Theorem, there exists $\xi_k \in (k-1/2, k)$ such that $k^\alpha - \left(k - \frac{1}{2}\right)^\alpha = \frac{1}{2}\alpha\xi_k^{\alpha-1}$. So,

$$\frac{1}{n^\alpha}\sum_{k=1}^n k^\alpha - \left(k - \frac{1}{2}\right)^\alpha = \frac{\alpha}{2n^\alpha}\sum_{k=1}^n \xi_k^{\alpha-1}.$$

Since

$$\sum_{k=1}^n (k-1)^{\alpha-1} \le \sum_{k=1}^n \xi_k^{\alpha-1} \le \sum_{k=1}^n k^{\alpha-1}$$

and

$$\lim_{n\to\infty}\frac{1}{n^\alpha}\sum_{k=1}^n (k-1)^{\alpha-1} = \lim_{n\to\infty}\frac{1}{n^\alpha}\sum_{k=1}^n k^{\alpha-1} = \frac{1}{\alpha},$$

we deduce

$$\lim_{n\to\infty}\sum_{k=1}^{2n}(-1)^k\left(\frac{k}{2n}\right)^\alpha = \frac{1}{2}$$

by the Squeeze Theorem. □

EXERCISE 11 (*CRUX MATHEMATICORUM* JUNE 2024 PROBLEM 4951). Let n be a positive integer. Prove that the sums

$$\sum_{k=1}^n \frac{(-1)^{k-1}}{k}\binom{n}{k}^{-1} \quad\text{and}\quad \sum_{k=1}^n \frac{(-1)^{k-1}}{k}\binom{n}{k-1}$$

are equal and find their common value.

The common value of the sum is $\frac{1+(-1)^{n+1}}{n+1}$.

PROOF. From the identity

$$\binom{n}{k-1} = \frac{k}{n+1}\binom{n+1}{k},$$

we deduce

$$\sum_{k=1}^n \frac{(-1)^{k-1}}{k}\binom{n}{k-1} = \frac{1}{n+1}\sum_{k=1}^n (-1)^{k-1}\binom{n+1}{k}.$$

From the Binomial Theorem,

$$0 = \sum_{k=0}^{n+1}(-1)^k\binom{n+1}{k} = 1 + (-1)^{n+1} - \sum_{k=1}^n (-1)^{k-1}\binom{n+1}{k},$$

from which we deduce

$$\sum_{k=1}^n \frac{(-1)^{k-1}}{k}\binom{n}{k-1} = \frac{1}{n+1}\sum_{k=1}^n (-1)^{k-1}\binom{n+1}{k} = \frac{1+(-1)^{n+1}}{n+1}.$$

Next, let

$$S_n := \sum_{k=1}^n \frac{(-1)^{k-1}}{k\binom{n}{k}}.$$

Using $k\binom{n}{k} = n\binom{n-1}{k-1}$, we deduce

$$S_n = \frac{1}{n}\sum_{k=1}^n \frac{(-1)^{k-1}}{\binom{n-1}{k-1}}.$$

Again using $\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}$ and $\binom{n}{k-1} = \frac{n}{n+1-k}\binom{n-1}{k-1}$ gives

$$S_n = \frac{1}{n+1}\sum_{k=1}^{n}(-1)^{k-1}\left(\frac{1}{\binom{n}{k-1}} + \frac{1}{\binom{n}{k}}\right).$$

But,

$$\frac{1}{n+1}\sum_{k=1}^{n}(-1)^{k-1}\frac{1}{\binom{n}{k-1}} = \frac{(-1)^{n+1}}{n+1} + S_{n+1}$$

and

$$\frac{1}{n+1}\sum_{k=1}^{n}(-1)^{k-1}\frac{1}{\binom{n}{k}} = \frac{1}{n+1} - S_{n+1},$$

which gives the conclusion:

$$S_n = \frac{1+(-1)^{n+1}}{n+1}.$$

□

Exercise 12. Let (X,d) be a compact metric space and $f : X \to X$ be an isometry of X. Then $f(X) = X$.

Proof. Define a sequence of functions $\{g_n\}$ recursively by $g_1 = f$ and $g_{n+1} = f \circ g_n$. Since f is an isometry, each $\{g_n\}$ is an isometry, so

$$d(g_n(x), g_n(y)) = d(x,y)$$

for all $x, y \in X$. Therefore, the sequence $\{g_n\}$ is uniformly bounded and equicontinuous. By the Arzela-Ascoli Theorem, there is a uniformly convergent subsequence $\{g_{n_j}\}$. Then for all $x \in X$,

$$0 = \lim_{j\to\infty} d(g_{n_j}(x), g_{n_{j+1}}(x)) = \lim_{j\to\infty} d(x, g_{n_{j+1}-n_j}(x)),$$

where we used the fact that each g_n is an isometry and $g_{n+1} = f \circ g_n$. If $m_j = n_{j+1} - n_j$ it follows that $g_{m_j}(x) \to x$ as $j \to \infty$. Fix $x \in X$. By passing to a subsequence if necessary, we can suppose that $\{g_{m_j-1}(x)\}$ converges to some $y \in X$ as $j \to \infty$. Since $g_{m_j}(x) = f(g_{m_j-1}(x))$, taking $j \to \infty$ gives $x = f(y)$, which proves $X \subset f(X)$. □

Corollary 13. *If X is a compact metric space and f is an isometry of X, then there is a sequence $\{n_k\}$ such that $f^{n_k}(x) \to x$ uniformly as $k \to \infty$, where f^{n_k} is the n_k-fold composition of f with itself.*

Exercise 14. Prove the Arzela-Ascoli Theorem: If X is a compact metric space and $\mathcal{F}$ is a uniformly bounded and equicontinuous family of functions on X, then there is a uniformly convergent subsequence in $\mathcal{F}$.

Proof. Let $\{x_n\}$ be a countable dense subset of X. Since $\mathcal{F}$ is uniformly bounded, the set $\{f(x_1) : f \in \mathcal{F}\}$ is pre-compact. In particular, there is a sequence of functions $\{f_{n_j(1)}\} \subset \mathcal{F}$ such that $\{f_{n_j(1)}(x_1) : j \geq 1\}$ converges to some point, say $g(x_1)$ The set $\{f_{n_j(1)}(x_2) : j \geq 1\}$ is also pre-compact, so there is a subsequence $\{f_{n_j(2)}(x_2) : j \geq 1\}$ which converges, say to $g(x_2)$. We continue inductively in this fashion, for each $k \geq 1$ obtaining a subsequence $\{f_{n_j(k+1)}\} \subset \{f_{n_j(k)}\}$ such that $\{f_{n_j(k)}(x_k)\}$ converges to some $g(x_k)$ as $j \to \infty$. If $m_k = n_k(k)$, it follows that $f_{m_k}(x_j) \to g(x_j)$ for every $j \geq 1$. We show now that the function g defined on the points $\{x_n\}$ is continuous. Indeed, since $\mathcal{F}$ is equicontinuous, if $\epsilon > 0$ there is $\delta > 0$ such that

$$|f_{m_k}(x_m) - f_{m_k}(x_n)| \leq \epsilon$$

for all $k \geq 1$ if $d(x_m, x_n) < \delta$. Taking $k \to \infty$ yields $|g(x_m) - g(x_n)| \leq \epsilon$, so g is uniformly continuous on $\{x_n\}$. Since $\{x_n\}$ is dense in X, it follows that g can be extended uniquely to a uniformly continuous function on X. Thus, for a general $x \in X$, the number $g(x)$ is well-defined. To see that $f_{m_k}(x) \to g(x)$ for each $x \in X$, fix $\epsilon > 0$ and $\delta > 0$ as before. If $x \in X$, let $x_n \in X$ be chosen so that $d(x_n, x) < \delta$. For this n, for all sufficiently large j we have $|f_{m_j}(x_n) - g(x_n)| \leq \epsilon$. Then

$$|f_{m_j}(x) - g(x)| \leq |f_{m_j}(x) - f_{m_j}(x_n)| + |g(x_n) - g(x)| + |f_{m_j}(x_n) - g(x_n)| \leq 3\epsilon.$$

Since $\epsilon > 0$ was arbitrary, it follows $f_{m_k} \to g$ on X. To see that the convergence is uniform, we just appeal to the following lemma:

Lemma 15. *If $\{f_n\}$ is an equicontinuous sequence of functions converging to pointwise some f_∞ on a compact metric space X, then the convergence is uniform.*

Proof. Fix $\epsilon > 0$ and let $\delta > 0$ be chosen so that $|f_n(x) - f_n(y)| \leq \epsilon$ if $d(x, y) < \delta$ for all $1 \leq n \leq \infty$. For each $x \in X$, let N_x be an integer so large that $|f_n(x) - f_\infty(x)| \leq \epsilon$ if $n \geq N_x$. Since X is compact, we can cover X by finitely many balls of radius δ. Let $x_1, \ldots, x_m$ be the centers of these balls and let $N = \max\{N_{x_1}, \ldots, N_{x_m}\}$. For a fixed $x \in X$, choose x_j so that x is in the δ-ball about x_j. Then for $n \geq N$, our choice of δ implies

$$|f_n(x) - f_\infty(x)| \leq |f_n(x) - f_n(x_j)| + |f_\infty(x) - f_\infty(x_j)| + |f_n(x_j) - f_\infty(x_j)| \leq 3\epsilon,$$

and since $\epsilon > 0$ was arbitrary and N does not depend on x, this shows $f_n \to f_\infty$ uniformly on X. □

□

Exercise 16. Suppose that $f : \mathbb{C} \to \mathbb{C}$ is an entire function such that for every $z_0 \in \mathbb{C}$, if $f(z) = \sum c_n(z - z_0)^n$, there is some n such that $c_n = 0$. Then f is a polynomial.

Proof. Baire Category. □

Exercise 17 (General Maximum Modulus). Suppose that $\Omega \subset \mathbb{C}$ is an open set such that $\partial\Omega$ is compact, and that $f : \overline{\Omega} \to \mathbb{C}$ is a continuous function holomorphic in Ω. Then $|f|$ achieves its extremal values in $\partial\Omega$. If Ω is connected and there is a point $z_0 \in \Omega$ such that $|f(z_0)| = \max_z |f(z)|$, then f is identically a constant.

Proof. To each $\epsilon > 0$, let $\Omega_n = \{z \in \Omega : d(x, \partial\Omega) > n^{-1}\}$. Then $\Omega_{n+1} \subset \Omega_n \subset \Omega$ is open for each n, and by the ordinary maximum modulus principle, $|f|$ attains a maximum value in $\partial\Omega_\epsilon = \{z \in \Omega : d(x, \partial\Omega) = n^{-1}\}$. Let $z_n \in \partial\Omega_n$ be the value where this maximum is attained. By the maximum modulus principle, if $n > m$, then $|f(z_m)| < |f(z_n)|$, otherwise $|f(z_n)| = |f(z_m)|$ and this implies that f is constant. Now, since $\{z \in \Omega : d(x, \partial\Omega) \leq 1\}$ is compact, we can choose a convergent subsequence $\{z_{n_j}\}$. Let z_0 be the sub-sequential limit. Then $d(z_{n_j}, \partial\Omega) \leq n_j^{-1}$ implies that $d(z_0, \partial\Omega) = 0$, so $z_0 \in \partial\Omega$. Moreover, it holds $|f(z_0)| > |f(z)|$ for all $z \in \Omega$ by definition of z_0, and this implies that f attains its maximal value in $\partial\Omega$.

Finally, to obtain a contradiction, suppose there is $z \in \Omega$ that attains the maximum value of $|f|$. Then there is n large enough that $z \in \Omega_n$, and this means that $|f(z_n)| = |f(z)|$, which means f is constant on Ω_n, hence is constant on Ω since Ω is connected. □

Exercise 18. Let A be the disk algebra, i.e., the continuous functions $f : \overline{\mathbb{D}} \to \mathbb{C}$ such that f is holomorphic in $\mathbb{D}$ equipped with $\|\cdot\|_\infty$. If $f_1, \ldots, f_n \in A$ do not share any common zeros, then there are $g_1, \ldots, g_n \in A$ such that $\sum_{j=1}^n g_j f_j = 1$.

Proof. Let

$$I = \left\{ \sum_{j=1}^n g_j f_j : g_1, \ldots, g_n \in A \right\}.$$

Clearly, I is an ideal. If $I \neq A$, then there is a maximal ideal M (possibly I itself) such that $I \subseteq M$. Since the maximal ideals of A are precisely the kernels of evaluation maps, there is $|z_0| \leq 1$ with $f(z_0) = 0$ for all $f \in I$. However, since $f_1, \ldots, f_n \in I$, this means that $f_1(x_0) = \cdots = f_n(x_0) = 0$, contradicting the fact that $\{f_j\}$ share no common zeros. The contradiction is resolved if $I = A$, which is equivalent to the stated exercise. $\square$

Exercise 19. Let $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ be the n-th harmonic number. Show that H_n is never an integer for $n \geq 2$.

Proof. For $n \geq 2$, observe that

$$H_n = \frac{P_n}{n!},$$

where

$$P_n = \sum_{j=1}^{n} \prod_{i=1, i\neq j}^{n} i.$$

So, for any $2 \leq k \leq n$, it holds

$$P_n = \frac{n!}{k} \mod k.$$

Let $p(n)$ be the largest prime less than n. If $p(n) < n$, then for any $p(n) < k \leq n$, all the prime factors of k are strictly smaller than $p(n)$. Otherwise, if there is a prime divisor q of k with $q > p(n)$, then $p(n) < 2p(n) < 2q \leq k \leq n$. By Bertrand's postulate, there is another prime p' with $p(n) < p' < 2p(n)$, contradicting the maximality of $p(n)$. As a result, if $p(n) < n$, it holds

$$P_n = \frac{n!}{p(n)} \neq 0 \mod p(n),$$

since $p(n)$ does not divide the product $n(n-1)\cdots(p(n)+1)$ and it does not divide the product $(p(n)-1)!$. If $p(n) = n$, then

$$P_n = (p(n)-1)! \neq 0 \mod p(n).$$

Therefore, if P_n were divisible by $n!$, it would be divisible by $p(n)$, but we have shown that P_n is not divisible by $p(n)$. It follows that $H_n = P_n/n!$ is not an integer for $n \geq 2$. $\square$

Remark 20. Is there a way to do this without Bertrand's postulate?

Exercise 21. Let p be a polynomial with degree at least 2 and all real coefficients such that for some number a, it holds $p(a) \neq 0$ but $p'(a) = p''(a) = 0$. Then p has a nonreal root.

Lemma 22. *Suppose that $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ is a polynomial such that*

(1) $n \geq 2$,
(2) $a_0 \neq 0$,
(3) all the roots of p are real.

Then $a_1^2 > 2a_0a_2$.

Proof. We can assume without loss of generality that $a_n = 1$. The proof is by induction on $n \geq 2$. When $n = 2$, let r_1 and r_2 be the roots of p. Then $a_2 = 1$, $a_1 = -(r_1 + r_2)$, and $a_0 = r_1r_2$, which yields

$$a_1^2 = (r_1 + r_2)^2 = r_1^2 + r_2^2 + 2r_1r_2 > 2r_1r_2 = 2a_0a_2.$$

The inequality is strict since $r_1^2 + r_2^2 > 0$. To complete the induction, suppose that the conclusion holds for all polynomials of degree $n \geq 2$ satisfying the given hypotheses, and that p is a degree $n+1$ polynomial satisfying the given hypotheses. Then we can write

$$p(x) = (x - \alpha)(x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0)$$

where $x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0$ satisfies the hypotheses of the lemma and α is a nonzero real root of p. Then it holds $b_1^2 > 2b_0 b_2$ and

$$a_0 = -\alpha b_0, \quad a_1 = b_0 - \alpha b_1, \quad a_2 = b_1 - \alpha b_2.$$

Therefore,

$$a_1^2 - 2a_0 a_2 = \alpha^2(b_1^2 - 2b_0 b_2) + b_0^2 > 0$$

since $b_0 \neq 0$ and $b_1^2 - 2b_0 b_2 > 0$. This completes the induction, proving the lemma. $\square$

Proof of 21. By replacing p with $p(\cdot + a)$, we can assume without loss of generality that $a = 0$. Writing $p(x) = a_n x^n + \cdots + a_2 x^2 + a_1 x + a_0$ for real $a_0, \ldots, a_n$, the condition $p(0) \neq 0$ and $p'(0) = p''(0) = 0$ is equivalent to $a_0 \neq 0$ and $a_1 = a_2 = 0$. If all roots of p were real, then p satisfies the hypotheses of lemma 22, which means $a_1^2 > 2a_0 a_2$. But this is clearly false, since $a_1^2 = 0 = 2a_0 a_2$. So p has a non-real root. $\square$

Exercise 23 (Riemann Rearrangement). Suppose that $\{a_n\}$ is a sequence whose series converges absolutely. Then for any permutation σ of $\mathbb{N}$, $\sum a_{\sigma(n)} = \sum a_n$.

Proof. We first assume that $a_n \geq 0$ for all n. Fix $\epsilon > 0$. Then there exists N sufficiently large that

$$\sup_{m \geq N} \sum_{n=N}^{m} a_n < \epsilon.$$

Let $E = \{n : \sigma(n) \leq N\}$, and let $N' = \max E$. Since σ is a permutation of $\mathbb{N}$, it holds E contains only finitely many numbers, so $N' < \infty$. If $n > N'$, then we have $\sigma(n) > N$, so

$$\sum_{n=N'}^{m} a_{\sigma(n)} \leq \sum_{n=\sigma(N)}^{\sigma(m)} a_n \leq \sup_{m \geq N} \sum_{n=N}^{m} a_n < \epsilon.$$

Thus

$$\sup_{m \geq N'} \sum_{n=N'}^{m} a_{\sigma(n)} < \epsilon.$$

This proves that $\sup_j \sum_1^j a_n < \infty$, i.e., the series $\{a_{\sigma(n)}\}$ converges.

Now, if we choose $M_N > N'$ large enough that $\{1, \ldots, N\} \subset \{\sigma(1), \ldots, \sigma(M_N)\}$, then

$$\left| \sum_{j=1}^{N} a_n - \sum_{j=1}^{M_N} a_{\sigma(n)} \right| = \sum_{\{j \leq M_N : \sigma(j) > N\}} a_{\sigma(n)}.$$

Now, if we put $E_N := \{j \leq M_N : \sigma(j) > N\}$, then $\min E_N \to \infty$ as $N \to \infty$. Thus if N is sufficiently large, we have

$$\left| \sum_{j=1}^{N} a_n - \sum_{j=1}^{M_N} a_{\sigma(n)} \right| = \sum_{\{j \leq M_N : \sigma(j) > N\}} a_{\sigma(n)} < \epsilon.$$

Taking $N \to \infty$ and using $M_N \to \infty$,

$$\left| S - S' \right| \leq \epsilon,$$

where $S = \sum a_n$ and $S' = \sum a_{\sigma(n)}$. Since $\epsilon > 0$ was arbitrary, $S = S'$.

For the full statement, we now assume that $\{a_n\}$ is any sequence (not necessary non-negative) that tends to $+\infty$ as $n \to \infty$. Break a_n into a_n^+ and a_n^- and apply the preceding work to each part and then combine to obtain the theorem. $\square$

Exercise 24. Let

$$p_n(x) = \sum_{j=0}^{n} \frac{x^j}{j!}.$$

Then p_n has n distinct roots $z_1, \ldots, z_n$ and

$$\sum_{k=1}^{n} z_k^{-j} = 0, \quad 2 \leq j \leq n.$$

Proof. Clearly, 0 is not a root of p_n for any $n \geq 1$, so the n complex roots $z_1, \ldots, z_n$ are all nonzero. Consider the reversed polynomial

$$q_n(z) = \sum_{j=0}^{n} \frac{x^j}{(n-j)!}.$$

Then q_n has roots $x_j = 1/z_j$ for $1 \leq j \leq n$. For $0 \leq j \leq n$, define

$$e_j = (-1)^j \sum_{0 \leq i_1 < \cdots < i_j \leq n} x_{i_1} \cdots x_{i_j}$$

and $e_0 = 1$. Then by Vieta's formulas we have $e_j = \frac{1}{j!}$. We now prove the identity $p_j := x_1^j + \cdots + x_n^j = 0$ by induction on j. First, we observe that $e_1 = -p_1 = 1$. To see that $p_2 = 0$, by Newton's identity we have

$$2e_2 + p_2 + p_1 = 0.$$

Since $2e_2 = 1$ and $p_1 = -1$, we see $p_2 = 0$. Fix $2 \leq j < n$, and suppose we know $p_i = 0$ for $2 \leq i \leq j$. Again applying Newton's identity, we see

$$(k+1)e_{k+1} + p_{k+1} + p_1 e_k = 0$$

by the induction hypothesis $p_2 = \cdots = p_k = 0$. However, $(k+1)e_{k+1} = \frac{1}{k!}$ and $p_1 e_{k-1} = -\frac{1}{k!}$, which again leaves us with $p_{k+1} = 0$, completing the induction.

To see that the roots are all distinct, if there is a repeated root z_j, then z_j is a root of both p_n and p_n'. But this is impossible, since this would mean

$$0 = p_n'(z_j) = p_n(z_j) - z_j^n/n! = -z_j^n/n!,$$

and in particular $z_j = 0$. Since 0 is not a root, we have a contradiction, so $z_1, \ldots, z_n$ are all distinct. □

Let's try this exercise:

Exercise 25. Suppose that x is an algebraic number. Then there exists a constant $c > 0$ and a positive integer d such that for each $p/q \in \mathbb{Q}$ with $q > 0$ written in lowest terms,

$$\left| x - \frac{p}{q} \right| \geq \frac{c}{q^d}.$$

Proof. We can suppose that $x \neq 0$. Let f be the minimal polynomial of x and let d be the degree of f. Since f is a polynomial of lowest degree with x as a root, it holds $f'(x) \neq 0$. Let $\delta = |f'(x)|$. There is $\epsilon > 0$ such that $|f'(\alpha)| \geq \delta/2$ if $\alpha \in (x-\epsilon, x+\epsilon)$. By shrinking ϵ if necessary, we can assume that f has no other real roots in the closed interval $[x-\epsilon, x+\epsilon]$. If $p/q \in \mathbb{Q}$ satisfies $|x - p/q| < \epsilon$, then there is $\alpha \in (x-\epsilon, x+\epsilon)$ such that

$$x - \frac{p}{q} = \frac{-f(p/q)}{f'(\alpha)}.$$

On the other hand, since f is a monic polynomial of degree d, and f has no other real roots in the interval $[x-\epsilon, x+\epsilon]$, there is a constant $D > 0$ such that $|f(z)| \geq D|z|^d$ if $|x-z| < \epsilon$. Since p/q is such a number it holds

$$\left|x - \frac{p}{q}\right| = \left|\frac{f(p/q)}{f'(\alpha)}\right| \geq \frac{2Dp^d}{\delta q^d} \geq \frac{C'}{q^d}$$

where $C' = \frac{2D}{\delta}$. If $|x - p/q| > \epsilon$, then since $\epsilon > \epsilon q^{-d}$ for any $q \geq 1$, by taking $C = \min\{\epsilon, C'\} > 0$, we find

$$\left|x - \frac{p}{q}\right| \geq \frac{C}{q^d},$$

as required.

□

Exercise 26. Suppose that f is a differentiable function. Show that there exists $\alpha \in (0, 2\pi)$ such that the vector $(f(\alpha), f'(\alpha)+1)$ is perpendicular to $(\cos\alpha, \sin\alpha)$.

Proof. We need to show that there is $\alpha \in (0, 2\pi)$ such that

$$f(\alpha)\cos\alpha + f'(\alpha)\sin\alpha + \sin\alpha = 0.$$

This is the same as

$$(f(\alpha)(\sin\alpha - \cos\alpha))' = 0.$$

Observe that $f(\beta_0)(\sin\beta_0 - \cos\beta_0) = f(\beta_1)(\sin\beta_1 - \cos\beta_1) = 0$ when $\beta_0 = \pi/4$ and $\beta_1 = 5\pi/4$, so by Rolle's theorem there is $\alpha \in (\pi/4, 5\pi/4)$ such that

$$(f(\alpha)(\sin\alpha - \cos\alpha))' = 0.$$

□

Exercise 27. Let $K : [0,1] \times [0,1] \to \mathbb{R}$ be a continuous operator where $K(x,y) = K(y,x)$ for all x, y. Then the operator $T : L^2([0,1]) \to L^2([0,1])$

$$Tf(x) := \int_0^1 K(x,y) f(y)\, dy$$

is a compact operator. In particular, there exists a countable sequence $\{\lambda_k\}$ of real eigenvalues converging to 0 and functions $f_k \in L^2([0,1])$ such that

$$\lambda_k f_k(x) = \int_0^1 K(x,y) f_k(y)\, dy$$

for all $x \in [0,1]$. Moreover, the functions f_k are orthogonal.

Proof. By dominated convergence, Tf is continuous, and by Minkowski's integral inequality T is bounded. If $\|f\|_2 \leq 1$, then

$$|Tf(x) - Tf(y)| \leq \|K(x,\cdot) - K(y,\cdot)\|_2$$

and $\|Tf\|_\infty \leq \|K\|_\infty$ by Cauchy Schwarz, which means that any sequence in $\{Tf : \|f\|_2 \leq 1\} =: T(B)$ has a uniformly convergent subsequence by Arzela-Ascoli. The convergence also happens in L^2, which proves that $T(B)$ has a compact closure in L^2. Thus T is compact, and the theorem follows. □

Exercise 28 (Cantor's Theorem). Let X and Y be sets, and let $X \sim Y$ denote the existence of a bijection $f : X \to Y$. Moreover, let 2^X denote the power set of X. Then we have $2^X \not\sim X$ for any X.

Proof. We can assume that $X \neq \emptyset$. First, we claim that $2^X \sim \{0,1\}^X$, where $\{0,1\}^X$ is the set of functions $f : X \to \{0,1\}$. Define a map $T : \{0,1\}^X \to 2^X$ by $T(f) = \{x \in X : f(x) = 1\} \in 2^X$. This is evidently an injection, since $T(f) = T(g)$ implies that $f(x) = g(x) = 1$ whenever $x \in T(f)$ and moreover $f(x) = g(x) = 0$ for all $x \notin T(f)$. Thus $f = g$. It is a surjection, since if $A \in 2^X$, we have $Tf = A$, where $f(x) = 1$ if $x \in A$ and $f(x) = 0$ if $x \notin A$. This proves $2^X \sim \{0,1\}^X$.

So all that we need to do is show that $\{0,1\}^X \not\sim X$. Given a map $T : X \to \{0,1\}^X$, define a function $g \in \{0,1\}^X$ by $g(x) = 1 - (T(x))(x)$. Then $g \neq T(x)$ for any $x \in X$, since $g(x) \neq T(x)(x)$. This proves that any map $T : X \to \{0,1\}^X$ is not surjective, completing the proof. $\square$

As a corollary, we deduce that $\mathbb{R}^{\mathbb{R}} \not\sim \mathbb{R}$, since

$$\mathbb{R}^{\mathbb{R}} \sim \{0,1\}^{\mathbb{N}\times\mathbb{R}} \sim \{0,1\}^{\mathbb{R}} \sim 2^{\mathbb{R}}.$$

Exercise 29. Let $f \in C^1(\mathbb{R}^n)$ be a map such that $f'(x)$ is non-singular for each $x \in \mathbb{R}^n$. If $f^{-1}(K)$ is compact for each compact set K, then $f(\mathbb{R}^n) = \mathbb{R}^n$.

Proof. If $y \in f(\mathbb{R}^n)$, then fix $x \in \mathbb{R}^n$ where $f(x) = y$. Since $f'(x)$ is nonsingular, the inverse function theorem furnishes neighborhoods U and V of x and Y respectively such that $f(U) = V$ and $f|_U$ is a bijection. It follows that y is an interior point of $f(\mathbb{R}^n)$, and since y was an arbitrary element of $f(\mathbb{R}^n)$ we deduce $f(\mathbb{R}^n)$ is open.

On the other hand, $f(\mathbb{R}^n)$ is closed. Indeed, if $f(x_n) \to y$ for some $y \in \mathbb{R}^n$ as $n \to \infty$, we claim that $y \in f(\mathbb{R}^n)$. Indeed, for each $n \geq 1$ let $K_n = \{x : |x - y| \leq n^{-1}\}$, which is compact for all n. Then we have

$$\bigcap_{n\geq 1} f^{-1}(K_n) = \{x : f(x) = y\}.$$

Each $f^{-1}(K_n)$ is nonempty since it contains infinitely many of the points $\{x_k\}$. Since $f^{-1}(K_{n+1}) \subset f^{-1}(K_n)$ and $f^{-1}(K_n)$ is compact for all n, we see

$$\bigcap_{n\geq 1} f^{-1}(K_n) \neq \emptyset,$$

so there is a point $x \in \mathbb{R}^n$ where $f(x) = y$.

This proves $f(\mathbb{R}^n)$ is both open and closed, and since $\mathbb{R}^n$ is connected and $f(\mathbb{R}^n) \neq \emptyset$, we conclude that $f(\mathbb{R}^n) = \mathbb{R}^n$. $\square$

Remark 30. One can replace $\mathbb{R}^n$ by any open connected set U provided that $f : U \to U$.

Exercise 31. Let G be a finite group and K be a field. Let $K[G]$ be the set of functions $f : G \to K$, which is a vector space over K when endowed with pointwise operations and is also a group when endowed with multiplication

$$(\alpha\beta)(g) := \sum_{u\in G} \alpha(u)\beta(u^{-1}g), \quad \alpha, \beta \in K[G].$$

Show that the center of this group is a vector subspace of $K[G]$ whose dimension is the number of distinct conjugacy classes in G.

Lemma 32. *The center of $K[G]$ is the set of functions $f : G \to K$ such that $f(xy) = f(yx)$ for all $x, y \in G$.*

Proof. Fix $x, y \in G$ and let $\beta : G \to K$ be the function where $\beta(y^{-1}) = 1$ and $\beta(z) = 0$ if $z \neq y^{-1}$. If f is in the center of $K[G]$, then

$$f(xy) = \sum_{u\in G} f(u)\beta(u^{-1}x) = f\beta(x) = \beta f(x) = \sum_{u\in G} \beta(u) f(u^{-1}x) = f(yx).$$

On the other hand, if $f(xy) = f(yx)$ for all $x, y \in G$, then for any $h \in K[G]$, it holds

$$fh(x) = \sum_{u \in G} f(u)h(u^{-1}x) = \sum_{u \in G} f(xu^{-1})h(u) = \sum_{u \in G} f(u^{-1}x)h(u) = hf(x).$$

Thus f is in the center of $K[G]$. $\square$

Proof. It is clear from lemma 32 that the center of $K[G]$ is a vector subspace. For convenience, let V denote the center of $K[G]$. For $x \in G$, let $C_x := \{gxg^{-1} : g \in G\}$ be the conjugacy class of x. Let $Q_x : G \to K$ be the function such that $Q_x(z) = 1$ if $z \in C_x$ and $Q_x(z) = 0$ if $z \notin C_x$. If there are M distinct conjugacy classes, then there are M elements in the set $\{Q_x : x \in G\}$. So, if we can show $\{Q_x : x \in G\}$ is a basis of V, then we are done.

To this end, first we show $Q_x \in V$. Indeed, if $Q_x(ab) = 1$, then $ab = gxg^{-1}$ for some $g \in G$, so $ba = (bg)x(bg)^{-1} \in C_x$ and therefore $Q_x(ba) = 1$. The same reasoning shows $Q_x(ab) = 0$ implies $Q_x(ba) = 0$, and so $Q_x(ab) = Q_x(ba)$ for all $a, b \in G$ and by lemma 32 we conclude $Q_x \in V$.

To establish linear independence, fix $x_1, \dots, x_M \in G$ where $\{Q_x : x \in G\} = \{Q_{x_1}, \dots, Q_{x_m}\}$, and suppose $\alpha_1, \dots, \alpha_m \in K$ are chosen so that

$$\sum_{i=1}^{M} \alpha_i Q_{x_i} = 0.$$

By taking $x = x_k$ for each $k = 1, \dots, M$, and since distinct conjugacy classes are disjoint, we see $\alpha_1 = \cdots = \alpha_M = 0$, so $\{Q_x : x \in G\}$ is linearly independent.

Finally, we claim

$$f = \sum_{i=1}^{M} f(x_i) Q_{x_i}.$$

Indeed, if $x \in G$, find the unique x_j such that $x \in C_{x_j}$. Writing $x = gx_jg^{-1}$, we have $f(x) = f(gx_jg^{-1}) = f(g^{-1}gx_j) = f(x_j)$ by lemma 32. Since $\sum_{i=1}^{M} f(x_i)Q_{x_i}(x) = f(x_j)$, this proves

$$f(x) = \sum_{i=1}^{M} f(x_i) Q_{x_i}(x),$$

for all $x \in G$, as required. $\square$

Theorem 33. *Every conformal map $f : \mathbb{D} \to \mathbb{D}$ is of the form*

$$f(z) = \zeta \frac{\alpha - z}{1 - \overline{\alpha} z}$$

where α, ζ are constants satisfying $|\alpha| < |\zeta| = 1$.

Proof. Clearly all such functions of the form $\zeta \frac{\alpha - z}{1 - \overline{\alpha} z}$ are conformal. Let $g(z) = f^{-1}(z)$ and $\alpha = g(0)$. For a fixed $z_0 \in \mathbb{D}$, let $\beta_{z_0}(z) = \frac{z_0 - z}{1 - \overline{z}_0 z}$. Then we have $f \circ \beta_\alpha(0) = 0$ and $|f \circ \beta_\alpha| \leq 1$ on $\mathbb{D}$, so by Schwarz's lemma

$$|f \circ \beta_\alpha(z)| \leq |z|$$

for all $z \in \mathbb{D}$. It follows

$$|f(z)| \leq |\beta_\alpha(z)|$$

for all $z \in \mathbb{D}$. Moreover, if equality holds for some $z \in \mathbb{D}$, then there is $|\zeta| = 1$ such that $f(z) = \zeta \beta_\alpha(z)$ for all $z \in \mathbb{D}$. So we need to show that $|f(z)| \geq |\beta_\alpha(z)|$. To this end, observe that since $|g(z)| \leq 1$, we have $\beta_\alpha \circ g$ has $\beta_\alpha \circ g(0) = 0$ and $|\beta_\alpha \circ g(z)| \leq 1$ on $|z| < 1$, so another application of Schwarz's lemma implies $|\beta_\alpha \circ g(z)| \leq |z|$, hence $|\beta_\alpha(z)| \leq |f(z)|$. The conclusion now follows. $\square$

Exercise 34. Let f be analytic on the closed unit disk and assume that $|f(z)| \leq 1$ on this set. Suppose also that $f(1/2) = f(i/2) = 0$. Prove that $|f(0)| \leq 1/4$.

Lemma 35. *Suppose that $f : \mathbb{D} \to \mathbb{D}$ satisfies $f(\alpha) = 0$ for some $\alpha \in \mathbb{D}$. Then $|f(0)| \leq |\alpha|$.*

Proof. Apply Schwarz's lemma to the function

$$f\left(\frac{\alpha - z}{1 - \overline{\alpha} z}\right)$$

to see $|f(z)| \leq \left|\frac{\alpha - z}{1-\overline{\alpha}z}\right|$ for all $z \in \mathbb{D}$. Take $z = 0$ to recover the stated inequality. $\square$

Solution. Let m be the multiplicity of the zero at $1/2$. Then $\frac{f(z)}{(z-1/2)^m}$ can be extended to a holomorphic function on $\mathbb{D}$ such that $f(1/2) \neq 0$. It follows that the function

$$g(z) = \frac{f(z)}{\left(\frac{\alpha - z}{1-\overline{\alpha}z}\right)^m}$$

is a holomorphic function such that $g(i/2) = 0$ and $g(1/2) \neq 0$. By the lemma, $|g(0)| \leq |i/2| = 1/2$, and since $|g(0)| = 2|f(0)|$, we see $|f(0)| \leq 1/4$. $\square$

Theorem 36. *Suppose that f is a nonzero holomorphic on the annulus $\{a < |z| < b\}$ and continuous on $\{a \leq |z| \leq b\}$, where $0 < a < b$. For $a \leq t \leq b$, let*

$$M_t = \max_{|z|=t} |f(z)|, \quad m_t = \min_{|z|=t} |f(z)|.$$

Then

$$m_a^{1-p(t)} m_b^{p(t)} \leq m_t \leq M_t \leq M_a^{1-p(t)} M_b^{p(t)},$$

where $p(t) = \frac{\log t - \log a}{\log b - \log a}$.

Proof. You can use Brownian motion, or you can apply max modulus to the harmonic function

$$g(z) = \log|f(z)| - \frac{\log b - \log|z|}{\log b - \log a} \log M_a - \frac{\log|z| - \log a}{\log b - \log a} \log M_b.$$

$\square$

Exercise 37. Suppose that $f : \mathbb{D} \to \mathbb{D}$ is a holomorphic function continuous on $\overline{\mathbb{D}}$ such that $|f(z)| = 1$ when $|z| = 1$. Then either f is constant or of the form

$$f(z) = \zeta \prod_{i=1}^{n} \left(\frac{\alpha_i - z}{1 - \overline{\alpha_i} z}\right)^{m_i}$$

where $\alpha_1, \ldots, \alpha_n \in \mathbb{D}$, $m_1, \ldots, m_n$ are positive integers, and $|\zeta| = 1$ are constants.

Lemma 38. *If $f : \mathbb{D} \to \mathbb{D}$ is a holomorphic function, continuous on $\mathbb{D}$, such that $|f(z)| = 1$ whenever $|z| = 1$, then f has no zeroes in $\mathbb{D}$ if and only if f is constant.*

Proof. Trivially f has no zeros in $\mathbb{D}$ if it is a constant. For the converse, suppose that f has no zeros in $\mathbb{D}$. Since $|f(z)| = 1$ for $|z| = 1$, it holds $|f(z)| \leq 1$ on $\overline{\mathbb{D}}$. Since f has no zeros on $\overline{\mathbb{D}}$, there is $\delta > 0$ such that $|f(z)| \geq \delta$. This implies $z \mapsto 1/\overline{f(1/\overline{z})}$ is holomorphic on $\mathbb{C} \setminus \overline{\mathbb{D}}$, and has $|1/\overline{f(1/\overline{z})}| = 1$ for $|z| = 1$. By the maximum modulus principle, it thus holds $1 \leq |1/\overline{f(1/\overline{z})}| \leq 1$. This proves $|f(z)| = 1$ on $\mathbb{D}$. Thus f is a constant by the open mapping theorem. $\square$

Proof. First, we will assume that f has a zero. Since $\mathbb{D}$ is bounded, f has only finitely many zeros in $\mathbb{D}$ otherwise the zero set has a limit point. This limit point would belong to $\mathbb{D}$ since $|f(z)| = 1$ on $|z| = 1$, so that $f \equiv 0$ on $\mathbb{D}$. So, let $\alpha_1, \ldots, \alpha_n$ be the zeros of f, and let $m_1, \ldots, m_n$ be the respective multiplicities. Then

$$\frac{f(z)}{(\alpha_1 - z)^{m_1} \cdots (\alpha_n - z)^{m_n}}$$

can be extended to a non-vanishing holomorphic function, and in particular

$$\frac{f(z)}{\prod_{i=1}^{n}\left(\frac{\alpha_i - z}{1-\overline{\alpha_i} z}\right)^{m_i}}$$

has no zeros and has modulus 1 on $|z| = 1$. By the lemma, it is a constant, and the theorem follows. $\square$

Exercise 39. Fix real numbers $0 < a_1 < b_1$ and $0 < a_2 < b_2$ and let $A_1 = \{z : a_1 < |z| < b_1\}$ and $A_2 = \{z : a_2 < |z| < b_2\}$. Prove that A_1 and A_2 are conformally equivalent if and only if $\frac{b_1}{a_1} = \frac{b_2}{a_2}$. Moreover, show that any such conformal map $f : A_1 \to A_2$ is of the form

$$f(z) = c\frac{a_2}{a_1}z$$

where $|c| = 1$.

Proof. Let $r_1 = \frac{b_1}{a_1}$ and $r_2 = \frac{b_2}{a_2}$. If $r_1 = r_2 = r$, then there is a conformal map from A_1 to $\{1 < |z| < r\}$ given by $f(z) = z/a_1$. For the same reason there is a conformal map from A_2 to $\{1 < |z| < r\}$, hence A_1 and A_2 are conformally equivalent.

For the converse, since A_1 is conformally equivalent to $\{1 < |z| < r_1\}$ and A_2 is conformally equivalent to $\{1 < |z| < r_2\}$, we can suppose without loss of generality $A_1 = \{1 < |z| < r_1\}$ and $A_2 = \{1 < |z| < r_2\}$. Suppose that $f : A_1 \to A_2$ is conformal. For $1 < |z| < r_1$, let $g(z) = \frac{f(z)}{z}$. Given $1 < r < r_1$, on $|z| = r$ we have

$$\left|\frac{f(z)}{z}\right| \leq \frac{r_2}{r},$$

which, by maximum modulus, implies

$$|f(z)| \leq \frac{r_2}{r}|z|$$

on $|z| \leq r$. Taking $r \to r_1$,

$$|f(z)| \leq \frac{r_2}{r_1}|z|$$

for all $r_1 < |z| < r_2$. Similarly,

$$|f^{-1}(z)| \leq \frac{r_1}{r_2}|z|,$$

so

$$|f(z)| \geq \frac{r_2}{r_1}|z|.$$

This proves $|\frac{r_2}{r_1}g(z)| = 1$ on A_1, and by the open mapping theorem this proves $f(z) = c\frac{r_2}{r_1}z$ for some constant $|c| = 1$. Since $|f(z)| > 1$, we see $r_2 \geq r_1$. Since $|f^{-1}(z)| > 1$, this shows $r_1 \geq r_2$. Thus $r_1 = r_2$ and $f(z) = cz$ for some $|c| = 1$.

For the final claim, we revert to assuming $A_1 = \{z : a_1 < |z| < b_1\}$ and $A_2 = \{z : a_2 < |z| < b_2\}$. Given a conformal map $f : A_1 \to A_2$, it holds $a_2^{-1}f(a_1 z)$ is a conformal map $\{1 < |z| < r\} \to \{1 < |z| < r\}$, which by our previous argument shows $a_2^{-1}f(a_1 z) = cz$ for some $|c| = 1$. Thus $f(z) = c\frac{a_2}{a_1}z$, finishing the proof. $\square$

Exercise 40. Suppose that f is a complex valued function defined in an open neighborhood Ω containing the closed unit disk. Suppose further that f is holomorphic in Ω except at a pole z_0 where $|z_0| = 1$. Then

$$\lim_{n\to\infty}\frac{a_n}{a_{n+1}} = z_0,$$

where the a_n are chosen so that $f(z) = \sum_{n\geq 0} a_n z^n$ in a neighborhood of the origin.

Theorem 41. *Suppose that f is holomorphic in a domain Ω except at a pole $z_0 \in \Omega$ of order m. Then f can be written in the form*

$$f(z) = \sum_{j=1}^{m} \frac{b_j}{(z_0 - z)^j} + g(z)$$

for all $z \in \Omega \setminus \{z_0\}$, where g is analytic in Ω. The function $\sum_{j=1}^{m} \frac{b_j}{(z_0-z)^j}$ is called the principle part of f at z_0.

Proof. The function $(z - z_0)^m f(z)$ can be analytically continued onto all of Ω, hence

$$(z - z_0)^m f(z) = b_m + b_{m-1}(z - z_0) + \cdots + b_1(z - z_0)^{m-1} + r(z),$$

where b_j is the $(m - j)$-th Taylor coefficient of $(z - z_0)^m f(z)$, so that r is an analytic function with a zero at z_0 of order m. It follows that $r(z) = (z - z_0)^m g(z)$ on Ω for some analytic function g by analytic continuation. Dividing both sides by $(z - z_0)^m$ and changing the signs of the b_j as necessary completes the proof. $\square$

Theorem 42. *Suppose that f is a holomorphic function in some domain Ω containing the closed unit disk, except possibly at a point z_0. Suppose that*

$$f(z) = \sum_{n \geq 0} a_n z^n$$

in a neighborhood of the origin and that the principle part of f at z_0 is $\sum_{j=1}^{m} \frac{b_j}{(z_0-z)^j}$. Then

$$a_n = \sum_{j=1}^{m} \binom{n+j-1}{j-1} \frac{b_j}{z_0^{n+j}} + r_n,$$

where r_n is a bounded sequence of complex numbers.

Proof. Write

$$f(z) = \sum_{j=1}^{m} \frac{b_j}{(z_0 - z)^j} + g(z)$$

for some analytic function g. Then

$$n! a_n = f^{(n)}(0) = n! \sum_{j=1}^{m} \binom{n+j-1}{j-1} \frac{b_j}{z_0^{n+j}} + g^{(n)}(0).$$

It follows

$$a_n = \sum_{j=1}^{m} \binom{n+j-1}{j-1} \frac{b_j}{z_0^{n+j}} + \frac{g^{(n)}(0)}{n!}.$$

Letting $r_n = \frac{g^{(n)}(0)}{n!}$, we see that the sequence r_n is bounded since Cauchy's inequality implies

$$|r_n| \leq \max_{|z|=1} |g(z)|$$

for all $n \geq 0$. $\square$

Solution to the Problem. We will first assume that $z_0 = 1$. In this case, we have

$$a_n = \sum_{j=1}^{m} \binom{n+j-1}{j-1} b_j + r_n$$

for some bounded sequence $\{r_n\}$ of complex numbers. Thus

$$\frac{a_n}{a_{n+1}} = \frac{\sum_{j=1}^m \binom{n+j-1}{j-1} b_j + r_n}{\sum_{j=1}^m \binom{n+j}{j-1} b_j + r_{n+1}}.$$

Observe that

$$\binom{n+j-1}{j-1}$$

is a degree $j-1$ polynomial of n with leading coefficient $\frac{1}{(j-1)!}$, so

$$\frac{\sum_{j=1}^m \binom{n+j-1}{j-1} b_j}{\sum_{j=1}^m \binom{n+j}{j-1} b_j}$$

is a ratio of two degree m polynomials in n. The leading coefficient of both polynomials is $\frac{b_m}{(m-1)!}$, so

$$\lim_{n\to\infty} \frac{\sum_{j=1}^m \binom{n+j-1}{j-1} b_j}{\sum_{j=1}^m \binom{n+j}{j-1} b_j} = 1.$$

Since the sequence $\{r_n\}$ is bounded, it thus holds

$$\lim_{n\to\infty} \frac{\sum_{j=1}^m \binom{n+j-1}{j-1} b_j + r_n}{\sum_{j=1}^m \binom{n+j}{j-1} b_j + r_{n+1}} = 1.$$

This proves

$$\lim_{n\to\infty} \frac{a_n}{a_{n+1}} = 1.$$

In the case $z_0 \neq 1$, consider the function $g(z) = f(z_0 z)$, which is holomorphic in Ω except at the pole $z = 1$. Since $g(z) = \sum_{n\geq 1} (z_0^n a_n) z^n$, our prior work shows

$$\lim_{n\to\infty} \frac{z_0^n a_n}{z_0^{n+1} a_{n+1}} = 1.$$

Rearranging, we get

$$\lim_{n\to\infty} \frac{a_n}{a_{n+1}} = z_0,$$

as required. □

Remark 43. If instead f had multiple poles $z_1, \ldots, z_n$ on $|z| = 1$, this same proof would establish the following result:

$$\lim_{n\to\infty} \frac{a_n}{a_{n+1}} = \frac{\sum_{j=1}^N \frac{r_j}{z_j^M}}{\sum_{j=1}^N \frac{r_j}{z_j^{M+1}}}$$

where M is the maximal order of all the poles, and by re-indexing we assume $z_1, \ldots, z_N$ are the poles on $|z| = 1$ with order M, and r_j is the residue of z_j. If $\sum_{j=1}^N \frac{r_j}{z_j^{M+1}} = 0$, then the limit may not exist, since then you are dealing with the asymptotics of a rational function where the degree of the numerator is potentially greater than the degree of the denominator. As an example of this phenomenon, consider $f(z) = \frac{1}{1-z^2}$.

Corollary 44. *Suppose that f is a complex valued function defined in an open neighborhood Ω containing the closed unit disk. Suppose further that f is holomorphic in Ω except at a pole z_0 of order m where $|z_0| = 1$. If $f(z) = \sum_{n\geq 1} a_n z^n$ in a neighborhood of the origin*

$$a_n = \sum_{j=1}^m \binom{n+j-1}{j-1} \frac{b_j}{z_0^{n+j}} + o(1),$$

where $o(1)$ denotes a quantity that tends to 0 as $n \to \infty$.

PROOF. We know that

$$a_n = \sum_{j=1}^{m} \binom{n+j-1}{j-1} \frac{b_j}{z_0^{n+j}} + \frac{g^{(n)}(0)}{n!},$$

where g is an analytic function in Ω. We can pick $R > 1$ such that $B_R(0) \subset \Omega$. Since

$$\left| \frac{g^{(n)}(0)}{n!} \right| \leq \frac{1}{R^n} \max_{|z|=R} |g(z)|,$$

as $n \to \infty$ it follows $\frac{g^{(n)}(0)}{n!} \to 0$, and taking $r_n = \frac{g^{(n)}(0)}{n!}$ finishes the proof. □

COROLLARY 45. *Suppose that f is a complex valued function defined in an open neighborhood Ω containing the closed unit disk. Suppose further that f is holomorphic in Ω except at a simple pole at $z = z_0$. If $f(z) = \sum_{n\geq 1} a_n z^n$ in a neighborhood of the origin, then*

$$\lim_{n\to\infty} z_0^n a_n = -b$$

where b is the residue at $z = 1$ of the function $z \mapsto f(z_0 z)$.

EXERCISE 46. For each $n \geq 0$ let b_n be the number of ones in the binary expansion of n. Then the generating function f of b_n is given by

$$f(z) = \frac{g(z) - 2g(z^2)}{1-z},$$

where $|z| < 1$ and

$$g(z) = \sum_{n\geq 0} \frac{z^{2^n}}{1 - z^{2^n}}.$$

PROOF. First, observe that

$$b_n = \sum_{k=0}^{\infty} \left\lfloor \frac{n}{2^k} \right\rfloor - 2\left\lfloor \frac{n}{2^{k+1}} \right\rfloor.$$

Thus, it suffices to show that

$$g(x) = (1-x) \sum_{n\geq 0} \sum_{k\geq 0} \left\lfloor \frac{n}{2^k} \right\rfloor x^n.$$

Changing the order of summation,

$$\sum_{n\geq 0} \sum_{k\geq 0} \left\lfloor \frac{n}{2^k} \right\rfloor x^n = \sum_{k\geq 0} \sum_{n\geq 0} \left\lfloor \frac{n}{2^k} \right\rfloor x^n.$$

Now,

$$\sum_{n\geq 0} \left\lfloor \frac{n}{2^k} \right\rfloor x^n = \sum_{j\geq 0} j \sum_{k=2^n j}^{2^n(j+1)-1} x^n = \sum_{j\geq 0} j \frac{x^{2^n j} - x^{2^n(j+1)}}{1-x} = \frac{1}{1-x} \sum_{j\geq 1} x^{2^n j} = \frac{1}{1-x} \frac{x^{2^n}}{1-x^{2^n}}.$$

This completes the proof. □

The above function has an essential singularity at every dyadic rational on the unit circle.

EXERCISE 47. Suppose that Ω is a simply connected domain and f is a holomorphic function on Ω which does not vanish anywhere on Ω. Then there exists a holomorphic function g on Ω such that $e^{g(z)} = f(z)$ for all $z \in \Omega$.

Proof. Fix a point $p \in \Omega$. Since $f(p) \neq 0$, there is a branch of the logarithm such that $\log f(p)$ is well defined. Let

$$g(z) = \log f(p) + \int_{\Gamma(z)} \frac{f'(s)}{f(s)}\, ds,$$

where $\Gamma(z)$ denotes any continuous piecewise smooth curve in Ω from p to z. Since Ω is simply connected, such a path exists. Moreover, since f'/f is holomorphic on Ω and all such paths $\Gamma(z)$ are homotopic, the integral $\int_{\Gamma(z)} \frac{f'(s)}{f(s)}\, ds$ depends only on p and z. In particular, g is well-defined.

Next, we claim that g is holomorphic. Given $z \in \Omega$, let $\delta > 0$ be chosen small enough that $z + h \in \Omega$ whenever $|h| \leq \delta$. Given such h, and a path $\Gamma(z)$ from p to z, we can construct a path from p to $z_0 + h$ by appending the straight line segment $[z, z+h]$ to $\Gamma(z)$. The resulting path is a continuous and piecewise smooth path from p to $z+h$, so

$$\frac{g(z+h) - g(z)}{h} = \frac{1}{h} \int_{[z,z+h]} \frac{f'(s)}{f(s)}\, ds = \int_0^1 \frac{f'(z+th)}{f(z+th)}\, dt.$$

Now, since f does not vanish in Ω, the function

$$h \mapsto \frac{f'(z+h)}{f(z+h)}$$

for $|h| \leq \delta$ is uniformly bounded for a given z, so by dominated convergence

$$\lim_{h \to 0} \int_0^1 \frac{f'(z+th)}{f(z+th)}\, dt = \int_0^1 \lim_{h \to 0} \frac{f'(z+th)}{f(z+th)}\, dt = \frac{f'(z)}{f(z)}.$$

This proves that $g'(z) = \frac{f'(z)}{f(z)}$ for all $z \in \Omega$, hence g is holomorphic.

Finally, to see that $e^{g(z)} = f(z)$, let $r(z) = e^{g(z)} - f(z)$. Since $r'(z)f(z) = f'(z)r(z)$, by taking n derivatives we see that

$$\sum_{k=0}^{n} \binom{n}{k} r^{(k+1)}(z) f^{(n-k)}(z) = \sum_{k=0}^{n} \binom{n}{k} r^{(k)}(z) f^{(n-k+1)}(z).$$

Since $r(p) = 0$, if we have shown $r^{(k)}(p) = 0$ for $0 \leq k \leq n$ for some $n \geq 0$, it follows that

$$r^{(n+1)}(p) f(p) = 0.$$

Since $f(p) \neq 0$, we conclude $r^{(n+1)}(p) = 0$, and so by induction it follows $r^{(n)}(p) = 0$ for all n. This implies $r(z) = 0$ for all $z \in \Omega$, and we are done. □

Corollary 48. *If Ω is a simply connected set not containing 0, there is a holomorphic function $g(z)$ such that $e^{g(z)} = z$ for all $z \in \Omega$.*

Exercise 49. Suppose that $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$. Show that $|x| = |y|$ if and only if there is an orthogonal matrix Q such that $Qx = y$.

Proof. If there is an orthogonal matrix Q with $Qx = y$, then $|x| = |Qx| = |y|$. On the other hand, suppose $|x| = |y|$. We can assume that $x \neq 0$. Let $x_1 = x/|x|$ and $y_1 = y/|y|$. By the Gram-Schmidt procedure, we can furnish an orthonormal basis $x_2, \ldots, x_n$ of the $n-1$ dimensional space $\{z : x \cdot z = 0\}$ and similarly an orthonormal basis $y_2, \ldots, y_n$ of $\{z : z \cdot y = 0\}$. Then $x_1, \ldots, x_n$ form an orthonormal basis of $\mathbb{R}^n$, and similarly with $y_1, \ldots, y_n$. Let X be the $n \times n$ matrix whose columns are $x_1, \ldots, x_n$ and analogously define the matrix Y. Since the $x_1, \ldots, x_n$ are an orthonormal basis, it holds X is orthogonal, and so is Y. Let $Q = YX^\top$. Clearly, Q is orthogonal since $Q^\top Q = XY^\top YX^\top = XX^\top = I$ and $QQ^\top = YX^\top XY^\top = YY^\top = I$. Moreover, we claim $Qx = y$. Indeed, since $QX = Y$, it holds $Qx_1 = QXe_1 = Ye_1 = y_1$ where e_1 is the first standard basis vector. Multiplying both sides by $|x| = |y|$, we see $Qx = y$, and we are done. □

Exercise 50. Let F be a field. Then $F[x]$ is a PID. In particular, if I is a nonzero ideal, then there is a unique monic polynomial $p \in F[x]$ such that $I = (p)$.

Proof. Fix an ideal I of $F[x]$. We can suppose that $I \neq 0$ and $I \neq F[x]$. Then the only constant in I is 0, so there exists a nonconstant polynomial $p \in I$ where

$$\deg p = \min_{q \in I \setminus \{0\}} \deg q.$$

Clearly, $(p) \subset I$. If $q \in I$ and $q \neq 0$, then $\deg q \geq \deg p$. Therefore, we can find polynomials $r, s \in F[x]$ where $\deg r < \deg p$ and

$$q = sp + r.$$

Since $sp \in I$ and $q \in I$, it holds $q - sp \in I$. However, $\deg r < \deg p$, which means that $r = 0$. Thus $q = sp \in (p)$, which proves $I = (p)$.

Next, observe that by dividing p by its leading coefficient if necessary and using the fact I is invariant under scaling, we can assume that p is monic. Suppose that q is another monic polynomial which generates I. Then $\deg q = \deg p$, since otherwise I contains only polynomials of degree at least $\deg q > \deg p$, which would contradict the fact $p \in I$. Thus $p - q \in I$. However, $\deg(p - q) < \deg p$, which by construction of p implies that $p - q = 0$. Thus $p = q$, and we are done. $\square$

Exercise 51. Let F be a field and let A be a $n \times n$ matrix with entries in F. Then there exists a unique monic polynomial $m_A \in F[x]$ such that $m_A(A) = 0$ and if $p \in F[x]$ satisfies $p(A) = 0$, then $m_A \mid p$. Moreover, the roots of m_A in F are precisely the eigenvalues of A in F.

Proof. We can assume that $A \neq 0$. Let $I = \{p \in F[x] : p(A) = 0\}$. Clearly, I is an ideal of $F[x]$, so we need to show that $I \neq 0$. Since the vector space of square matrices has dimension n^2, the vectors $I, A, \ldots, A^{n^2}$ are linearly dependent, hence there are constants $\alpha_0, \ldots, \alpha_{n^2}$, not all zero, such that $\alpha_0 I + \cdots + \alpha_{n^2} A^{n^2} = 0$. If $p(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_{n^2} x^{n^2}$, clearly p is nonzero and $p \in I$, which shows that $I \neq 0$.

Thus there is a unique monic polynomial m_A such that $I = (m_A)$. All we need to show is that $\lambda \in F$ satisfies $m_A(\lambda) = 0$ if and only if $A - \lambda I$ is singular. First suppose that $A - \lambda I$ is singular. Then there is a nonzero vector $v \in F^n$ such that $Av = \lambda v$. Thus $0 = m_A(A)v = p(\lambda)v$. Since $v \neq 0$, this implies $m_A(\lambda) = 0$ and hence λ is a root of m_A. On the other hand, suppose that $m_A(\lambda) = 0$. Then there exists a monic polynomial p with $\deg p < \deg m_A$ such that $m_A(x) = (x - \lambda)p(x)$ for all $x \in F$. It follows $(A - \lambda I)p(A) = 0$. If $A - \lambda I$ is non-singular, then we must have $p(A) = 0$ and thus $m_A \mid p$, which contradicts the fact $\deg p < \deg m_A$. Thus $A - \lambda I$ is singular, and in particular λ is an eigenvalue of A in F. $\square$

Exercise 52. Let F be a field and K be an extension of F. Let A be a matrix with entries in F and m_A^K, m_A^F be the minimal polynomials of A over K and F, respectively. Then $m_A^K = m_A^F$.

Proof. Let $n = \deg m_A^K$. Clearly we have $m_A^K \mid m_A^F$ and in particular $n \leq \deg m_A^F$. Fix a (possibly infinite) basis $\mathcal{B}$ of the extension K/F. Suppose that

$$m_A^K(x) = x^n + \alpha_{n-1} x^{n-1} + \cdots + \alpha_0,$$

where $\alpha_0, \ldots, \alpha_{n-1} \in K$. By definining $\alpha_n := 1$, we can write

$$\alpha_j = \sum_{i=1}^{N_j} \beta_{ij} x_{ij}$$

where $\beta_{ij} \in F$ for each $j = 0, \ldots, n$ and some $x_{ij} \in \mathcal{B}$. Since there are only finitely many α_j, there are only finitely many (say, N) basis elements x_{ij} used to represent all of the $\alpha_1, \ldots, \alpha_n$, so by re-indexing the x_{ij} we can find $\alpha_{ij} \in F$ such that

$$\alpha_j = \sum_{i=1}^{N} \alpha_{ij} x_i$$

where $x_i \in \mathcal{B}$ for $1 \leq i \leq N$. Then it holds

$$0 = m_A^K(A) = \sum_{i=1}^{N} \left[\sum_{j=0}^{n} \alpha_{ij} A^j \right] x_i.$$

Since $\alpha_{ij} \in F$, all the entries of A are in F, and the x_i are linearly independent over F, it holds

$$\sum_{j=0}^{n} \alpha_{ij} A^j = 0$$

for each $1 \leq i \leq N$. Letting $q_i(x) = \alpha_{in} x^n + \cdots + \alpha_{i0} \in F[x]$, we see $q_i(A) = 0$, so $m_A^F \mid q_i$. Since $\deg q_i \leq n$, it follows $\deg m_A^F \leq n$, and thus we have $\deg m_A^F = n$. But m_A^K is the unique monic polynomial in $K[x]$ with degree n such that $m_A^K(A) = 0$, and this proves $m_A^F = m_A^K$. □

LEMMA 53. *The roots of the minimal polynomial of A are precisely the eigenvalues of A.*

PROOF. Let K be a field extension of F which contains all the eigenvalues of A and all roots of the minimal polynomial of A. Let p be the minimal polynomial of A, which is the same in both K and F. If $\lambda \in K$ is an eigenvalue of A, there is a nonzero vector v with $Av = \lambda v$. Then $0 = p(A) = p(\lambda)v$, and since $v \neq 0$ it follows that $p(\lambda) = 0$. On the other hand, if r is a root of p, then we can write $p(x) = (x - r)q(x)$ for some polynomial q with degree strictly less than the degree of p. It follows that $q(A) \neq 0$, hence there exists a nonzero vector w such that $q(A)w \neq 0$. It follows that $v := q(A)w$ satisfies $(A - rI)v = p(A)w = 0$, hence r is an eigenvalue of A. Thus the roots of p are precisely the eigenvalues of A. □

LEMMA 54. *The degree of the minimal polynomial p of A has degree at most $n = \dim V$.*

PROOF. First, suppose that there does not exist $x_1 \in V$ such that $x_1, Ax_1, \ldots, A^{n-1}x_1$ are linearly dependent, so that $x_1, \ldots, A^{n-1}x_1$ form a basis of V. Then there exist constants $c_0, \ldots, c_n$ such that $A^n x_1 = c_n A^{n-1} x_1 + \cdots + c_0 x_1$, and thus $A^n(A^m x_1) = c_n A^{n-1}(A^m x_1) + \cdots + c_0(A^m x_1)$ for $0 \leq m < n$. By linearity it holds $A^n x = c_n A^{n-1} x + \cdots + c_0 x$ for all $x \in V$ since $x_1, \ldots, A^{n-1}x_1$ form a basis of V. Thus the minimal polynomial of A divides $x^n - c_n x^{n-1} - \cdots - c_0$ and hence has degree n, so there is nothing to prove.

So, let us suppose that there exists $x_1 \in V$ such that $x_1, Ax_1, \ldots, A^{n-1}x_1$ are linearly dependent. We claim that we can decompose

$$V = \bigoplus_{i=1}^{k} W_i,$$

where each W_i is nontrivial and satisfies $A(W_i) \subset W_i$ for each i. Indeed, if $n = 1$, then the result is trivial, and so suppose that the result has been established for all vector spaces of dimension at most $n \geq 1$. Let $1 \leq m \leq n$ be the largest integer such that $x_1, Ax_1, \ldots, A^{m-1}x_1$ are linearly independent. Let us first suppose that $m < n$. Then the space W_1 spanned by $\{x_1, Ax_1, \ldots, A^{m-1}x_1\}$ is $m < n$ dimensional and A-invariant since $x_1, \ldots, A^{m-1}x_1, A^m x_1$ are linearly dependent. Pick a space W' such that $V = W_1 \oplus W'$. If $AW' \subset W'$ then we are done. Otherwise, if $k = \dim W'$, there exists $x_2 \in W$ such that $x_2, \ldots, T^{k-1}x_2$ are linearly dependent, otherwise W' is A-invariant. By the induction hypothesis $W' = \bigoplus_{i=2}^{k} W_i$ for A-invariant subspaces W_i, and the induction is complete.

To complete the proof, note that if $n = 1$ the proof is trivial. So suppose $n \geq 2$. By induction, if p_i is the minimal polynomial of $A|_{W_i}$, it then holds $\deg p_i \leq \dim W_i$. Moreover, if $q = p_1 \cdots p_k$, then $q(A)|_{W_i} = 0$ for each i, hence $q(A) = 0$. Thus $p \mid q$, hence $\deg p \leq \deg p_1 + \cdots + \deg p_k \leq \dim W_1 + \cdots + \dim W_k = n$, and the proof is done. $\square$

Corollary 55. *Let F be a field and suppose that L is some field extension of F. If $\alpha \in L$ and $\beta \in L$ are algebraic over F (i.e., the root of some polynomial with coefficients in F), then so are α^{-1}, β^{-1}, $\alpha\beta$ and $\alpha + \beta$.*

Proof. That α^{-1} and β^{-1} are algebraic is trivial. Let $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ be the minimal polynomial of α and $q(x) = x^m + b_{m-1}x^{m-1} + \cdots + b_0$ be the minimal polynomial of β. Let P be the rational canonical form corresponding to the polynomial p and Q be the rational canonical form corresponding to the polynomial q. Then P has characteristic polynomial p, and analogously for Q and q. This implies α is an eigenvalue of P and β is an eigenvalue for Q. To see that $\alpha\beta$ is algebraic over F, note that the tensor product $P \otimes Q$ has entries in F and $\alpha\beta$ as an eigenvalue, so the minimal polynomial $m_{P\otimes Q}(x)$ has $\alpha\beta$ as a root and coefficients in F.

Finally, to see that $\alpha + \beta$ is algebraic over F, let I_n be the $n \times n$ identity matrix and I_m be the $m \times m$ identity matrix. Let $X = P \otimes I_m + I_n \otimes Q$. Let $v \in L^n$ and $w \in L^m$ be eigenvectors of P and Q, respectively, with eigenvalues α and β. Then

$$X(v \otimes w) = P \otimes I_m(v \otimes w) + I_n \otimes Q(v \otimes w) = (\alpha + \beta)(v \otimes w).$$

Thus X is a matrix whose minimal polynomial m_X has $\alpha + \beta$ as a root, finishing the proof. $\square$

Corollary 56. *The algebraic integers are defined to be the set of real numbers which are roots of monic polynomials with integer coefficients. The algebraic integers are an integral domain containing $\mathbb{Z}$.*

Proof. If α, β are algebraic integers, then there are matrices with integer coefficients possessing α, β as eigenvalues. Then matrices possessing $\alpha\beta$ and $\alpha + \beta$ as eigenvalues can be formed through taking tensor products and sums of such matrices, and such matrices therefore have integer coefficients. It follows $\alpha\beta$ and $\alpha + \beta$ are the roots of the characteristic polynomials of these matrices, which much have integer coefficients since all the matrix entries are integers. $\square$

More generally, we have the following:

Corollary 57. *Let F be a field and $R \subset F$ be a subring containing 1. Let $F[R]$ be the set of monic polynomials with coefficients in R. Then the the elements of F which are roots of polynomials in $F[R]$ is an integral domain containing R.*

Example 58. Let $\alpha = \sqrt{2}$ and β be the golden ratio. Then we have the representation

$$\alpha \sim \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}, \quad \beta \sim \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix},$$

which implies that

$$\alpha\beta \sim \begin{bmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 2 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}, \quad \alpha + \beta \sim \begin{bmatrix} 0 & 1 & 2 & 0 \\ 1 & 1 & 0 & 2 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

Interestingly, this actually yields two new problem solving techniques. For starters, it gives you one way to compute minimal polynomials (or multiples thereof), since the minimal polynomial of the matrix representations of $\alpha\beta$ and $\alpha + \beta$ are multiples of the minimal polynomials of $\alpha\beta$ and $\alpha + \beta$.

Exercise 59. Let

$$f(z) = \sum_{n \geq 0} a_n z^n,$$

where all the a_n are non-negative reals, and the series has radius of convergence 1. Show that f cannot be analytically continued to a neighborhood of 1.

Lemma 60. *Let $\{a_n\}$ be a non-negative sequence of real numbers. Then*

$$\lim_{r \to 1^-} \sum_{n \geq 0} a_n r^n = \sum_{n \geq 0} a_n,$$

where the limit may be infinite.

Proof. Let $S_k = \sum_{n=0}^{k} a_n$. If $0 < r < 1$, then

$$\sum_{n \geq 0} a_n r^n \geq \sum_{n=0}^{m} a_n r^n \geq S_m r^m.$$

Thus

$$\liminf_{r \to 1^-} \sum_{n \geq 0} a_n r^n \geq S_m,$$

and since m was arbitrary,

$$\liminf_{r \to 1^-} \sum_{n \geq 0} a_n r^n \geq \sum_{n \geq 0} a_n.$$

On the other hand, since $r < 1$,

$$\limsup_{r \to 1^-} \sum_{n \geq 0} a_n r^n \leq \sum_{n \geq 0} a_n,$$

and this proves the claim:

$$\lim_{r \to 1^-} \sum_{n \geq 0} a_n r^n = \sum_{n \geq 0} a_n.$$

$\square$

Solution to Problem. To obtain a contradiction, suppose that f can be analytically continued to a neighborhood of 1. Then we can find coefficients $b_0, b_1, \ldots$ and $\delta > 0$ small enough that

$$f(z) = \sum_{n \geq 1} b_n (z-1)^n$$

whenever $|z - 1| < \delta$. By continuity, we have $f^{(k)}(1) = \lim_{r \to 1^-} f^{(k)}(r)$. By the lemma, this implies

$$b_k = \frac{f^{(k)}(1)}{k!} = \lim_{r \to 1} \sum_{n \geq k} \binom{n}{k} a_n r^{n-k} = \sum_{n \geq k} \binom{n}{k} a_n.$$

If we choose $x > 1$ such that $|x - 1| < \delta$, then

$$f(x) = \sum_{k \geq 0} \sum_{n \geq k} \binom{n}{k} a_n (x-1)^k = \sum_{n \geq 0} a_n \sum_{k=0}^{n} \binom{n}{k} (x-1)^k = \sum_{n \geq 0} a_n x^n,$$

where summation interchange was justified by the fact each term in the series is non-negative. By hypothesis, $f(x)$ is well-defined by analytic continuation, so $\sum a_n x^n$ is a finite real number. But $x > 1$, contradicting the fact that the the radius of convergence of $\sum_{n \geq 0} a_n z^n$ is 1. The contradiction is resolved only if f cannot be analytically continued to a neighborhood of 1. $\square$

Exercise 61. Suppose that a non-constant entire function f takes real values on two intersecting lines in the plane. Show that the measure of the angle formed by either line is a rational multiple of π.

PROOF. First, we observe that if f is a function which is real valued on the real axis and on the line $re^{i\theta}$ for $r \in \mathbb{R}$, then f is also real valued on the line $re^{-i\theta}$. Indeed, by the Schwartz reflection principle, we have $f(z) = \overline{f(\overline{z})}$, so $f(re^{-i\theta}) = \overline{f(re^{i\theta})} \in \mathbb{R}$.

Now let us assume that f is the function given in the problem. By translating and rotating the domain of f if necessary, we can assume that the lines intersect at the origin and that one of the lines is the real axis. If f assumes real values on the line $\{re^{i\theta} : r \in \mathbb{R}\}$ for some $0 < \theta < \pi$, then it is enough to show that θ is a rational multiple of π. We will now show by induction that f is real valued on each line $\{re^{in\theta} : r \in \mathbb{R}\}$ for all $n \geq 1$. If this is shown, then since f is real valued on the real axis, and by the first argument of this proof, f must be real valued on the lines $\{re^{in\theta} : r \in \mathbb{R}\}$ for all $n \in \mathbb{Z}$. To this end, the base case $n = 1$ is by assumption. Suppose we have shown f is real valued on the line $\{re^{ik\theta} : r \in \mathbb{R}\}$ for each $1 \leq k \leq n$. Let $g(z) = f(e^{in\theta}z)$. Then g is real valued on the real axis by the induction hypothesis, and also on the line $re^{-i\theta}$. By the first observation we made in this proof, g must also be real valued on the line $re^{i\theta}$. But this means $g(re^{i\theta}) = f(re^{i(n+1)\theta})$ is real valued for $r \in \mathbb{R}$, so the induction is complete.

Finally, to obtain a contradiction, suppose that θ is not a rational multiple of π. Then the set of points $\{e^{in\theta} : n \in \mathbb{Z}\}$ is dense in the unit circle in $\mathbb{C}$. Since f is real valued on each of the lines $re^{in\theta}$ for $n \in \mathbb{Z}$, by continuity it follows that f is real valued on every line $re^{i\phi}$ for any $0 \leq \phi < 2\pi$. But this means f is only real valued, which means f is constant by the CR equations. The contradiction is resolved if θ is a rational multiple of π, and the proof is done. □

EXERCISE 62.

(1) Let f be a complex function which is analytic in the disk $\{|z| < 1\}$ and continuous on $\{|z| \leq 1\}$. Suppose further that f is real valued when $|z| = 1$. Show that f is constant.
(2) Find a non-constant function which is analytic at every point in $\mathbb{C}$ except for a single point on the unit circle $\{|z| = 1\}$, and which is real valued at every other point of the unit circle.

SOLUTION.

(1) Let $\phi(z) = \frac{1+iz}{1-iz}$ for $z \neq -i$. Note that ϕ conformally maps the upper half plane $\{\mathrm{Im} z \geq 0\}$ to the unit disk $\{|z| \leq 1\}$, and moreover $|\phi(z)| = 1$ iff $\mathrm{Im} z = 0$. Thus the function $f \circ \phi$ is continuous on $\{\mathrm{Im} z \geq 0\}$ and analytic on $\{\mathrm{Im} z > 0\}$, and is real valued on the real axis. By the Schwartz reflection principle, $f \circ \phi$ extends to an entire function. But f is continuous on $\{|z| \leq 1\}$ and is thus bounded, so by Liouville's theorem $f \circ \phi$ is constant. Since ϕ is invertible, it follows that f is constant.
(2) Let $f(z) = i\frac{z+1}{z-1}$. If $|z| = 1$ and $z \neq 1$ then

$$f(z) = i\frac{(z+1)(\overline{z}-1)}{|z-1|^2} = i\frac{\overline{z}-z}{|z-1|^2} \in \mathbb{R}$$

since $i(\overline{z} - z) = -2\mathrm{Im} z \in \mathbb{R}$. Clearly f is analytic except at $z = 1$.

□

EXERCISE 63. Let F be a polynomial of degree $d \geq 1$ and let S be its root set. If R is a rational function whose poles are contained in S, then there exists a unique choice of integers $m \leq n$ and polynomials $a_m, \ldots, a_n$ of degree at most d such that

$$R = \sum_{k=m}^{n} a_k F^k.$$

PROOF. First suppose that R is a polynomial of degree $N \geq 0$. We will now prove by induction on $N \geq 0$ that for every polynomial R of degree N, there exists a unique $n \geq 0$ and unique polynomials $a_0, \dots, a_n$ of degree at most d such that

$$R = \sum_{k=0}^{n} a_k F^k.$$

When $N = 0$, it holds R is constant and thus $n = 0$ by the fundamental theorem of algebra. Thus a_0 is a constant which must be identically R, proving the base case. If $N \geq 1$ and the hypothesis has been established for all polynomials of degree strictly less than N, then by the polynomial division algorithm we can uniquely write

$$R = qF + r$$

where q, r are polynomials with $\deg r < d$ and $\deg q < N$. By the induction hypothesis we can uniquely write

$$q = \sum_{k=0}^{n} a_{k+1} F^k$$

for some polynomials $a_1, \dots, a_n$ of degree at most d and $a_n \neq 0$. Putting $a_0 := r$, it follows that

$$R = \sum_{k=0}^{n+1} a_k F^k.$$

To see that this expansion is unique, suppose that $R = \sum_{k=0}^{n'+1} b_k F^k$. Letting $q' = \sum_{k=0}^{n'} b_{k+1} F^k$, we see that

$$(q - q')F + (a_0 - b_0) = 0.$$

By uniqueness of the polynomial division algorithm, $q = q'$ and hence $a_0 = b_0$. However, since $\deg q = \deg q' < N$, by the induction hypothesis $\sum_0^n a_{k+1} F^k = \sum_0^{n'} b_{k+1} F^k$ implies that $n = n'$ and $a_k = b_k$ for each k, completing the induction.

We can now assume that R is a rational function whose poles are in the zero set of F. Then we can find a unique smallest integer m such that $F^m R$ is a polynomial. Then we can uniquely write

$$F^m R = \sum_{k=0}^{n+m} a_k F^k$$

for some integer $n \geq -m$ and polynomials $a_0, \dots, a_{n+m}$ of degree at most d. Dividing through by F^m, we get

$$R = \sum_{k=-m}^{n} a_{k+m} F^k.$$

The expansion is unique since the expansion for $F^m R$ is unique, completing the proof. □

Exercise 64. Suppose that $x_1, \dots, x_n$ elements of a complex vector space X. Let

$$G(x_1, \dots, x_n) = \det \begin{bmatrix} (x_1, x_1) & \cdots & (x_1, x_n) \\ \vdots & \ddots & \vdots \\ (x_n, x_1) & \cdots & (x_n, x_n) \end{bmatrix}.$$

The points $x_1, \dots, x_n$ are linearly dependent if and only if $G(x_1, \dots, x_n) = 0$. Moreover, if $x_1, \dots, x_n$ are linearly independent, then for any $x \in V$,

$$d(x, V)^2 = \frac{G(x_1, \dots, x_n, x)}{G(x_1, \dots, x_n)},$$

where V is the subspace generated by $x_1, \dots, x_n$ and $d(x, V) = \inf\{|x - y| : y \in V\}$.

PROOF. Suppose that the $x_1, \dots, x_n$ are linearly dependent. By re-indexing if necessary, we can assume that $x_n = \sum_{j=1}^{n-1} \lambda_j x_j$. Then the matrix

$$A := \begin{bmatrix} (x_1, x_1) & \cdots & (x_1, x_n) \\ \vdots & \ddots & \vdots \\ (x_n, x_1) & \cdots & (x_n, x_n) \end{bmatrix}$$

is rank deficient, since the n-th column is in the span of the other $n-1$:

$$\begin{bmatrix} (x_1, x_n) \\ \vdots \\ (x_n, x_n) \end{bmatrix} = \sum_{j=1}^{n-1} \lambda_j \begin{bmatrix} (x_1, x_j) \\ \vdots \\ (x_n, x_j) \end{bmatrix}$$

Thus $G(x_1, \dots, x_n) = \det A = 0$. If $G(x_1, \dots, x_n) = 0$, then the matrix A from before is rank deficient, so again by re-indexing if necessary we can assume that

$$\begin{bmatrix} (x_1, x_n) \\ \vdots \\ (x_n, x_n) \end{bmatrix} = \sum_{j=1}^{n-1} \lambda_j \begin{bmatrix} (x_1, x_j) \\ \vdots \\ (x_n, x_j) \end{bmatrix}.$$

This implies

$$\left(x_j, x_n - \sum_{j=1}^{n-1} \lambda_j x_j \right) = 0$$

for each $j = 1, \dots, n$. It follows that

$$\left\| x_n - \sum_{j=1}^{n-1} \lambda_j x_j \right\| = \left(x_n, x_n - \sum_{j=1}^{n-1} \lambda_j x_j \right) - \sum_{j=1}^{n-1} \lambda_j \left(x_j, x_n - \sum_{j=1}^{n-1} \lambda_j x_j \right) = 0,$$

and hence x_n is in the span of the $x_1, \dots, x_{n-1}$.

For the final conclusion, observe that

$$G(x_1, \dots, x_n, x) = \det \begin{bmatrix} (x_1, x_1) & \cdots & (x_1, x_n) & (x_1, Px) \\ \vdots & \ddots & \vdots & \vdots \\ (x_1, Px) & \cdots & (x_n, Px) & \|Px\|^2 + \|P^\perp x\|^2 \end{bmatrix}$$

where Px is the orthogonal projection of x onto V and $P^\perp x = x - Px$. Since $Px \in \operatorname{span}\{x_1, \dots, x_n\}$, by the previous observation we have

$$\det \begin{bmatrix} (x_1, x_1) & \cdots & (x_1, x_n) & (x_1, Px) \\ \vdots & \ddots & \vdots & \vdots \\ (x_1, Px) & \cdots & (x_n, Px) & \|Px\|^2 \end{bmatrix} = 0,$$

and by the co-factor expansion formula for the determinant it holds

$$G(x_1, \dots, x_n, x) = G(x_1, \dots, x_n, Px) + G(x_1, \dots, x_n)\|P^\perp x\|^2 = G(x_1, \dots, x_n)\|P^\perp x\|^2.$$

Thus

$$\frac{G(x_1, \dots, x_n, x)}{G(x_1, \dots, x_n)} = \|P^\perp x\|^2 = d(x, V)^2,$$

as required. $\square$

COROLLARY 65. *Let X be a $m \times n$ matrix where $m \geq n$. Then $X^\top X$ is invertible if and only if the columns of X are linearly independent.*

PROOF. Let $x_1, \ldots, x_n$ be the columns of X. Then $\det X^\top X = G(x_1, \ldots, x_n) \neq 0$ iff the $x_1, \ldots, x_n$ are linearly independent. $\square$

COROLLARY 66. *Let $m \geq n$ be integers and $M_{m\times n}$ be the set of $m \times n$ matrices. Then the set L of rank n matrices is open in $M_{m\times n}$, and moreover there is a continuous mapping $T : L \to M_{n\times m}$ such that $T(A)A = I_n$ for all $A \in L$.*

PROOF. Let $E \subset M_{n\times n}$ be the set of invertible matrices. It is known that E is open in $M_{n\times n}$. Note that $L = \{X \in M_{m\times n} : X^\top X \in E\}$. Since the map $X \mapsto X^\top X$ is continuous, it follows L is open. Let $T(A) = (A^\top A)^{-1} A^\top$. It is easy to check that T is continuous (since $A \mapsto A^{-1}$ is continuous in E) and $T(A)A = I_{n\times n}$ whenever $A \in L$. $\square$

EXERCISE 67. Suppose that $\{p_j\}_{j\geq 1}$ is a sequence of polynomials, all of which are of degree at most $n \geq 1$. Suppose also that $x_0, \ldots, x_n$ is a sequence of distinct points such that

$$\lim_{j\to\infty} p_j(x_k)$$

exists for each $k = 0, \ldots, n$. Show that $\lim_{j\to\infty} p_j(x)$ exists for every $x \in \mathbb{R}$ and that the function $p(x) = \lim_{j\to\infty} p_j(x)$ is a polynomial of degree at most n.

PROOF. Vandermonde. $\square$

EXERCISE 68. Suppose that $z_1, \ldots, z_n$ are points in $\mathbb{C}$. Show that there exists a subset $J \subset \{1, \ldots, n\}$ such that

$$\left|\sum_{j\in J} z_j\right| \geq \frac{1}{4\sqrt{2}} \sum_{j=1}^{n} |z_j|.$$

PROOF. For each $k = 1, 2, 3, 4$ let Q_k be those indices i such that z_i is in the k-th quadrant of $\mathbb{C}$. Note that

$$\sum_{i=1}^{n} |z_i| \leq \sum_{i\in Q_1} |z_i| + \sum_{i\in Q_2} |z_i| + \sum_{i\in Q_3} |z_i| + \sum_{i\in Q_4} |z_i| \leq 4 \max_k \sum_{i\in Q_k} |z_i|.$$

Since both sides of the inequality we seek to prove are invariant under rotations, we can assume that the maximum is achieved when $k = 1$. Moreover, by re-indexing if necessary, we can suppose that $Q_1 = \{1, \ldots, m\}$ for some $1 \leq m \leq n$. Then since $\mathrm{Re} z_i \geq 0$ and $\mathrm{Im} z_i \geq 0$, it holds $|z_i| \leq \mathrm{Re} z_i + \mathrm{Im} z_i$. Moreover, since $\sqrt{a+b} \geq \frac{1}{\sqrt{2}}(\sqrt{a} + \sqrt{b})$ when $a \geq 0$ and $b \geq 0$, it holds

$$\sum_{i=1}^{m} \mathrm{Re} z_i + \sum_{i=1}^{m} \mathrm{Im} z_i \leq \sqrt{2}\sqrt{\left(\sum_{i=1}^{m} \mathrm{Re} z_i\right)^2 + \left(\sum_{i=1}^{m} \mathrm{Im} z_i\right)^2} = \sqrt{2}\left|\sum_{i=1}^{m} z_i\right|.$$

In total, this yields

$$\sum_{i=1}^{n} |z_i| \leq 4\sum_{i=1}^{m} |z_i| \leq 4\sum_{i=1}^{m} \mathrm{Re} z_i + \mathrm{Im} z_i \leq 4\sqrt{2}\left|\sum_{i=1}^{m} z_i\right|,$$

which proves the claim. $\square$

EXERCISE 69. Let p be an odd prime number and $\mathbb{F}_p$ be the field on p letters. Show that there are exactly $\frac{p+1}{2}$ perfect squares in $\mathbb{F}_p$.

PROOF. Let $\mathbb{F}_p^*$ be the group of units in $\mathbb{F}_p$. Define an endomorphism $\phi : \mathbb{F}_p^* \to \mathbb{F}_p^*$ by $\phi(x) = x^2$. Then the number of distinct squares in $\mathbb{F}_p$ is $1 + \#\mathrm{Im}\phi$, where the 1 is added since 0 is a square element not contained in $\mathbb{F}_p^*$. Suppose that $x \in \ker\phi$. Then $x^2 = 1$, so $(x-1)(x+1) = 0$. Since $\mathbb{F}_p$ is an integral domain, this forces either $x = 1$ or $x = p - 1$. Both elements are distinct and in

$\ker T$, so $\# \ker T = 2$. This proves $\#\mathrm{Im}\phi = \frac{p-1}{2}$ since $\#\mathbb{F}_p^* = p-1$. We deduce that the number of distinct perfect squares in $\mathbb{F}_p$ is $1 + \frac{p-1}{2} = \frac{p+1}{2}$. $\square$

Exercise 70. Compute
$$\int_0^\infty \frac{1}{1+x^a}\,dx$$
for each real number $a > 1$.

Proof. Let C_R be the contour which encloses the sector $\{re^{i\theta} : 0 \le r \le R, 0 \le \theta \le 2\pi/a\}$. Integrate $\int_{C_R} \frac{1}{1+z^a}\,dz$ using the residue theorem (the branch of z^a is chosen to not intersect the contour). One then obtains
$$\int_0^\infty \frac{dx}{1+x^a} = \frac{\pi/a}{\sin(\pi/a)}$$
$\square$

Exercise 71. Suppose that f is an non-constant entire function and that there exist $a, b \in \mathbb{C}$ such that $f(az+b) = f(z)$. Then there exists $n \ge 1$ such that $a^n = 1$.

Proof. If $a = 1$, then we are done. Otherwise, assume that $a \ne 1$. Then for every $n \ge 1$, it holds
$$f\left(a^n z + b\frac{a^n - 1}{a-1}\right) = f(z)$$
for all $z \in \mathbb{C}$. If $|a| < 1$ then taking $n \to \infty$ implies that $f(z) = f\left(\frac{b}{1-a}\right)$, which contradicts the fact f is not constant. If $|a| > 1$, fix $k \ge 1$ large enough that $f^{(k)}(0) \ne 0$ (such a k exists otherwise f is constant). Then
$$a^{nk} = f^{(k)}(0)^{-1} f^{(k)}\left(-\frac{b}{a^n}\frac{a^n-1}{a-1}\right)$$
for all $n \ge 1$. Note that
$$\lim_{n\to\infty}\left|f^{(k)}(0)^{-1} f^{(k)}\left(-\frac{b}{a^n}\frac{a^n-1}{a-1}\right)\right| = \left|f^{(k)}(0)^{-1} f^{(k)}\left(\frac{b}{1-a}\right)\right|,$$
but $|a^{nk}| = |a|^{nk} \to \infty$ as $n \to \infty$, a contradiction. This forces $|a| = 1$. Suppose that $a = e^{i\theta\pi}$ for some irrational θ. Then $\{a^n : n \ge 1\}$ is dense in the unit circle, so if $|z| = 1$ we can find a subsequence $\{n_k\}$ such that $a^{n_k} \to z^{-1}$. Then it holds
$$f(z) = f\left(1 + b\frac{z^{-1} - 1}{a-1}\right)$$
for all $|z| = 1$. By analytic continuation the equality holds for all $z \ne 0$. However, this means that f is bounded away from the origin, and since f is bounded near the origin, f must be bounded on $\mathbb{C}$. Thus f is constant, a contradiction. The contradiction is resolved if θ is rational, which means there is $n \ge 1$ such that $a^n = 1$. $\square$

Easier Proof. Suppose $a \ne 1$. Let $c = \frac{b}{1-a}$ so that $ac + b = c$. Then the function $g(z) := f(z+c)$ has $g(az) = f(az+c) = f(az+ac+b) = f(z+c) = g(z)$. Thus $g^{(k)}(0) = a^k g^{(k)}(0)$ for all $k \ge 0$. Since $g^{(k)}(0) \ne 0$ for some $k \ge 1$ large enough (otherwise g hence f is constant), it holds $a^k = 1$ for such k, and the proof is done. $\square$

Exercise 72. Let V be a nonzero finite dimension vector space over $\mathbb{C}$. Show that there do not exist complex square matrices A, B such that $AB - BA = I$. However, show by way of example that this conclusion does not hold if V is infinite dimensional.

PROOF. In the finite dimensional case, $\mathrm{Tr}(AB - BA) = 0 \neq \mathrm{Tr}(I)$ so no such operator exists. In the infinite dimensional setting let V be the vector space of complex valued C^∞ functions on $\mathbb{R}$. Note that $V \neq 0$ by Urysohn's lemma. Let D be the differential operator and

$$Sf(x) = xf(x).$$

Then $DSf(x) = f(x) + xf'(x)$ and $SDf(x) = xf'(x)$, which implies

$$(DS - SD)f(x) = f(x),$$

hence $DS - SD = I$. □

EXERCISE 73. Suppose that X is a compact metric space and $f_1, f_2, \ldots$ are a sequence of real-valued, uniformly bounded, and equicontinuous functions on X. Let

$$g_n = \max\{f_1, \ldots, f_n\}.$$

Show that $\{g_n\}$ converges uniformly.

PROOF. Clearly the sequence $\{g_n\}$ is uniformly bounded since f_n is uniformly bounded. Thus the pointwise limit $g(x) := \lim_{n\to\infty} g_n(x) = \sup_n g_n(x)$ exists for each $x \in X$. Since the supremum of a family of continuous functions is lower semi-continuous, it is enough to show that g is upper semi-continuous. To this end, we need to show that for each $c \in \mathbb{R}$, it holds $\{g < c\}$ is open. Suppose that $g(x) < c$. Fix $\epsilon > 0$ small enough that $g(x)+\epsilon < c$. By equicontinuity, we can find $\delta > 0$ small enough that $d(x, y) < \epsilon$ implies $|f_n(x) - f_n(y)| < \epsilon$ for all $n \geq 1$. Then $f_n(y) < f_n(x) + \epsilon \leq g(x) + \epsilon < c$. It follows $g(y) < c$ whenever $g(x, y) < \delta$, which shows that x is an interior point of $\{g < c\}$. Since x was arbitrary, it follows $\{g < c\}$ is open, finishing the proof. □

COROLLARY 74. *Suppose that $g : [0, 1] \times [0, 1] \to \mathbb{R}$ is a continuous function. Show that*

$$f(x) = \max_{y\in[0,1]} g(x, y)$$

is a continuous function.

PROOF. Taking $r_1, r_2, \ldots$ to be an enumeration of rationals in $[0, 1]$, it holds

$$f(x) = \lim_{n\to\infty} \max\{g(x, r_1), g(x, r_2), \ldots, g(x, r_n)\}.$$

The functions $\{g(\cdot, r_n) : n \geq 1\}$ are uniformly bounded and equicontinuous so by the previous exercise the convergence is uniform and thus f is continuous. □

EXERCISE 75. Let n be a fixed positive integer, and define two $n \times n$ matrices to be equivalent if there is a non-singular real matrix C with $CAC^\top = B$. How many equivalence classes are there?

PROOF. By Sylvester's law of inertia this is just the number of matrices of the form

$$\begin{bmatrix} I_p & & \\ & -I_r & \\ & & 0_z \end{bmatrix}$$

where $p + r + z = n$ for $p, r, z \geq 0$. As usual, I_j denotes the $j \times j$ identity matrix and 0_z is the $z \times z$ zero matrix. So the answer is just $\binom{n+2}{2} = \frac{(n+2)(n+1)}{2}$.

□

EXERCISE 76. Suppose that $g : [0, 1] \times [0, 1] \to \mathbb{R}$ is a continuous function such that for each $x \in [0, 1]$ there is a unique $y_x \in [0, 1]$ such that $g(x, y_x) = \max\{g(x, y) : y \in [0, 1]\}$. Then the mapping $f(x) = y_x$ is continuous.

Proof. Let $h(x) = g(x, f(x))$. As we know, h is continuous. So, if $\{x_n\}$ is a sequence convering to $x \in [0, 1]$, then $h(x_n) \to h(x)$. Fix a subsequence $\{x_{n_k}\}$. Then there is a subsequence $f(x_{n_{k_j}})$ which converges to a limit $y \in [0, 1]$. It holds $h(x_{n_{k_j}}) \to g(x, y) = h(x)$. Since $g(x_n, f(x_n)) \geq g(x_n, z)$ for all $z \in [0, 1]$, it holds $g(x, y) \geq g(x, z)$ for all z, so $y = y_x$. Thus every subsequence of $\{f(x_n)\}$ has a further subsequence converging to y_x, which proves $\lim_n f(x_n) = y_x = f(x)$, hence f is continuous. □

Exercise 77. Suppose that G is a finite group of order pq, where p and q are primes with $p < q$ and $q \neq 1 \mod p$. Then G is cyclic.

Proof. Let P be a Sylow group of order p and Q be a Sylow subgroup of order q. Since $n_p | q$ it holds that $n_p \in \{1, q\}$. We cannot have $n_p = q$ since $n_p = 1 \mod p$, so $n_p = 1$ and hence P is normal. Moreover, since $n_q \in \{1, p\}$ and $n_q = 1 \mod q$,

Thus the map $T : G \to \mathrm{Aut}(P)$ by $T(g) = x \mapsto gxg^{-1}$ is a homomorphism and $\ker T = C_G(P) := \{g \in G : gx = xg\ \forall x \in P\}$. Since P is cyclic it holds $P \leq C_G(P)$ hence $\#C_G(P) \in \{p, pq\}$. $\ker T | p - 1$.

Since $G/C_G(P) \lesssim \mathrm{Aut}(P)$ it holds $\#G/C_G(P) \leq p - 1$, which means that $\#C_G(P) \neq p$ since otherwise $\#G/C_G(P) = q \leq p - 1$. Thus $C_G(P) = G$. Now, since P and Q are both cyclic, there are $x \in P$ and $y \in Q$ such that $P = (x)$ and $Q = (y)$. Since x and y commute and are of co-prime orders in G, the order of xy is pq, hence $G = (xy)$. □

Exercise 78. Let A be an $n \times n$ matrix over a field F and I the $n \times n$ identity matrix. Show that the $2n \times 2n$ matrix

$$\begin{bmatrix} A & I \\ 0 & A \end{bmatrix}$$

is not diagonalizable.

Proof. Let $E = \begin{bmatrix} A & I \\ 0 & A \end{bmatrix}$. For any polynomial p, it holds

$$p(E) = \begin{bmatrix} p(A) & p'(A) \\ 0 & p(A) \end{bmatrix}.$$

So $p(E) = 0$ if and only if $p(A) = 0$ and $p'(A) = 0$. If λ is any eigenvalue of A, this forces $p(\lambda) = 0$ and $p'(\lambda) = 0$. Thus the minimal polynomial of E is not a product of distinct linear factors, since each eigenvalue of A is a double root of the minimal polynomial of E. This proves that E cannot be diagonalized. □

This would have been a useful theorem to know.

Exercise 79. Let H be a normal subgroup of a finite group G and assume that $\#H = p$. Then H is contained in every p-Sylow subgroup of G.

Proof. Let P be a Sylow p-subgroup of G. Since H is normal, P acts on the cosets of H via left multiplication, namely $g(xH) = (gx)H$ for $g \in P$ and a coset xH. We look at the orbit of H under P. Since $(G : H) = p^{n-1}$, by orbit stabilizer we have

$$(P : P_H) \,|\, p^{n-1}.$$

Since $\#P = p^n$, this implies that $P_H := \{g \in P : gH = H\}$ is nontrivial. Fix an element $g \neq 1$ such that $gH = H$. Then $g \in H$, so $g = x^k$ for some $1 \leq k < p$, where $x \in H$ is a generator of H. Choosing $1 \leq r < p$ such that $rk = 1 \mod p$, we have

$$P \ni g^r = x.$$

Since P was an arbitrary Sylow p group, we are done. □

Exercise 80. Let G be a group of order 120, and let H be a subgroup of order 24. Suppose that there is at least one left coset of H (other than H itself) which equals a right coset of H. Prove that H is a normal subgroup of G.

Proof. Let N_H be the normalizer of H, and observe that $5 = (G : H) = (G : N_H)(N_H : H)$. If we can show that $(N_H : H) \neq 1$, since 5 is prime it will hold $(N_H : H) = 5$ and $(G : N_H) = 1$, which means $N_H = G$ and thus H is normal. To this end, a left coset xH and a right coset Hy such that $xH = Hy$. Then $xy^{-1}H = xHx^{-1}$, and thus we can find $h \in H$ such that $xy^{-1}h = 1$. It follows that $xy^{-1} \in H$, and so $x^{-1}H = y^{-1}H = Hx^{-1}$. This proves $x^{-1}Hx = H$, and since $xH \neq H$ we see that $N_H \neq H$ and the proof is done. $\square$

Corollary 81. *The set of real numbers $\{\sqrt{m} - \sqrt{n} : m \geq n \geq 0\}$ is dense in $(0, \infty)$.*

Proof. We claim that

$$\lim_{k\to\infty} \sqrt{[(kx)^2]} - \sqrt{[((k-1)x)^2]} = x.$$

Indeed,

$$\begin{aligned}\sqrt{[(kx)^2]} - \sqrt{[((k-1)x)^2]} - x &= \sqrt{[(kx)^2]} - kx - \left(\sqrt{[((k-1)x)^2]} - (k-1)x\right)\\ &= \frac{[(kx)^2] - (kx)^2}{\sqrt{[(kx)^2]} + kx} - \frac{[((k-1)x)^2] - ((k-1)x)^2}{\sqrt{[((k-1)x)^2]} + (k-1)x}.\end{aligned}$$

Since $0 \leq [(kx)^2] - (kx)^2 \leq 1$ we have $0 \leq \frac{[(kx)^2]-(kx)^2}{\sqrt{[(kx)^2]}+kx} \leq \frac{1}{\sqrt{[(kx)^2]}+kx}$ hence $\frac{[(kx)^2]-(kx)^2}{\sqrt{[(kx)^2]}+kx} \to 0$ as $k \to \infty$, and thus also $\frac{[((k-1)x)^2]-((k-1)x)^2}{\sqrt{[((k-1)x)^2]}+(k-1)x} \to 0$ as $k \to \infty$. Therefore

$$\lim_{k\to\infty} \sqrt{[(kx)^2]} - \sqrt{[((k-1)x)^2]} - x = 0,$$

as required. $\square$

Exercise 82 (Putnam 2017 A3). Let f and g be positive continuous functions on $[0, 1]$. Suppose $\int_0^1 f = \int_0^1 g$ but $f \neq g$. Given $n \in \mathbb{Z}$, define

$$I_n = \int_0^1 \frac{f(x)^{n+1}}{g(x)^n}\, dx.$$

Show that the sequence $I_{-1}, I_0, I_1, \ldots$ is increasing and $\lim_{n\to\infty} I_n = \infty$.

Proof. By definition $I_{-1} = I_0$, so we prove that $I_{n-1} \leq I_n$ by induction on $n \geq 0$. Suppose we have shown that $I_{-1} \leq I_0 \leq \cdots \leq I_{n-1} \leq I_n$. Then we have

$$I_n = \int_0^1 \frac{f(x)^{n+1}}{g(x)^n}\, dx = \int_0^1 \frac{f(x)^{\frac{n}{2}}}{g(x)^{\frac{n-1}{2}}} \frac{f(x)^{\frac{n+2}{2}}}{g(x)^{\frac{n+1}{2}}}\, dx \leq \sqrt{I_{n-1} I_{n+1}}$$

by the Cauchy-Schwarz inequality. But, we know $I_{n-1} \leq I_n$, and by rearranging we get $\sqrt{I_n} \leq \sqrt{I_{n+1}}$ and therefore $I_n \leq I_{n+1}$.

Finally, to show that $I_n \to \infty$, note that if $f(x) \leq g(x)$ for all $x \in (0, 1)$ then the condition $\int f = \int g$ implies $f = g$, so there exists $x_0 \in (0, 1)$ such that $\frac{f(x_0)}{g(x_0)} > 1$. Since f/g is continuous, we can therefore find an interval (a, b) containing x_0 and $\alpha > 1$ such that $f(x)/g(x) \geq \alpha$ on (a, b). Note that this implies $f(x) > 0$ on (a, b) so that $\int_a^b f(x)\, dx > 0$ and moreover

$$I_n \geq \int_a^b f(x) \left(\frac{f(x)}{g(x)}\right)^n dx \geq \alpha^n \int_a^b f(x)\, dx$$

for all $n \geq 1$. But $n \geq 1$, and therefore $I_n \to \infty$ as $n \to \infty$. $\square$

EXERCISE 83. Let $a \in (0,1)$ and suppose $f : \mathbb{R} \to \mathbb{R}$ satisfies

$$\lim_{x\to 0} f(x) = 0, \quad \lim_{x\to 0} \frac{f(x) - f(ax)}{x} = 0.$$

Then $\lim_{x\to 0} \frac{f(x)}{x} = 0$.

PROOF. Fix $\epsilon > 0$ and let $\delta > 0$ be small enough that

$$\left| \frac{f(x) - f(ax)}{x} \right| < \epsilon, \quad |x| < \delta.$$

Then since $0 \le a < 1$, it holds $\left| \frac{f(a^k x) - f(a^{k+1}x)}{x} \right| < \epsilon a^k$ for any $|x| < \delta$ and $k \ge 1$. Thus $|x| < \delta$ implies

$$\left| \frac{f(x)}{x} \right| \le \left| \frac{f(a^{N+1}x)}{x} \right| + \sum_{k=0}^{N} \left| \frac{f(a^k x) - f(a^{k+1}x)}{x} \right| < \left| \frac{f(a^{N+1}x)}{x} \right| + \epsilon \frac{1 - a^{N+1}}{1 - a}$$

for all $N \ge 1$. Using the fact $f(x) \to 0$ as $x \to 0$, by taking $N \to \infty$ we retrieve

$$\left| \frac{f(x)}{x} \right| \le \frac{\epsilon}{1-a}$$

for $|x| < \delta$. Since $\epsilon > 0$ was arbitrary, we are done. □

EXERCISE 84. Suppose that $f(z) = \sum_{n\ge 0} c_n z^n$ with radius of convergence $R > 0$. If $s_k(z) = \sum_{n=0}^{k} c_n z^n$, then if $|z| < r < R$ it holds

$$|f(z) - s_k(z)| \le C_r \frac{\left|\frac{z}{r}\right|^{k+1}}{1 - |z/r|},$$

where $C_r = \max_{|w|=r} |f(w)|$. In particular,

$$\sum_{k=0}^{\infty} |f(z) - s_k(z)| \le \max_{|w|=|z|} |f(w)|$$

so that

$$\sum_{k=0}^{\infty} f(z) - s_k(z)$$

converges absolutely and uniformly on compact subsets of $\{|z| < R\}$.

PROOF. By Cauchy's formula, we know that

$$c_n = \frac{1}{2\pi i} \int_{|w|=r} \frac{f(w)}{w^{n+1}}\, dw.$$

Therefore,

$$f(z) = \frac{1}{2\pi i} \int_{|w|=r} \sum_{n=0}^{\infty} \left(\frac{z}{w}\right)^n f(w) \frac{dw}{w}, \quad s_k(z) = \frac{1}{2\pi i} \int_{|w|=r} \sum_{n=0}^{k} \left(\frac{z}{w}\right)^n f(w) \frac{dw}{w}$$

and in particular

$$f(z) - s_k(z) = \frac{1}{2\pi i} \int_{|w|=r} \sum_{n=k+1}^{\infty} \left(\frac{z}{w}\right)^n f(w) \frac{dw}{w} = \frac{1}{2\pi i} \int_{|w|=r} \left(\frac{z}{w}\right)^{k+1} \frac{f(w)}{w - z}\, dw.$$

Applying the triangle inequality, we get

$$|f(z) - s_k(z)| \le \left|\frac{z}{r}\right|^{k+1} \frac{1}{2\pi} \int_0^{2\pi} \frac{r|f(w)|}{|re^{i\theta} - z|}\, d\theta \le \left|\frac{z}{r}\right|^{k+1} \frac{rC_r}{r - |z|},$$

as required. $\square$

Exercise 85. Let p be a degree $n \geq 2$ polynomial with n distinct roots $z_1, \dots, z_n$. Then

$$\sum_{j=1}^{n} \frac{z_j^{n-1}}{p'(z_j)} = \alpha^{-1},$$

where α is the leading coefficient of p, and for any $2 \leq k \leq n$,

$$\sum_{j=1}^{n} \frac{z_j^{n-k}}{p'(z_j)} = 0.$$

Proof. Let $R > 0$ be large enough that $|z_j| < R$ for each $1 \leq j \leq n$. We evaluate the integral

$$I_R = \frac{1}{2\pi i} \int_{|z|=R} \frac{z^{n-k}}{p(z)}\, dz.$$

On the one hand, we can write

$$\frac{1}{p(z)} = \sum_{j=1}^{n} \frac{1}{p'(z_j)} \frac{1}{z - z_j},$$

so that

$$I_R = \sum_{j=1}^{n} \frac{1}{2\pi i\, p'(z_j)} \int_{|z|=R} \frac{z^{n-k}}{z - z_j}\, dz = \sum_{j=1}^{n} \frac{z_j^{n-k}}{p'(z_j)}.$$

On the one hand, by parameterizing I_R, we see that

$$I_R = \frac{1}{2\pi} \int_0^{2\pi} \frac{R^{n-k+1} e^{i(n-k+1)\theta}}{p(Re^{i\theta})}\, d\theta.$$

If $k \geq 2$, then I_R decays like R^{1-k} as $R \to \infty$. Therefore $I_R \to 0$ as $R \to \infty$, and in particular,

$$\sum_{j=1}^{n} \frac{z_j^{n-k}}{p'(z_j)} = 0.$$

If $k = 1$, then

$$\lim_{R\to\infty} \frac{1}{2\pi} \int_0^{2\pi} \frac{R^n e^{in\theta}}{p(Re^{i\theta})}\, d\theta = \alpha^{-1}.$$

Thus we have

$$\sum_{j=1}^{n} \frac{z_j^{n-1}}{p'(z_j)} = \alpha^{-1},$$

and we are done. $\square$

Corollary 86. *Let $p(z) = 1 + z + \cdots + \frac{z^n}{n!}$. Then p has n distinct roots $z_1, \dots, z_n$, and*

$$\sum_{j=1}^{n} z_j^{-1} = -1, \quad \sum_{j=1}^{n} z_j^{-k} = 0, \quad 2 \leq k \leq n.$$

Proof. Since $p(z) = p'(z) + \frac{z^n}{n!}$, it holds $p'(z_j) = -\frac{z_j^n}{n!}$, and since $z_j \neq 0$ we see that $p'(z_j) \neq 0$. Thus p has no repeated roots. By the preceding exercise,

$$n! = \sum_{j=1}^{n} \frac{z_j^{n-1}}{p'(z_j)} = -n! \sum_{j=1}^{n} z_j^{-1}, \quad 0 = \sum_{j=1}^{n} \frac{z_j^{n-k}}{p'(z_j)} = -n! \sum_{j=1}^{n} z_j^{-k}.$$

This is the stated claim. $\square$

Exercise 87. For each non-negative integer j, show that

$$S(n) = \sum_{k=0}^{n} k^j$$

is a degree $j+1$ polynomial in n. *Hint: Show that the polynomials $(x+1)^{k+1} - x^{k+1}$ for $k \geq 0$ are a basis in the set of polynomials.*

Proof. Note that linear combinations of polynomials are again polynomials, so it is enough to show that $\sum_{k=0}^{n} b(k)$ is a degree at most $j+1$ polynomial in n for each polynomial b belonging to a basis of degree j polynomials. For each $0 \leq i \leq j$, let

$$b_i(x) = (x+1)^{i+1} - x^{i+1}.$$

Note that $b_0, \dots, b_j$ form a basis of the polynomials of degree at most j. Indeed, if $c_0, \dots, c_j$ are such that

$$c_0 b_0 + \cdots + c_j b_j = 0,$$

then $c_j = 0$ since b_j is the only polynomial with a degree j term, and by induction it follows that $c_{j-1} = \cdots = c_0 = 0$. Since the dimension of the space of degree j polynomials is $j+1$, it follows that $b_0, \dots, b_j$ are a basis. On the other hand, observe that

$$\sum_{k=0}^{n} b_i(k) = \sum_{k=0}^{n} (k+1)^{i+1} - k^{i+1} = (n+1)^{i+1}$$

is a degree at most $j+1$ polynomial for each $0 \leq i \leq j$. To see it is degree exactly $j+1$, we observe that

$$\frac{1}{j+1} n^{j+1} = \int_0^n x^j \, dx \leq S(n) \leq \int_0^{n+1} x^j \, dx = \frac{1}{j+1}(n+1)^{j+1},$$

and the only way for this estimate to hold for all n is if it is degree exactly $j+1$. □

Corollary 88. *Let S be the polynomial defined by $S(n) = \sum_{k=0}^{n} k^m$. If $m \geq 2$, then S is divisible by $x(x+1)$ and has leading coefficient $\frac{1}{m+1}$.*

Proof. We can find constants $c_0, \dots, c_m$ such that $x^m = c_0 + c_1((x+1)^2 - x^2) + \cdots + c_m((x+1)^{m+1} - x^{m+1})$, and since $(x+1)^{m+1} - x^{m+1}$ is the only polynomial which contains an x^m term it follows that $c_m \binom{m+1}{m} x^m = x^m$, hence $c_m = \frac{1}{m+1}$. Since $S(0) = 0$ it follows S is divisible by x. Then we have

$$\begin{aligned} S(n) &= \sum_{k=0}^{n} c_0 + c_1((x+1)^2 - x^2) + \cdots + c_m((x+1)^{m+1} - x^{m+1}) \\ &= c_0(n+1) + \cdots + c_m(n+1)^{m+1}. \end{aligned}$$

Thus the leading coefficient of $S(n)$ is $c_m = \frac{1}{m+1}$. Moreover, $S(n)/(n+1)$ is a polynomial in n for each $n \geq 0$, and therefore S is divisible by $x+1$. It is divisible by x since $S(0) = 0$, and we are done. □

Exercise 89. Suppose that f is an entire function on the Riemann sphere such that $f(\infty) = \infty$. Then f is a polynomial.

Proof. Let $g(z) = f(1/z)$. If the singularity of g at $z = 0$ is essential, then by Casorati-Weierstrass we can find a sequence of points $x_1, x_2, \dots$ tending to 0 such that $|g(x_n)| < 1$ for all n, contradicting the fact that $f(\infty) = \infty$. Therefore g has a pole at 0, let n be its order. Write $f(z) = \sum_{k \geq 0} a_k z^k$, and consider the function

$$\frac{f - \sum_{k=0}^{n} a_k z^k}{z^{n+1}}.$$

Since $f(z)/z^n$ tends to a finite limit as $|z| \to \infty$, it holds $f(z)/z^{n+1} \to 0$ as $n \to \infty$, and therefore $\frac{f(z)-\sum_{k=0}^n a_k z^k}{z^{n+1}} \to 0$ as $z \to \infty$. However, since

$$\lim_{z\to 0} \frac{f(z) - \sum_{k=0}^n a_k z^k}{z^{n+1}} = a_{n+1},$$

it follows that $\frac{f-\sum_{k=0}^n a_k z^k}{z^{n+1}}$ extends to an entire function which tends to 0 as $|z| \to \infty$. By Liouville's theorem, $\frac{f-\sum_{k=0}^n a_k z^k}{z^{n+1}} = 0$ for all z, hence $f(z) = \sum_{k=0}^n a_k z^k$ and thus f is a polynomial. □

Corollary 90. *Every conformal map $f : \mathbb{C} \to \mathbb{C}$ is non-constant affine function.*

Proof. Let $g = f^{-1}$. Fix a sequence $z_n \in \mathbb{C}$ which tends to ∞ as $n \to \infty$. Then $z_n = g(f(z_n))$. If $|f(z_{n_k})| \le R$ for some subsequence $\{z_{n_k}\}$, then $g(f(z_{n_k}))$ is bounded since g is holomorphic. This implies $|z_{n_k}|$ is bounded, a contradiction. Therefore $f(z_n) \to \infty$ as $n \to \infty$, hence $f(\infty) = \infty$. By the preceding lemma, f is a polynomial, and by the fundamental theorem of algebra we deduce that f is linear. □

Corollary 91. *Every holomorphic automorphism of the Riemann sphere is a fractional linear transformation.*

Proof. If $f(\infty) = \infty$ then by the fundamental theorem of algebra f is a linear polynomial. Otherwise, let $g = \frac{1}{f-f(\infty)}$. Then g is a holomorphic automorphism of the sphere with $g(\infty) = \infty$, and g is a linear polynomial. Rearranging, we get $f(z) = f(\infty) + \frac{1}{cz+d}$, which implies f is a fractional linear transformation. □

Exercise 92. Let $0 < r_1 < r_2$ and suppose that f is a holomorphic function on the annulus $\{r_1 < |z| < r_2\}$. Then

$$f = g(z) + b(z)$$

, where g is holomorphic on $|z| < r_2$ and b is holomorphic on $|z| > r_1$. Moreover, the choices of g and b are unique up to the addition of an entire function.

Proof. Fix $r_1 < r < r_2$. For $n \ge 0$, let

$$a_n = \frac{1}{2\pi i} \int_{|w|=r} \frac{f(w)}{w^{n+1}}\, dw.$$

Note that a_n is independent of the choice of r since f is analytic in the annulus. Define a function g by

$$g(z) = \sum_{n\ge 0} a_n z^n.$$

Note that the series defining g converges absolutely whenever $|z| < r_2$. Indeed, if we fix $|z| < r < r_2$, it holds

$$\sum_{n=0}^m a_n z^n = \frac{1}{2\pi i} \int_{|w|=r} \sum_{n=0}^m \left(\frac{z}{w}\right)^n f(w) \frac{dw}{w},$$

and since $|z/w| < 1$ it holds $\sum_{n=0}^m (z/w)^n$ converges absolutely and uniformly in w, from which it follows that $\sum_{n=0}^m a_n z^n$ converges as $m \to \infty$. Thus g is a well-defined analytic function on $|z| < r_2$. For $n \ge 1$, let

$$b_n = \frac{1}{2\pi i} \int_{|w|=r} w^{n-1} f(w)\, dw,$$

and define

$$b(z) = \sum_{n\ge 1} b_n z^{-n}.$$

Note that the choice of b_n does not depend on $r_1 < r < r_2$, so for a given z in the annulus and choosing $r_1 < r < |z|$ we see that b converges to a holomorphic function on $|z| > r_1$. We just need to show that

$$f(z) = g(z) + b(z).$$

To this end, we show that

$$f\left(\frac{1}{z}\right) - g\left(\frac{1}{z}\right) = \sum_{n\geq 1} b_n z^n$$

if $r_2^{-1} < |z| < r_1^{-1}$. Note that

$$\frac{1}{2\pi i}\int_{|w|=r} \frac{f(w^{-1}) - g(w^{-1})}{w}\, dw = -\frac{1}{2\pi i}\int_{|w|=1/r} \frac{f(w) - g(w)}{w}\, dw = 0$$

if $r_2^{-1} < r < r_1^{-1}$ by construction of g, and that for $n \geq 1$ it holds

$$\frac{1}{2\pi i}\int_{|w|=r} \frac{f(w^{-1}) - g(w^{-1})}{w^{n+1}}\, dw = -\frac{1}{2\pi i}\int_{|w|=1/r} w^{n-1}(f(w) - g(w))\, dw = b_n$$

by definition of b_n and since g is analytic on $|z| < r_2$. This proves $f(1/z) - g(1/z) = \sum_{n\geq 1} b_n z^n$, and thus $f(z) = g(z) + b(z)$.

For uniqueness up to the addition of an entire function, suppose that $f = g' + b'$ where g' is analytic in $|z| < r_2$ and b' is analytic in $|z| > r_1$. Then $H := g - g' = b - b'$ is entire since $g - g'$ is analytic on $|z| < r_2$ and on $|z| > r_1$. Thus $g = g' + H$ and $b = b' + H$, and we are done. $\square$

Corollary 93. *There exist constants $\{a_n\}$ for $n \in \mathbb{Z}$ such that*

$$f(z) = \sum_{n\in\mathbb{Z}} a_n z^n,$$

and the convergence is uniform on compact subsets of $\{r_1 < |z| < r_2\}$.

Exercise 94. If f is holomorphic in the annulus $\{r_1 < |z| < r_2\}$ and non-vanishing, then there exists $n \geq 0$ and a holomorphic function g on the annulus such that $f(z) = z^n e^{g(z)}$.

Proof. Let n be the winding number of f about the origin, and let $u(z) = f(z)/z^n$. Then u is non-vanishing on the annulus and has winding number 0, and thus u'/u has a holomorphic primitive g. Using the standard arguments one deduces that $u = e^g$, and therefore $f(z) = z^n e^{g(z)}$. $\square$

I just wanted to review a basic fact about complex functions.

Exercise 95. Let Ω be a connected open set, and suppose that f is holomorphic in Ω. If $E := \{x \in \Omega : f(z) = 0\}$ has a limit point in Ω, then $f = 0$. In other words, either $f = 0$ or zeros of a holomorphic function are isolated.

Proof. Since E is closed and Ω is connected, we just need to show that E is open. Let $z_0 \in \Omega$ be a limit point of E. Since f is holomorphic in Ω and $f(z_0) = 0$, for all z close enough to z_0 we can write

$$f(z) = \sum_{n=1}^{\infty} a_k (z - z_0)^k.$$

To obtain a contradiction, suppose that f is not identically zero in any neighborhood of z_0. Then there exists a smallest integer $k \geq 1$ such that $a_k \neq 0$, and then it holds $f(z) = (z - z_0)^k h(z)$ for a holomorphic function h such that $h(z_0) = a_k \neq 0$. Thus there is $\delta > 0$ such that $|h(z)| > 0$ if $|z - z_0| < \delta$. Since z_0 is a limit point, there is a sequence $z_1, z_2, \dots \in E$ of distinct points with $z_n \to z_0$ as $n \to \infty$. As soon as $|z_0 - z_n| < \delta$, we have $0 = f(z_n) = (z_n - z_0)^n h(z_n)$, and since $z_n \neq z_0$ it thus holds $h(z_n) = 0$, a contradiction. Therefore there exists a neighborhood of z_0 such that $f(z) = 0$ in this neighborhood and we conclude $E = \Omega$. $\square$

Exercise 96. Suppose that Ω is an open set, $z_0 \in \Omega$, and f is holomorphic function in $\Omega \setminus \{z_0\}$. Suppose further that f is bounded in a neighborhood of z_0. Then f can be extended to a holomorphic function on Ω.

Proof. Let $g(z_0) = 0$ and if $z \neq z_0$ let $g(z) = (z - z_0)f(z)$. We claim that g is holomorphic in Ω. This is verified from the easily fact the integral of g about any closed loop about z_0 vanishes. If a closed loop runs through z_0, approximate it by a loop which does not contain z_0.

In this fashion one sees that

$$g(z) = \sum_{n \geq 1} a_{n-1}(z - z_0)^n$$

in a neighborhood of z_0 for some coefficients $a_0, a_1 \ldots$, and this shows that

$$f(z) = \sum_{n \geq 0} a_n (z - z_0)^n$$

in that same neighborhood. Thus f is analytic at z_0. □

Revisiting some old problem for practice...

Exercise 97. All non-constant holomorphic functions f on the unit disk satisfying $f(z^k) = f(z)^k$ for all integers $k \geq 1$ of the form $f(z) = z^n$ for some $n \geq 1$.

Proof. First, if f is constant, then clearly $f \in \{0, 1\}$. Next, if $|z| < 1$, then since $f(z^k) = f(z)^k$ we can take $k \to \infty$ to see

$$\lim_{k \to \infty} f(z)^k = f(0).$$

The only way for the limit to exist for all z is if $|f(z)| \leq 1$ for all z. Therefore, by the maximum modulus principle, if $|f(z)| = 1$ for any $|z| < 1$, then f is the constant function 1. Therefore, $|f(z)| < 1$ for all $|z| < 1$, so that $f(0) = 0$. Let n be the order of the zero. Then $g(z) := f(z)/z^n$ can be extended to a holomorphic function on the unit disk with $g(0) \neq 0$ and also satisfying $g(z^k) = g(z)^k$ for all $k \geq 1$. By repeating our previous analysis with g instead of f, we see that $g(0) \neq 0$ implies $|g(z)| = 1$ for some z, hence g is the constant function 1. Thus we have $f(z) = z^n$, and we are done. □

Exercise 98. Suppose that f is a non-vanishing holomorphic function on a simply connected domain Ω. Then there exists a holomorphic function g such that $f = e^g$.

Proof. Fix $z_0 \in \Omega$. Since f is non-vanishing we can choose a holomorphic primitive g of f'/f such that $g(z_0) = \log f(z_0)$, where the log is chosen on an arbitrary branch. We need to show that $f = e^g$. Let $h = f - e^g$. Then $h' = \frac{f'}{f} h$, hence $h'(z_0) = 0$ since $h(z_0) = 0$. By the Leibniz formula,

$$h^{(n+1)}(z_0) = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{f'}{f}\right)^{(n-k)} (z_0) h^{(k)}(z_0),$$

and by induction it follows that $h^{(n)}(z_0) = 0$ for all $n \geq 0$. But since h is analytic this forces $h(z) = 0$ in Ω, and therefore $f = e^g$. □

Exercise 99. Suppose that $f : [a, b] \to \mathbb{R}$ is a C^1 function. Then

$$\sup_{x \in [a,b]} |f(x)| \leq \frac{1}{b - a} \int_a^b |f(x)| \, dx + \int_a^b |f'(x)| \, dx.$$

Proof. By the mean value theorem there exists $y \in (a, b)$ such that $f(y) = \frac{1}{b-a} \int_a^b f(t) \, dt$. Then

$$f(x) = \frac{1}{b - a} \int_a^b f(t) \, dt + \int_y^x f'(t) \, dt.$$

By the triangle inequality, it follows

$$|f(x)| \le \frac{1}{b-a}\int_a^b |f(t)|\,dt + \int_a^b |f'(t)|\,dt$$

and we are done. ☐

The above is actually a pretty crude inequality since one can easily strengthen it to $\sup|f| \le \inf|f| + \int_a^b |f'|$.

Exercise 100. Let H be a proper subgroup of a finite group G. Then $G \neq \bigcup_{x\in G} xHx^{-1}$.

Proof. The number m of distinct conjugacy classes xHx^{-1} is given by $\#G/\#\{x : xHx^{-1} = H\}$. If $m = 1$ then $H = G$, a contradiction. Thus $m \ge 2$. Let $x_1, \dots, x_m$ be representatives from each distinct conjugacy class. Then

$$\#\bigcup_{x\in G} xHx^{-1} \le \sum_{i=1}^{m} \#H = m\#H.$$

Since $\frac{\#H}{\#\{x:xHx^{-1}=H\}} < 1$, it holds $m\#H < \#G$, and we are done. ☐

Exercise 101. Let F be a finite field and n be a positive integer. Show that there exist $n\times n$ matrices A, B with entries in F such that $AB - BA = I$ if and only if the characteristic of F divides n.

Proof. Let p be the characteristic of F. If such matrices exist, then $\mathrm{Tr}(AB - BA) = 0 = n$, so $p \,|\, n$. Conversely, let $V = F[x]/(x^n)$. Clearly V is n dimensional, and define operators A and B by $Af(x) = f'(x)$ and $Bf(x) = xf(x)$. By the product rule and the fact $p|n$, it holds $AB - BA = I$, and we are done. ☐

Exercise 102. Let F be an infinite field and let K be an extension of F. If A and B are two matrices with entries in F which are similar over K, then they are similar over F.

Proof. Let X be an invertible matrix such that $AX = XB$ and $\mathcal{B}$ be a basis of K/F. Let $\alpha_1, \dots, \alpha_m \in \mathcal{B}$ be chosen such that every entry in X can be written as a linear combination of the α_i. Then we can write $X = X_1\alpha_1 + \cdots X_m\alpha_m$ where the X_i have coefficients in F. It follows that $AX_i = X_iB$ for each i, hence

$$A(c_1X_1 + \cdots + c_mX_m) = B(c_1X_1 + \cdots + c_mX_m), \quad c_1, \dots, c_m \in K.$$

If we can find $c_1, \dots, c_m \in F$ such that $c_1X_1 + \cdots + c_mX_m$ is invertible, then we are done. To this end, since $p(c_1, \dots, c_m) := \det[c_1X_1 + \cdots + c_mX_m]$ is a polynomial in $c_1, \dots, c_m$ and is nonzero when $(c_1, \dots, c_m) = (\alpha_1, \dots, \alpha_m)$, and since F is infinite, there exists $c_1, \dots, c_m \in F$ such that $p(c_1, \dots, c_m) \neq 0$. Thus $c_1X_1 + \cdots + c_mX_m$ is invertible, and we are done. ☐

Exercise 103. Let V be a finite dimensional vector space over an algebraically closed field and $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of a nontrivial linear map $A : V \to V$. Let p_A be the minimal polynomial of A and χ_A be the characteristic polynomial of A. If m_i is the algebraic multiplicity of λ_i as a root of p_A, then

$$V = \bigoplus_{i=1}^{k} \ker(A - \lambda_i)^{m_i}.$$

Moreover, the dimension of $\ker(A - \lambda_i)^{m_i}$ is the algebraic multiplicity of λ_i as a root of χ_A

Lemma 104. *Continuing with the notation of the exercise, suppose that $p_A = p_1p_2$, where p_1 and p_2 are co-prime and both are degree at least 1. Then $V = \ker p_1(A) \oplus \ker p_2(A)$.*

Proof. Let $V_1 = \ker p_2(A)$ and $V_2 = \ker p_1(A)$. Note that since p is the minimal polynomial it holds $V_1 \notin \{0, V\}$ and $V_2 \notin \{0, V\}$, so the spaces V_1 and V_2 are proper. Also, clearly $p_1(A)V_1 = 0$ and $p_2(A)V_2 = 0$. Let $U = V_1 \cap V_2$ and suppose that q is the minimal polynomial of A on U. Then $(fp_1 + gp_2)(A)|_U = 0$ for any polynomials f and g, hence $q|(fp_1 + gp_2)$. But p_1 and p_2 are co-prime, so by Bezout's lemma it holds $q|1$. But this means q is constant, hence $U = 0$. To show that $V = V_1 \oplus V_2$, apply Bezout's lemma once more to find polynomials q_1 and q_2 such that $q_1p_1 + q_2p_2 = 1$. Then $q_1(A)p_1(A) + q_2(A)p_2(A) = I$, so that $q_1(A)p_1(A)x + q_2(A)p_2(A)x = x$ for any $x \in V$. Since $p_2(A)(q_1(A)p_1(A))x = q_1(A)p(A)x = 0$ and $p_1(A)(q_2(A)p_2(A)x) = q_2(A)p(A)x = 0$, this proves $x \in V_1 \oplus V_2$, completing the proof. □

Solution to the Exercise. Write $p_A(x) = (x - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k}$. By the lemma and induction, it follows that we can write

$$V = \bigoplus_{i=1}^{k} \ker(A - \lambda_i)^{m_i}.$$

Note that $\ker(A - \lambda_i)^{m_i}$ is an A-invariant subspace, and on this subspace the minimal polynomial of A only has one root, namely λ_i. So, if χ_i is the characteristic polynomial of A on $\ker(A - \lambda_i)^{m_i}$, then $\chi_i(x) = (x - \lambda_i)^{\dim \ker(A - \lambda_i)^{m_i}}$. Thus $\chi_A(x) = (x - \lambda_i)^{\dim \ker(A - \lambda_1)^{m_1}} \cdots (x - \lambda_k)^{\dim \ker(A - \lambda_k)^{m_k}}$, proving that $\dim \ker(A - \lambda_i)^{m_i}$ is the multiplicity of λ_i as a root of χ_A. □

Corollary 105 (Cayley-Hamilton). *Suppose that A is a square matrix with characteristic polynomial χ. Then $\chi(A) = 0$.*

Proof. Denote by p the minimal polynomial of A. If $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of A, we can write

$$V = \bigoplus_{i=1}^{k} \ker(A - \lambda_i)^{m_i},$$

where m_i is the multiplicity of λ_i as a root of p. First, we show that the minimal polynomial p_i of A restricted to $\ker(A - \lambda_i)^{m_i}$ is $(x - \lambda_i)^{m_i}$. Clearly it divides $(x - \lambda_i)^{m_i}$. If $q = p_1 \cdots p_k$, then $q(A) = 0$ so $p \,|\, q$. Since the p_i are mutually co-prime it follows that $(x - \lambda_i)^{m_i}$ divides p_i, hence $p_i(x) = (x - \lambda_i)^{m_i}$. Since p_i is minimal, there exists $x \in \ker(A - \lambda_i)^{m_i}$ such that $(A - \lambda_i)^{m_i - 1}x \neq 0$, and so the vectors $x, (A - \lambda_i)x, \ldots, (A - \lambda_i)^{m_i - 1}x$ are linearly independent. Thus $\dim \ker(A - \lambda_i)^{m_i} \geq m_i$, and in particular $(A - \lambda_i)^{\dim \ker(A - \lambda_i)^{m_i}} = 0$. Thus $\chi(A) = (A - \lambda_i)^{\dim \ker(A - \lambda_1)^{m_1}} \cdots (A - \lambda_k)^{\dim \ker(A - \lambda_k)^{m_k}} = 0$, as required. □

Corollary 106 (Diagonalizable Matrices). *A is diagonalizable if and only if the minimal polynomial of A is the product of distinct linear factors.*

Lemma 107. *Suppose that $V = \bigoplus_{i=1}^{k} W_i$, where each W_i is an invariant subspace. Let p_i be the minimal polynomial of $A|_{W_i}$. Then the minimal polynomial p of A is the least common multiple of $p_1, \ldots, p_k$.*

Proof. Let q be the least common multiple of $p_1, \ldots, p_k$. Since for each i there is a polynomial r_i such that $q = r_ip_i$, it holds $q(A)|_{W_i} = r_i(A)p_i(A)|_{W_i} = 0$ for each i, thus $q(A) = 0$. This proves $p|q$. On the other hand, since $p(A)|_{W_i} = 0$, it holds $p_i|p$ for all i and thus $q|p$. Since p and q are both monic, we deduce $p = q$, as required. □

Proof of the Diagonalization Theorem. Suppose that A is diagonalizable. Then there exists a basis of $v_1, \ldots, v_n$ consisting of eigenvectors of A. Consider the n eigenspaces $W_i = \operatorname{span} v_i$. Then $V = \bigoplus_{i=1}^{n} W_i$, and the minimal polynomial of A on each W_i is linear, hence by the lemma the minimal polynomial of A is a product of distinct linear factors.

For the converse, if $p(x) = (x - \lambda_1) \cdots (x - \lambda_k)$ for distinct $\lambda_1, \dots, \lambda_k$, then it holds

$$V = \bigoplus_{i=1}^{k} \ker(A - \lambda_i).$$

Thus V has a basis consisting of eigenvectors of A, proving A is diagonalizable. $\square$

Exercise 108. Suppose that V is a finite dimensional vector space over a field F and $A : V \to V$ is a diagonalizable map. If $W \subset V$ is a subspace satisfying $AW \subset W$, then $A|_W$ is diagonalizable over W.

Proof. Let q be the minimal polynomial of $A|_W$ and p be the minimal polynomial of A. Since $p(A)|_W = 0$ it holds $q|p$, and since p is a product of distinct linear factors it follows that q is a product of distinct linear factors. Thus $A|_W$ is diagonalizable. $\square$

Exercise 109 (A Master Theorem About Matrix Decompositions). Let K be an algebraically closed field, and let and V be a finite dimensional vector space over K. If $A : V \to V$ is a nonzero linear map, then

$$V = \bigoplus_{i=1}^{k} W_i,$$

where each W_i is an irreducible $A-$invariant subspace, meaning that $AW_i \subset W_i$ and each W_i cannot be decomposed into a direct sum of proper A-invariant subspaces. Moreover, the minimal polynomial of $A|_{W_i}$ is $(x - \lambda_i)^{\dim W_i}$, where λ_i is an eigenvalue of A.

Proof. The proof of (1) is by induction on $n = \dim V$. If $n = 1$, the proof is trivial. So suppose that $n \geq 2$ and the result has been shown for all vector spaces over K with dimension strictly smaller than n. If V cannot be decomposed into a direct sum of A-invariant subspaces, then we are done. Otherwise, $V = U \oplus W$, where U and W are both proper A-invariant subspaces. Since U and V are proper, by the inductive hypothesis U and V can be decomposed into irreducible A-invariant subspaces, thus proving the existence of such a decomposition.

To prove the "moreover," let p_i be the minimal polynomial of $A|_{W_i}$ and suppose that p_i has at least two distinct roots. Then $p_i = q_1 q_2$, where q_1 and q_2 are co-prime and are both at least degree 1. By the lemma, $W_i = U_1 \oplus U_2$ where $U_1 = \{x \in W_i : q_1(A) = 0\}$ and $U_2 = \{x \in W_i : q_2(A) = 0\}$, contradicting the irreducibility of W_i. It follows that $p_i(x) = (x - \lambda_i)^k$ for some $k \geq 1$, where λ_i is an eigenvalue of A. By the cyclic decomposition theorem, each W_i is a direct sum of $A - \lambda_i$ cyclic subspaces. But each such subspace is necessarily A-invariant, and so W_i must itself be $A - \lambda_i$ cyclic. In particular, $k = \dim W_i$, as required.

$\square$

Corollary 110 (Jordan Decomposition). *Let K be an algebraically closed field. A Jordan block is a square matrix of the form*

$$\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix}$$

for some $\lambda \in K$. In other words, J has λ on the diagonal and 1 on the super-diagonal.

The claim is that if A is any $n \times n$ matrix, then A is similar to a matrix of the form

$$\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix}$$

where each J_i is a Jordan matrix with an eigenvalue of A on the diagonal.

Proof. Decompose $V = \bigoplus_{i=1}^k W_i$, where each W_i is an irreducible A-invariant subspace. By inspecting the proof of the master theorem, we see that there exists a basis of W_i of the form $v, (A - \lambda_i I)v, \dots, (A - \lambda_i)^{\dim W_i - 1}v$, where λ_i is an eigenvalue of A. For $j = 1, \dots, \dim W_i$, let $v_j = (A - \lambda_i)^{j-1}v$, and let $v_{j+1} = 0$. Then $A|_{W_i}$ acts on this basis by

$$Av_j = v_{j+1} + \lambda v_j$$

hence the matrix representation of $A|_{W_i}$ with respect to this basis is given by a $\dim W_i \times \dim W_i$ Jordan block whose diagonal is λ_i. $\square$

Theorem 111. *Suppose that A and B are $n \times n$ diagonalizable matrices which such that $AB = BA$. Then A and B are simultaneously diagonalizable, meaning that there exists a basis $v_1, \dots, v_n$ which are eigenvectors for both A and B simultaneously.*

Proof. We can decompose

$$V = \bigoplus_{i=1}^k \ker(A - \lambda_i)$$

where $\lambda_1, \dots, \lambda_k$ are the distinct eigenvalues of A. Now, since $AB = BA$, it holds $\ker(A - \lambda_i)$ is B-invariant. Since B is diagonalizable, the minimal polynomial of B is a product of distinct linear factors, so the minimal polynomial of B restricted to $\ker(A - \lambda_i)$ is also a product of distinct linear factors. In other words, B is diagonalizable over $\ker(A - \lambda_i)$. So for each i there exists a basis $x_1^{(i)}, \dots, x_{k_i}^{(i)}$ of $\ker(A - \lambda_i)$ which are eigenvectors of B which are also eigenvectors of A. Thus the collection $\{x_1^{(i)}, \dots, x_{k_i}^{(i)} : 1 \le i \le k\}$ is a basis of V which are both eigenvectors for A and B simultaneously, completing the proof. $\square$

Exercise 112. Suppose that A is a Hermitian matrix. Then A is diagonalizable.

Proof. We know that

$$V = \bigoplus_{i=1}^j \ker(A - \lambda_i I)^{m_i}$$

where m_i is the multiplicity of the eigenvalue λ_i as a root of the minimal polynomial of A. If $m_i \ge 2$ for some i, then there is $y \ne 0$ such that $(A - \lambda_i I)y \ne 0$ but $(A - \lambda_i I)^2 y 0$. But that $|(A - \lambda_i I)y|^2 = ((A - \lambda_i I)y, (A - \lambda_i)y) = (y, (A - \lambda_i I)^2 y) = 0$, a contradiction. So $m_i = 1$ for all i, which means A is diagonalizable. $\square$

Exercise 113. Suppose that A is a normal operator, meaning $A^*A = AA^*$, where $*$ denotes the Hermitian conjugate. Then A is diagonalizable.

Proof. Label

$$S = \frac{A + A^*}{2},\ T = \frac{A - A^*}{2i}.$$

Note that S, T are both Hermitian, so are diagonalizable. Suppose that $S = P_1 D_1 P_1^{-1}$ and $T = P_2 D_2 P_2^{-1}$ with D_1 and D_2 diagonal. Since A is normal, S and T commute, hence are simultaneously diagonalizable. So there is a matrix P with $S = PD_1P^{-1}$ and $T = PD_2P^{-1}$, which means

$$A = S + iT = P(D_1 + iD_2)P^{-1},$$

and as $D_1 + iD_2$ is diagonal we are done. $\square$

Exercise 114. Suppose that A is a matrix over $\mathbb{C}$ and f is a polynomial. Then every eigenvalue of $f(A)$ is of the form $f(\lambda)$, where λ is an eigenvalue of A.

Proof. If v is an eigenvector of A with eigenvalue λ, then $f(A)v = f(\lambda)v$, so $f(\lambda)$ is an eigenvalue of $f(A)$. Conversely, if μ is an eigenvalue of $f(A)$, then on the A-invariant subspace $W := \ker(f(A) - \mu)$, the minimal polynomial p of $A|_W$ divides $f - \mu$. As p also divides the minimal polynomial of A, it has a root λ which is an eigenvalue of A, hence $f(\lambda) - \mu = 0$. $\square$

Exercise 115. If A is an invertible linear map such that A^2 is diagonalizable, then A is diagonalizable.

Proof. The minimal polynomial q of A^2 is of the form $q(x) = (x - \lambda_1^2)\cdots(x - \lambda_k^2)$ where $\lambda_i^2 \neq \lambda_j^2$ if $i \neq j$ and each λ_i is an eigenvalue of A. If p is the minimal polynomial of A, then $p \mid q(x^2)$, so in particular p divides $(x - \lambda_1)(x + \lambda_1)\cdots(x - \lambda_k)(x + \lambda_k)$. But since $\lambda_i^2 \neq \lambda_j^2$ for $i \neq j$ and $\lambda_i \neq 0$ for all i, it holds that $q(x^2)$ is a product of distinct linear factors. But this means p is a product of distinct linear factors, hence A is diagonalizable. $\square$

Exercise 116. Suppose that P and Q are square matrices such that $P^2 = P$, $Q^2 = Q$, and $1 - (P + Q)$ is invertible. Then P and Q have the same rank.

Proof. Let $A = 1 - (P + Q)$. Notice that $APx = -Qx$ and $AQx = -Px$. Thus $\mathrm{Im}(AP) \subset \mathrm{Im}(Q)$ and $\mathrm{Im}(AQ) \subset \mathrm{Im}P$, and in particular $\mathrm{rank}(AP) \leq \mathrm{rank}\,Q$ and $\mathrm{rank}(AQ) \leq \mathrm{rank}\,P$. But A is invertible, so $\mathrm{rank}(AP) = \mathrm{rank}\,P$ and $\mathrm{rank}(AQ) = \mathrm{rank}\,Q$, which yield $\mathrm{rank}\,P = \mathrm{rank}\,Q$. $\square$

Exercise 117. Let P_n be the space of polynomials of degree at most $2n + 1$, where $n \geq 0$. Then there exists unique constants $c_1, \ldots, c_n$ such that

$$\int_{-1}^{1} p(x)\,dx = 2p(0) + \sum_{k=1}^{n} c_k(p(k) + p(-k) - 2p(0))$$

for all $p \in P_n$.

Lemma 118. *Let V be a vector space over $\mathbb{C}$ and suppose that $T, T_1, \ldots, T_k \in V^*$ are such that $T_1x = \cdots = T_kx = 0$ implies $Tx = 0$. Then there exists constants $c_1, \ldots, c_k \in \mathbb{C}$ such that $T = c_1T_1 + \cdots + c_kT_k$. If the map $(T_1, \ldots, T_k) : V \to \mathbb{C}^k$ is surjective, then the constants are unique.*

Proof. Let $\Lambda = (T_1, \ldots, T_k)$, so that $\ker\Lambda \subset \ker T$. Define a linear map $\beta : \mathrm{Im}\Lambda \to \mathbb{C}$ by $\beta(y) = Tx$ if $y = \Lambda x$. The assumption $\ker\Lambda \subset \ker T$ implies that β is a well-defined linear map. Moreover, $\beta \circ \Lambda = T$. We can extend β to a linear map on $(\mathbb{C}^k)^*$, hence there are constants $c_1, \ldots, c_k$ such that $\beta(x_1, \ldots, x_k) = c_1x_1 + \cdots c_kx_k$ for all $(x_1, \ldots, x_k) \in \mathbb{C}^k$. To establish uniqueness when Λ is surjective, suppose that $\beta' : \mathbb{C}^k \to \mathbb{C}$ is another map such that $\beta' \circ \Lambda = T$. Then since Λ is surjective it follows $\beta = \beta'$, as required. $\square$

Proof. For $1 \leq k \leq n$ let $T_k : P_n \to \mathbb{C}$ be the map $T_kp = p(k) + p(-k) - 2p(0)$. Let $Tp = -2p(0) + \int_{-1}^{1} p\,dx$. First, we need to show that $(T_1, \ldots, T_n)$ is a surjective map onto $\mathbb{R}^k$. To this end, given any points $y_1, \ldots, y_k$, by inverting the Vandermonde matrix generated by the points j^2 for $0 \leq j \leq n$ we can find a unique polynomial q of degree at most n such that $q(0) = 0$ and $q(k^2) = y_k$ for each $1 \leq k \leq n$. Note that $q(x^2) = p(x) + p(-x) - 2p(0)$ for some polynomial p of degree at most $2n + 1$, and thus $T_kp = y_k$ for each k, proving $(T_1, \ldots, T_n)$ is surjective.

If we can show that $T_1p = \cdots = T_np = 0$ implies $Tp = 0$, then we are done. To this end, suppose that $T_1p = \cdots = T_np = 0$. Let $q(x) = p(x) + p(-x) - 2p(0)$. Then q is a degree at most $2n$ polynomial, and has $2n + 1$ roots 0 and $\pm k$ for $1 \leq k \leq n$. Thus $q = 0$. On the other hand, we can write

$$Tp = \int_{-1}^{1} \frac{q(x)}{2}\,dx + \int_{-1}^{1} \frac{p(x) - p(-x)}{2}\,dx.$$

Since $\frac{p(x)-p(-x)}{2}$ is an odd function, $\int_{-1}^{1} \frac{p(x)-p(-x)}{2}\,dx = 0$. Since $q = 0$, we thus have $Tp = 0$, completing the proof. $\square$

Exercise 119. Find all continuous bijections $f : [0,1] \to [0,1]$ such that

$$\int_0^1 g(f(x))\,dx = \int_0^1 g(x)\,dx$$

for all continuous functions $g : [0,1] \to \mathbb{R}$.

Proof. Since f is a continuous bijection, it either strictly increases or strictly decreases, and by replacing f with $1-f$ we can assume f strictly increases. With this assumption, it holds $f(0) = 0$ and $f(1) = 1$. Then we have

$$\int_0^1 f(x)^2\,dx = \frac{1}{3}, \quad \int_0^1 f^{-1}(x)^2\,dx = \frac{1}{3}$$

by taking $g(x) = x^2$ and $g(x) = f^{-1}(x)^2$. Since f strictly increases, it holds

$$\begin{aligned}\int_0^1 xf(x)\,dx &= \int_0^1 x \int_0^{f(x)} dy\,dx \\ &= \int_0^1 \int_{f^{-1}(y)}^1 x\,dx\,dy \\ &= \frac{1}{2}\int_0^1 1 - f^{-1}(y)^2\,dy \\ &= \frac{1}{3}.\end{aligned}$$

This proves

$$\int_0^1 (f(x)-x)^2\,dx = 0.$$

We deduce $f(x) = x$ for all x, hence the only measure preserving bijections are x and $1-x$. $\square$

Exercise 120. Let A be a $n \times n$ complex matrix with characteristic polynomial χ. Show that A is nilpotent if and only if $|\chi(z)| = 1$ whenever $|z| = 1$.

Proof. If A is nilpotent then the only eigenvalues of A are 0, hence $\chi(z) = z^k$ for some $1 \le k \le n$. Thus $|\chi(z)| = 1$ if $|z| = 1$. For the converse, if $|\chi(z)| = 1$ for $|z| = 1$ then $\chi(z) = z^n$. By Cayley-Hamilton it follows $A^n = 0$, thus A is nilpotent. $\square$

Exercise 121. Suppose that p is a non-constant complex polynomial such that $|p(z)| = 1$ whenever $|z| = 1$. Show that $p(z) = cz^k$ for a constant $|c| = 1$ and some positive integer k.

Proof. If p does not have a zero in $|z| < 1$, then p and $1/p$ are holomorphic in the unit disk and satisfy $|p(z)| \le 1$ and $|1/p(z)| \le 1$ for $|z| < 1$ by maximum modulus. Thus $|p(z)| = 1$ for $|z| < 1$, and by another application of maximum modulus it follows p is constant. If p has some (possibly repeating) roots $\alpha_1, \ldots, \alpha_k$ in $|z| < 1$, consider

$$q(z) = \frac{p(z)}{\frac{\alpha_1 - z}{1-\bar{\alpha}_1 z} \cdots \frac{\alpha_k - z}{1-\bar{\alpha}_k z}}.$$

Then $|q(z)| = 1$ for $|z| = 1$ and q can be extended to a non-vanishing holomorphic function in $|z| < 1$. Thus by our previous argument q is constant, which implies that

$$p(z) = c\prod_{j=1}^{k} \frac{\alpha_j - z}{1-\bar{\alpha}_j z}, \quad |c| = 1.$$

But p is a polynomial, and the only way for the right-hand side to be a polynomial is if $\alpha_1 = \cdots = \alpha_k = 0$ so that $p(z) = cz^k$. $\square$

Exercise 122. Suppose that X is a complete metric space and $T : X \to X$ is a map such that T^m is a contraction for some $m \geq 1$. Then T has a unique fixed point in X.

Proof. Since T^n is a contraction it has a unique fixed point $x \in X$. Thus $T^n x = x$, hence $T^{n+1}x = Tx$. This implies $T^n(Tx) = Tx$, and by uniqueness of x it follows $Tx = x$. $\square$

Exercise 123. Suppose that f is holomorphic on $\mathbb{C}\setminus\{0\}$ and is homogeneous of degree $\alpha \in \mathbb{R}$, meaning that $f(\lambda w) = \lambda^\alpha f(w)$ whenever $w \in \mathbb{C} \setminus \{0\}$ and $\lambda > 0$. Show that α is an integer and $f(z) = cz^\alpha$ for some constant c.

Proof. Let $f(z) = \sum_{n=-\infty}^{\infty} a_n z^n$ be the Laurent expansion of f. If f is not identically 0, then fix $n \in \mathbb{Z}$ such that $a_n \neq 0$. It follows that

$$a_n = \frac{1}{2\pi i}\int_{C_R} \frac{f(z)}{z^{n+1}}\,dz$$

where C_R is the disk $|z| = R$. By homogeneity we then have $a_n = R^{\alpha-n}a_n$ for all $R > 0$, which is only possible if $\alpha = n$. So for $k \neq n$ we have $a_k = R^{n-k}a_k$ hence $a_k = 0$. Thus $f(z) = a_n z^n$, as required. $\square$

Exercise 124. Let $M_2(\mathbb{Q})$ be the ring of all 2×2 matrices with coefficients in Q. Describe all field extensions K of $\mathbb{Q}$ such that there is an injective ring homomorphism $K \to M_2(\mathbb{Q})$. (Note: we take the convention that a ring homomorphism maps the multiplicative identity to the multiplicative identity.)

Claim. *There exists such a homomorphism if and only if* $[K : \mathbb{Q}] \leq 2$.

Proof. If $[K : \mathbb{Q}] = 1$ then there is nothing to show, so suppose that $[K : \mathbb{Q}] = 2$ so that $K = \mathbb{Q}(\alpha)$ for some α. Suppose that α has minimal polynomial $x^2 + bx + c$ for $b, c \in \mathbb{Q}$ and $c \neq 0$. Note that 1 and α define a basis of K/F, therefore by linearity we can uniquely define a map $T : K \to M_2(\mathbb{Q})$ by

$$T1 = I, \quad T\alpha = \begin{bmatrix} 0 & -c \\ 1 & -b \end{bmatrix}.$$

First, we show that T is injective. If $T(x + y\alpha) = 0$ for some $x, y \in \mathbb{Q}$, then

$$\begin{bmatrix} x & -yc \\ y & x - by \end{bmatrix} = 0.$$

This forces $x = y = 0$, so T is injective. It is also a ring homomorphism. Clearly it is linear, so we need to establish it is multiplicative. By linearity and the fact $T(1) = I$, it is enough to show that $T(\alpha^2) = T(\alpha)^2$. To this end,

$$T(\alpha)^2 = \begin{bmatrix} -c & bc \\ -b & -c + b^2 \end{bmatrix} = -b\begin{bmatrix} 0 & -c \\ 1 & -b \end{bmatrix} - cI = T(-b\alpha - c) = T(\alpha^2).$$

Thus T is the desired injective ring homomorphism. To prove the converse, suppose there is an injective ring homomorphism $T : K \to M_2(\mathbb{Q})$. Note that T is necessarily an injective linear map, so by the isomorphism theorem it follows that $[K : \mathbb{Q}] = \dim \mathrm{Im}T \leq \dim M_2(\mathbb{Q}) = 4$. Now, since $\mathrm{Im}T$ forms a commutative subgroup of $M_2(\mathbb{Q})$, we cannot have $\mathrm{Im}T = M_2(\mathbb{Q})$, so $[K : \mathbb{Q}] \leq 3$. If $[K : \mathbb{Q}] = 3$ and $1, \alpha, \beta$ are a basis of $K/\mathbb{Q}$, then note that the minimal polynomials of $T\alpha$ and $T\beta$ are quadratic, and since T is an injective ring homomorphism it follows the minimal polynomials of α and β are also quadratic. Thus since $\alpha\beta \in K$ we have $\alpha\beta = c_0 + c_1\alpha + c_2\beta$ for $c_0, c_1, c_2 \in \mathbb{Q}$, which implies that $\alpha = (\beta - c_1)^{-1}(c_0 + c_2\beta) = x_1 + x_2\beta$ for some $x_1, x_2 \in \mathbb{Q}$ since $\mathbb{Q}(\beta)$ is 2 dimensional. This contradicts the fact $1, \alpha, \beta$ are a basis of $K/\mathbb{Q}$, hence $[K : \mathbb{Q}] \leq 2$. $\square$

Does the above result hold if one replaces $M_2(\mathbb{Q})$ with $M_n(\mathbb{Q})$ with $n \geq 2$? My guess is not, I think one should probably get $[K : \mathbb{Q}] \leq n^2 - n$, since this is essentially the largest number of linearly independent commuting matrices you can have. The issues is that is becomes more difficult to prove

Exercise 125. Suppose that $\{f_\omega : |\omega| < 1\}$ is a family of entire functions such that $f_\omega(z)$ is analytic in ω for each $z \in \mathbb{C}$. Suppose further that f_ω is non-vanishing on $|z| = 1$ for all ω. Show that for any $k \geq 0$, the function

$$N(\omega) := \sum_{|z|<1 : f_\omega(z)=0} z^k$$

is analytic in ω.

Proof. Note that

$$N(\omega) = \frac{1}{2\pi i} \int_{|z|=1} \frac{f'_\omega(z)}{f_\omega(z)} z^k \, dz$$

by the residue theorem and the fact $f_\omega \neq 0$ on $|z| = 1$. Now, for each fixed ω, it holds f'_ω / f_ω is analytic in $|\omega| < 1$ since f_ω is non-vanishing. Thus $f_\omega(z) = \sum_{n \geq 0} a_n(z) \omega^n$, and we get

$$\frac{1}{2\pi i} \int_{|z|=1} \frac{f'_\omega(z)}{f_\omega(z)} z^k \, dz = \frac{1}{2\pi i} \sum_{n \geq 0} \omega^n \int_{|z|=1} z^j a_n(z) \, dz$$

□

Exercise 126. Suppose that A and B are $n \times n$ matrices over a field F such that $A^2 = A$ and $B^2 = B$. If A and B have the same rank, then they are similar.

Proof. Note that we can write $F^n = \ker A \oplus \mathrm{Im} A = \ker B \oplus \mathrm{Im} B$. Since A and B have the same rank there is an isomorphism $\phi_1 : \mathrm{Im} A \to \mathrm{Im} B$ and $\phi_2 : \ker A \to \ker B$. □

Exercise 127. Let F be a finite field of order q. Find the number of $n \times n$ matrices over F which have determinant 1.

Proof. The map $\det : A \mapsto \det A$ is a homomorphism from the multiplicative group of invertible $n \times n$ matrices over F to group of units in F. Thus the number k of such matrices satisfies

$$\frac{(q^n - 1)(q^n - q) \cdots (q^n - q^{n-1})}{k} = q - 1$$

since $(q^n - 1)(q^n - q) \cdots (q^n - q^{n-1})$ is the number of $n \times n$ invertible matrices over F. Thus

$$k = q^{n-1}(q^n - 1) \cdots (q^n - q^{n-2}).$$

□

Exercise 128. Suppose that f and g are entire functions such that

$$f^{(n)} + a_{n-1} f^{n-1} + \cdots + a_0 f = 0$$

$$g^{(m)} + b_{n-1} g^{m-1} + \cdots + b_0 g = 0$$

for some $m, n \geq 1$ and constants $a_0, \ldots, a_n, b_0, \ldots, b_m \in \mathbb{C}$. If $F = fg$, show that there are constants, $c_0, \ldots, c_{mn}$, not all zero, such that

$$c_{mn} F^{(mn)} + \cdots + c_0 F = 0.$$

Proof. Let D be the differentiation operator on the space V of entire functions. Label $p_1(z) = z^n + a_{n-1} z^{n-1} + \cdots + a_0$ and $p_2(z) = z + b_{n-1} z^{m-1} + \cdots + b_0$, so that $f \in \ker p_1(D) =: V_1$ and $g \in \ker p_2(D) =: V_2$. We note that V_1 and V_2 are both finite dimensional, with dimensions n and m, respectively (since $\ker D$ is one dimensional). Let $T : V_1 \times V_2 \to V$ be the bilinear map $T(u, v) = uv$. If W is the subspace generated by $T(V_1 \times V_2)$, then since $DT = T(D, \cdot) + T(\cdot, D)$ and since V_1 and V_2 are D-invariant spaces we see that $D : W \to W$. Note that W is spanned by the set $\{T(v_i, w_j) : 1 \le i \le n,\ 1 \le j \le m\}$ for a basis $\{v_i\}$ of V_1 and a basis $\{w_j\}$ of V_2, which shows $\dim W \le nm$. Therefore the minimal polynomial p of $D|_W$ is degree at most nm. Since $F \in W$, we have $p(D)F = 0$, as required. $\square$

Exercise 129. Let A be the set of positive integers which do not have the digit 9 in their decimal expansions. Prove that

$$\sum_{a \in A} \frac{1}{a} < \infty.$$

Proof. Let $[9] = \{0, \ldots, 8\}$. Then

$$\sum_{a \in A} a^{-1} \le \sum_{n \ge 0} \sum_{(a_0,\ldots,a_n) \in [9]^n \setminus \{0\}} \frac{1}{a_n 10^n + \cdots + a_0} \le \sum_{n \ge 0} \frac{9^n - 1}{10^n} < \infty.$$

$\square$

Exercise 130. What is the smallest order of a field F with characteristic 7 such that $x^{18} + \cdots + x + 1$ has a root in F?

Proof. Note that $x \in F$ solves $x^{18} + \cdots + x + 1 = 0$ if and only if $x^{19} = 1$ and $x \ne 1$. Since 19 is prime this means the unital group of x contains an element of order 19, hence $19 \,|\, (\#F - 1)$. Since $\#F = 7^k$ for some k, the smallest k where the divisibility relation holds is $k = 3$, and by Cauchy's theorem any group of order $7^3 - 1$ has an element of order 19. Thus $7^3 = 343$ is the smallest order. $\square$

Exercise 131. Let V be a real vector space of dimension n, and let $S : V \times V \to \mathbb{R}$ be a nondegenerate bilinear form. Suppose that W is a linear subspace of V such that the restriction of S to $W \times W$ is identically 0. Show that $\dim W \le n/2$.

Proof. Suppose $S(x, y) = x^\top Ay$. Define a map $T : V \to W^*$ by $Tx = (y \mapsto x^\top Ay)$. Since the map $x \mapsto (y \mapsto x^\top Ay)$ is an isomorphism of $V \to V^*$ from the fact S is nondegenerate, we know that $\mathrm{Im} T = W^*$. By rank nullity $n = \dim V = \dim W + \dim\ker T \ge 2\dim W$ since $W \subset \ker T$. Thus $\dim W \le n/2$. $\square$

Exercise 132. Let V be a finite dimensional vector space over a field F, and let A and B be linear endomorphisms of V. Prove that

$$\dim\ker AB \le \dim\ker A + \dim\ker B.$$

Proof. Define a map $T : \ker AB \to \ker A$ by $Tx = Bx$. Then $\ker AB/\ker T$ is isomorphic to a subspace of $\ker A$, hence $\dim\ker AB - \dim\ker T \le \dim A$. But $\dim\ker T \le \dim\ker B$ since $\ker T \subset \ker B$, thus $\dim\ker AB \le \dim\ker A + \dim\ker B$. $\square$

Exercise 133. Let X be a metric space and let $K \subset X$ be a compact subset. Let $\{U_\alpha\}_{\alpha \in I}$ be an open cover of K. Show that there exists $\epsilon > 0$ such that for any $x \in K$, the ball of radius ϵ about x is contained in some U_α.

Proof. Suppose that for each $n \ge 1$ there is $x_n \in K$ such that $B_{1/n}(x_n)$ is not contained in any of the U_α. Let $x \in K$ be the limit of a subsequence $\{x_{n_k}\}$, and such x exists by compactness. Then $x \in U_\alpha$ for some α. Fix $r > 0$ such that $B_r(x) \subset U_\alpha$. Then for all k sufficiently large $x_{n_k} \in B_{r/2}(x)$, and for such k it holds $B_{1/n_k}(x_{n_k}) \subset B_r(x) \subset U_\alpha$, a contradiction. The contradiction is resolved if the conclusion of the problem is true. $\square$

Exercise 134. Let c_n be the number of ways you can make n cents from pennies, nickels, dimes, and quarters. Compute

$$\lim_{n \to \infty} \frac{c_n}{n^3}.$$

PROOF. Let

$$f(z) = \sum_{n\geq 0} c_n z^n.$$

Then

$$\begin{aligned} f(z) &= (1+z+z^2+\cdots)(1+z^5+z^{10}+\cdots)(1+z^{10}+z^{20}+\cdots)(1+z^{25}+z^{50}+\cdots) \\ &= \frac{1}{(1-z)(1-z^5)(1-z^{10})(1-z^{25})}. \end{aligned}$$

So f is a meromorphic function with a pole of order 4 at $z=1$, and all the other poles have order at most 3. So

$$f(z) = \frac{c}{(1-z)^4} + \sum_p \frac{c_p}{(p-z)^{k_p}} + g(z),$$

where the sum is over all the poles of f of order at most 3, with $1 \leq k_p \leq 3$ being the multiplicities of each pole. The function g is an entire function, and c, c_p are constants. Taking derivatives yields

$$c_n = \frac{f^{(n)}(0)}{n!} = c\binom{n+3}{3} + \sum_p c_p \binom{n+k_p-1}{k_p-1} + \frac{g^{(n)}(0)}{n!}.$$

Thus

$$\frac{c_n}{n^3} \sim \frac{c}{n^3}\binom{n+3}{3}$$

since $1 \leq k_p \leq 3$ implies $\binom{n+k_p-1}{k_p-1}$ grows at most quadratically in n and since g is entire $\frac{g^{(n)}(0)}{n^3 n!} \to 0$. So we get

$$\lim_{n\to\infty} \frac{c_n}{n^3} = \frac{c}{3!}.$$

To compute c, we observe that $c = \lim_{z\to 1}(1-z)^4 f(z)$, and as

$$(1-z)^4 f(z) = \frac{1}{(z^4+\cdots+z+1)(z^9+\cdots+z+1)(z^{24}+\cdots+z+1)}$$

we get $c = \frac{1}{5\cdot 10\cdot 25}$. Thus

$$\lim_{n\to\infty} \frac{c_n}{n^3} = \frac{1}{3!\cdot 5\cdot 10\cdot 25} = \frac{1}{7500}.$$

□

EXERCISE 135. Let (X,d) be a compact metric space and let $C(X)$ be the set of continuous functions $X \to \mathbb{R}$. View $C(X)$ as a unital ring where addition and multiplication of two functions in $C(X)$ is interpreted pointwise. Find the maximal ideals of $C(X)$.

PROOF. The maximal ideals are of the form $\{f \in C(X) : f(x) = 0\}$ for $x \in X$. Clearly each such ideal is maximal (the quotient is isomorphic to $\mathbb{R}$), so given a maximal ideal M we need to show $M = \{f \in C(X) : f(x_0) = 0\}$ for some $x_0 \in X$. Suppose that to each $x \in X$ there is $f_x \in M$ such that $f_x(x) \neq 0$. Then there is a neighborhood U_x containing x such that $|f_x(y)| > \frac{1}{2}|f_x(x)| > 0$ for all $x \in U_x$. Then the $\{U_x : x \in X\}$ are a cover of X, and by compactness there is a finite sub-cover $U_{x_1}, \ldots, U_{x_n}$. Let

$$f(x) = \sum_{k=1}^n f_{x_k}(x)^2 \in M.$$

Then $f \in M$ since M is an ideal and moreover if $x \in X$, then $x \in U_{x_j}$ for some j, so $f(x) \geq f_j(x)^2 \geq \left(\frac{1}{2}|f_{x_j}(x_j)|\right)^2 > 0$. This implies $f \neq 0$ on X hence $1/f \in C(X)$ implies $(1/f)\cdot f = 1 \in M$, contradicting $M \neq C(X)$. So there is $x_0 \in X$ such that $f(x_0) = 0$ for all $f \in M$, and since M is maximal $M = \{f : f(x_0) = 0\}$. □

Exercise 136. Let $F_1, \ldots, F_n$ be fields and let $R = F_1 \times \cdots \times F_n$. Consider R as an abelian ring with component-wise addition and multiplication. Find all maximal ideals of R.

Proof. For $j = 1, \ldots, n$, let I_j be the ideal of R consisting of all elements where the j-th coordinate is 0. Note that $R/I_j \cong F_j$ so I_j is maximal. We just need to show that all such maximal ideals are of this form. Let M be a maximal ideal. There must be some index j such that every element of M is 0 in index j. Otherwise, for any $x_j \in F_j$, M contains the element $(0, \ldots, x_j, \ldots, 0)$ for each $1 \leq j \leq n$. Thus $(x_1, \ldots, x_n) \in M$ for any $(x_1, \ldots, x_n) \in R$, contradicting the fact $M \neq R$. It follows that $M \subset I_j$, and since M is maximal we have $M = I_j$, as required.

□

Corollary 137. *If $p \in F[x]$ for some field F, then there are as many maximal ideals of $F[x]/(p(x))$ as there are irreducible factors of p. Moreover, each maximal ideal in $F[x]/(p(x))$ has a quotient isomorphic to $F[x]/(p_k(x))$, where p_k is an irreducible factor of p.*

Remark 138. To be explicit, for example, if we $p(x) = x^3 - 1$ and $F = \mathbb{R}$ we get $\mathbb{R}[x]/(x^3 - 1)$ has two maximal ideals, corresponding to $\{c(x^2 + x + 1) : c \in \mathbb{R}\}$ and $\{(x - \alpha)(x - 1) : \alpha \in \mathbb{R}\}$.