

A book dedicated to lovers of mathematics

version 2023/11/25

# Contents

|   |           |
|---|-----------|
| <b>0. Problems</b>                      | <b>2</b>  |
| <b>2. Calculus</b>                      | <b>5</b>  |
| <b>3. Number Theory</b>                 | <b>10</b> |
| <b>4. Combinatorics and Probability</b> | <b>11</b> |
| <b>5. Abstract Algebra</b>              | <b>13</b> |
| <b>6. Analysis</b>                      | <b>17</b> |
| <b>-1. Solutions</b>                    | <b>20</b> |
| <b>-2. Notes</b>                        | <b>23</b> |

Table of contents:

Chapter 0. Problems

Chapter 1. Euclidean Geometry

Chapter 2. Calculus

Chapter 3. Number Theory

Chapter 4. Combinatorics and Probability

Chapter 5. Abstract Algebra

Chapter 6. Analysis

Chapter 7. Topology

Chapter -1. Solutions

Chapter -2. Notes

# 0. Problems

id aa. Consider the following stochastic process. First, a probability  $p$  is drawn from a uniform distribution on  $[0, 1]$ . Then 100 independent coins are drawn, each with probability  $p$  to show heads. Without integrals, determine the probability of seeing precisely  $k$  heads.

id ab. A group of  $n$  mathematicians will face the following challenge. Each will be given a hat showing an integer from  $1, \dots, n$  (repetitions and misses allowed), and will be able to see the numbers written on all hats but their own. Once they see each other, and without any communication between them, they will all have to simultaneously guess the number written on their hat. They will win if at least one of them makes a correct guess. They have today to form a strategy that will guarantee a win. What do you suggest?

id ad. Let  $f : \mathbf{Z}^N \rightarrow \mathbf{Z}$  be an additive mapping (namely  $f(x + y) = f(x) + f(y)$ ) from infinite integer sequences to integers. Show that  $f(x_1, x_2, x_3, \dots) = \alpha_1 x_1 + \dots + \alpha_N x_N$  for some finite integer sequence  $\alpha_1, \dots, \alpha_N$ .

id ah. Let  $f : [0, 1] \rightarrow \mathbf{R}$  be a continuous function for which  $\lim_{h \rightarrow 0} \frac{f(t_0 + h) + f(t_0 - h) - 2f(t_0)}{h^2} = 0$  for all  $t_0 \in (0, 1)$ . Then  $f(t) = \alpha t + \beta$  is linear.

id ai. Smooth (or even just continuous) homomorphisms  $\phi : \mathbf{R} \rightarrow \mathrm{GL}_d(\mathbf{C})$  are in 1 : 1 correspondence with  $A \in \mathbf{C}^{d \times d}$  via

$$\begin{aligned}\phi(t) &= \exp(At) \\ A &= \left. \frac{d\phi(t)}{dt} \right|_{t=0}\end{aligned}$$

id aj. Let  $a_n$  be a sequence for which  $\sum |a_n b_n| < \infty$  for all  $b_n \in \ell^2$ . Then  $a_n \in \ell^2$ .

id ak. A two player game goes as follows. Each player is given a number, drawn at random, independently, uniformly from  $[0, 1]$ . A player may keep their number, or ask for another - in which case they must keep the second number. Then the two players reveal their numbers, and the player with the bigger number wins. Find the optimal playing strategy.

id al.  $n$  ants are placed inside a circle of radius  $R$ , each is initially heading east, west, north, or south, and all ants have constant speed 1. When two ants facing opposite directions collide, they both immediately turn 90 degrees. Find the minimal length of time to guarantee all ants will leave the circle.

id am. Customers arrive to a store via a Poisson process (with some constant). For every arrival, the employee calculates the probability this will be the last customer of his shift. At the end of their shift, they write down the probability calculated for the last customer. Over many shifts, what is the distribution of probabilities?

id an. Two players take turns coloring the vertices of a graph. In the beginning, all vertices are white. Player 1 picks the first vertex and colors it black. Afterwards, the next player must color black a vertex that is currently white, and adjacent to the the last colored vertex. A player unable to do so losses the game. What property of

the graph is equivalent to the first/second player having a winning strategy, and what is their strategy?

id ao. A very long road contains  $n$  cars, each having a constant speed drawn at random from 20 kmh to 200 kmh. The road is narrow so there's no overtaking, so fast cars may have to slow down to match the speed of the car in front. Eventually, how many meshes of connected cars are to be expected?

id ap. Alice and Bob play the following game. There is an outer circle, fixed in position, containing  $n$  lamps, and an inner circle, which may be rotated, containing the numbers  $1..n$ . Alice wins when all the lights are off, and Bob wins if she gives up/the game goes forever. Bob chooses the initial on/off states of the lamps. Every turn, Alice specifies a set of numbers in  $1..n$  and tells Bob to switch the state (on  $\leftrightarrow$  off) of the lamps corresponding to those positions, but Bob may spin the inner circle of numbers before doing so. For which values of  $n$  can Alice/Bob win, and how?

id aq. A rectangular grid  $R$  is painted using 10 colors. A rectangle is called *special* if its four vertices have the same color. Find dimensions for  $R$  guaranteeing the existence of a special rectangle.

id ar. There exists an uncountable family  $\mathcal{F}$  of infinite subsets of  $\mathbf{N}$  where the intersection of any pair of sets in  $\mathcal{F}$  is finite.

id as. You find yourself in a huge circular train (say about the size of the equator) in which you can walk freely in both directions. You'll be free once you determine the exact number of cars, but you only have one guess. All the cars look exactly the same, and you can't leave any mark, except each car has a lamp with a switch which you may freely use. However, the initial state of the lamps is random. What do you do?

id at. An invisible frog lives on the integer number line, and each day it hops by a fixed unknown amount  $d$ . Every day you get one guess as to where the frog is, and if you hit it you win a prize. Find a strategy that guarantees you'll win the prize, eventually.

Part ii. Now the frog moved to the real line, and still hops every day by a fixed amount. It also gained an unknown small but positive length  $\varepsilon$ . What's your strategy?

id au. A standard 2-dimensional Gaussian is given in polar coordinates by  $\sqrt{2\log(1/u)}(\cos 2\pi\theta, \sin 2\pi\theta)$  where  $u, \theta$  are independent and uniform on  $[0, 1]$ .

id aw. There's a deck with 100 numbered  $1..100$  in some order. In each step we peek at the top card, say numbered  $k$ , take the first  $k$  cards from the deck, and shuffle them so that  $k$  becomes the bottom card, and then return them on top of the rest of the deck. In particular, after this move the  $k$ -th card now shows  $k$ . Show that eventually the top card becomes 1.

id ay. Let  $f : [0, 1] \rightarrow \mathbf{R}$  be a continuous function satisfying  $\int_0^1 f(x)x^k = 1$  for  $k = 0, \dots, n-1$ . What is the best lower bound on  $\int_0^1 f^2$ ?

id bd. From each vertex of a triangle, extend the two edges by the length of the opposite edge, then the six endpoints lie on a circle with center  $I$  and radius  $\sqrt{r^2 + s^2}$ .

id bq. If  $a_n \rightarrow a$  then  $\left(1 + \frac{a_n}{n}\right)^n \rightarrow e^a$ .

id br. A betting game with odds  $p > 1/2$  in your favour goes as follows. In each round, you may wager an

amount  $x$  of your capital. If the odds were right you get back  $2x$ , and if not you lose the initial  $x$ . You start with \$1 and the game goes for  $n$  rounds. What strategy should be played to maximize the expected value of your capital at the end of the game? What about maximizing the expected value of the log of your capital.

id bs. Compute the mean and variance of  $e^X$  where  $X \sim N(\mu, \sigma^2)$  is normal.

id bu. Given a sequence of distribution functions  $F_n$  there exists a right-continuous weakly-increasing function  $F : \mathbf{R} \rightarrow [0, 1]$  and a sub-sequence  $F_{s(n)}$  for which  $\lim_{n \rightarrow \infty} F_{s(n)}(x) = F(x)$  for all points of continuity  $x$  of  $F$ . Moreover,  $F$  is a distribution function if  $\forall \varepsilon \quad \exists a < b \quad \forall n \quad F_n(b) - F_n(a) \geq 1 - \varepsilon$ .

id cf.

$$\frac{1^5}{1+e^\pi} + \frac{3^5}{1+e^{3\pi}} + \frac{5^5}{1+e^{5\pi}} + \frac{7^5}{1+e^{7\pi}} \cdots = \frac{31}{504}$$

## 2. Calculus

This chapter emphasizes intuition over rigour. Proofs will later be given in the Analysis chapter.

id ac. Fact: For all  $x \in \mathbf{R}$  we have

$$\exp(x) = e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n$$

Nomenclature. The *Laplace transform* of a bounded continuous function  $f : [0, \infty) \rightarrow \mathbf{R}$  is

$$(\mathcal{L}[f])(\lambda) = \int_0^{\infty} e^{-\lambda x} f(x) dx$$

defined for  $\lambda > 0$ .

id bi. Fact: We have

$$\mathcal{L}[f'](\lambda) = \lambda \mathcal{L}[f] - f(0)$$

id ca. Fact: The De-Rham cohomology of  $\mathbf{R}^2 \setminus \{(0, 0)\}$  is  $H^0 \simeq \mathbf{R}$ ,  $H^1 \simeq \mathbf{R}$ ,  $H^2 = 0$ .

id ca. Explanation: The punctured plane is connected, so  $H^0 = \{\text{constant functions}\}$ . A closed 1-form  $\omega = u(x, y)dx + v(x, y)dy$  has  $u_y = v_x$ , and so  $\oint_{\partial R} \omega = 0$  for any rectangle  $R$  not including the origin. It is easy to see that there is a common value  $\nu(\omega)$  of  $\oint_{\partial R} \omega$  for all rectangles  $R$  around the origin. We'll show that  $\omega$  is exact iff  $\nu(\omega) = 0$ . Indeed, if  $\omega = df$  then  $\oint \omega = 0$  for any loop. On the other hand, if  $\nu(\omega) = 0$  then there is a well defined smooth function  $f(x, y) = \int_{\gamma} \omega$ , letting  $\gamma$  be any curve piece-wise parallel to the axes, going from  $(1, 0)$  to  $(x, y)$  not through the origin, and we have  $df = \omega$ . A closed 1-form which is not exact is  $\alpha = "d\theta" = \frac{x dy - y dx}{x^2 + y^2}$ . Indeed we have  $d\alpha = 0$  and  $\oint_{x^2+y^2=1} \alpha = 2\pi$ . It follows that any closed 1-form is uniquely a constant times  $\alpha$  plus an exact form, or  $H^1 = \mathbf{R}\alpha$ . Finally we must show any 2-form  $\eta = f(x, y)dx \wedge dy$  is exact. An engineer's solution would be to write  $\eta = \tilde{f}(r, \theta)rdr \wedge d\theta = d\left(\left(\int_1^r \tilde{f}(t, \theta)tdt\right)d\theta\right)$ . One can bring this back rigorously to  $x, y$  coordinates and verify it works.

id cp. Fact: For all  $z \in \mathbf{C}$  we have

$$\pi \cot(\pi z) = \frac{1}{z} + \sum_{m \geq 1} \frac{1}{z+m} + \frac{1}{z-m}$$

id cp. Explanation: Apply a logarithmic derivative to the sine product.

id cp. Explanation 2: It is easy to see the series converges to a meromorphic function of period 1 with simple poles at the integers with residues 1. Hence the difference between the two sides is an entire function. But one

can see that the difference is bounded, hence constant. Developing to first order at  $z = 0$  we find the two sides are equal.

## Elliptic functions

Nomenclature. The Weierstrass function of a 2 dimensional lattice  $L$  in  $\mathbf{C}$  is  $\wp(z) = z^{-2} + \sum_{\omega \in L \setminus 0} ((z - \omega)^{-2} - \omega^{-2})$

id cb. Fact:

i.  $\wp$  is a well defined  $L$ -periodic even meromorphic function of order 2 with double poles on the points of  $L$ . We have  $\wp'(\eta) = 0$  for  $\eta \in \frac{1}{2}L \setminus L$ .

ii. The expansion of  $\wp$  at the origin is  $\wp(z) = z^{-2} + \sum_{n \geq 1} (2n+1)G_{2n+2}z^{2n}$ , where  $G_m = \sum_{\omega \in L \setminus 0} \omega^{-m}$ .

iii. We have

$$\wp'(z)^2 = 4\wp(z)^3 - g_2\wp(z) - g_3 = 4(\wp(z) - e_1)(\wp(z) - e_2)(\wp(z) - e_3)$$

where  $g_2 = 60G_4, g_3 = 140G_6, e_1 = \wp(\frac{\omega_1}{2}), e_2 = \wp(\frac{\omega_2}{2}), e_3 = \wp(\frac{\omega_1 + \omega_2}{2})$  for  $L = \mathbf{Z}\omega_1 \oplus \mathbf{Z}\omega_2$

iv. We have

$$\begin{vmatrix} 1 & \wp(u) & \wp'(u) \\ 1 & \wp(z) & \wp'(z) \\ 1 & \wp(z+u) & -\wp'(z+u) \end{vmatrix} = 0$$

v.

$$p(u) + p(z) + p(u+z) = \frac{1}{4} \left( \frac{\wp'(u) - \wp'(z)}{\wp(u) - \wp(z)} \right)^2$$

id cb. Explanation: It's simple to verify  $\wp$  is a well defined even meromorphic function with double poles on the points of  $L$ . Clearly  $\wp'(z) = -2 \sum_{\omega \in L} (z - \omega)^{-3}$  is  $L$ -periodic. Hence for  $\omega \in L$ ,  $\wp(z + \omega) - \wp(z)$  is a constant.

But  $\wp$  is even, so this constant is zero and  $\wp$  is  $L$ -periodic of order 2. For  $\eta$  a proper half lattice point, we see that  $\wp'(\eta) = 0$  either from the series for  $\wp'$ , or by deriving the identity  $\wp(z) - \wp(2\eta - z) = 0$  at  $\eta$ . The expansion at 0 is clear, and the left=center identity follows because the difference is  $L$ -periodic without poles with no zero-th order terms. Since  $\wp$  has order 2, the values  $e_1, e_2, e_3$  are distinct (each is assumed with multiplicity 2). Thus  $e_1, e_2, e_3$  are precisely the three roots of the polynomial  $4t^2 - 60G_4t - 140G_6$ . Now, given (general)  $z_1, z_2$  we find  $a, b$  which solve  $b = a\wp(z_i) + \wp'(z_i)$ . Then  $f(z) = \wp'(z) - a\wp(z)$  is an elliptic function of  $L$  of order 3 with a triple pole at 0 and roots at  $z_1, z_2$ . Since the sum of the three roots minus poles of  $f$  (inside a parallelogram) is congruent to 0 mod  $L$ ,  $-z_1 - z_2$  is a root of  $f$ , yielding the determinant formula. Finally, the line  $y = ax + b$  and cubic  $y^2 = 4x^3 - g_2x - g_3$  intersect at  $(\wp(z_i), \wp'(z_i))$  and  $(\wp(-z_1 - z_2), \wp'(-z_1 - z_2))$ . It follows that  $4x^3 - a^2x^2 - (g_2 + 2ab)x - (g_2 + b^2)$  has roots at  $x = \wp(z_i), \wp(-z_1 - z_2)$  and so the sum of the roots is  $a^2/4$ , yielding v.

id cc. Fact: For  $\tau \in \mathbf{H}, q = e^{2\pi i\tau} \in \mathbf{D}^\times$  we have

$$\begin{aligned} \sum_{m \in \mathbf{Z}} (m + \tau)^{-4} &= \frac{8\pi^4}{3} \sum_{\ell=1}^{\infty} \ell^3 q^\ell \\ \sum_{m \in \mathbf{Z}} (m + \tau)^{-6} &= -\frac{8\pi^6}{15} \sum_{\ell=1}^{\infty} \ell^5 q^\ell \\ g_2(\tau) &= \frac{4\pi^4}{3} \left( 1 + 240 \sum_{\ell=1}^{\infty} \sigma_3(\ell) q^\ell \right) \\ g_3(\tau) &= \frac{8\pi^6}{27} \left( 1 - 504 \sum_{\ell=1}^{\infty} \sigma_5(\ell) q^\ell \right) \end{aligned}$$

id cc. Explanation: Recall [??]

$$\frac{1}{\tau} + \sum_{m \neq 0} \left( \frac{1}{\tau + m} - \frac{1}{m} \right) = \pi \cot(\pi\tau) = i\pi \frac{q+1}{q-1} = -i\pi \left( 1 + 2 \sum_{\ell=1}^{\infty} q^{\ell} \right)$$

repeated differentiation yields the first two identities. Once those are in place, we have  $G_{2j}(\tau) = \sum_{n,m} (m + n\tau)^{-2j} = 2\zeta(2j) + 2 \sum_{n,\ell=1}^{\infty} c\ell^{2j-1} q^{n\ell} = 2\zeta(2j) + 2c \sum_{k=1}^{\infty} \sigma_{2j-1}(k) q^k$

Nomenclature.  $\Delta = \Delta(4t^2 - g_2t - g_3) = g_2^3 - 27g_3^2$  and  $J = \frac{g_2^3}{g_2^3 - 27g_3^2}$ . These are functions of the lattice  $L$ , or functions of two generators  $\omega_1, \omega_2$  of  $L$ , or functions of one  $\tau \in \mathbf{H}$  with  $L = \mathbf{Z} + \tau\mathbf{Z}$ .

id ce. Fact: For  $\tau \in \mathbf{H}, q = e^{2\pi i\tau} \in \mathbf{D}^{\times}$  we have

$$\Delta(\tau) = (2\pi)^{12} \sum_{n=1}^{\infty} \tau(n) q^n$$

$$12^3 J(\tau) = q^{-1} + 744 + \sum_{n=1}^{\infty} c(n) q^n$$

where  $\tau(n), c(n) \in \mathbf{Z}$ .

id ce. Explanation: Using [cc] we have

$$(2\pi)^{-12} \Delta = (2\pi)^{-12} (g_2^3 - 27g_3^2) = \frac{1}{27 \cdot 64} ((1 + 240A)^3 - (1 - 504B)^2) = \frac{5A + 7B}{12} + C$$

So we need to show  $12 \mid 5\sigma_3(n) + 7\sigma_5(n)$  or simply  $12 \mid 5k^3 + 7k^5$ . Now we notice  $\tau(1) = 1, \tau(2) = -24$  and so

$$12^3 J = 12^3 \frac{g_2^3}{\Delta} = \frac{12^3 \cdot 64\pi^{12} (1 + 240A)^3}{27(2\pi)^{12} q(1 - 24q - Eq^2)} = \frac{(1 + 240A)^3}{q(1 - 24q - Eq^2)} = q^{-1} (1 + 3 \cdot 240q + \dots)(1 + 24q + \dots)$$

Nomenclature. A *modular function* is a meromorphic function on  $\mathbf{H} \cup i\infty$  invariant under  $\text{PSL}_2(\mathbf{Z})$ . Namely, if  $g$  is a meromorphic function on  $\mathbf{D}$ , and  $f(\tau) = g(e^{2\pi i\tau})$  satisfies  $f(\tau) = f(-1/\tau)$  then  $f$  is a *modular function*.

Example.  $J$  is a modular function.

Nomenclature.  $\mathcal{R} = \{\tau \in \mathbf{H} : |\tau| > 1, -0.5 \leq \Re(\tau) < 0.5\} \cup \exp(i[\pi/2, 2\pi/3])$  is the *fundamental region*.

id cg. Fact: Each  $\tau \in \mathbf{H}$  has exactly one  $\tau' \in \mathcal{R}$  for which  $\tau \simeq \tau'$  under  $\text{PSL}_2(\mathbf{Z})$ .

Moreover, if  $A\tau = \tau$  and  $A \neq \text{id}$  then  $\tau = i, e^{2\pi i/3}$ .

The solutions in case  $\tau = i$  are  $Az = z, \frac{-1}{z}$  and for  $\tau = e^{2\pi i/3}$  are  $Az = z, \frac{-1}{z+1}, \frac{-z-1}{z}$ .

id cg. Explanation. Let  $\tau'$  be an element of  $\text{PSL}_2(\mathbf{Z})\tau$  of maximal imaginary part. It exists because  $\Im A\tau = \frac{\Im \tau}{|c\tau + d|^2}$  and  $\{c\tau + d : \gcd(c, d) = 1\}$  is discrete and has a shortest element. Under  $\tau \mapsto \tau \pm 1$  we may assume  $-0.5 \leq \Re \tau' < 0.5$ . It must be the case that  $|\tau'| \geq 1$  for otherwise  $-1/\tau'$  would have a larger imaginary part. If  $|\tau'| > 1$  then  $\tau' \in \mathcal{R}$ , and if  $|\tau'| = 1$  then either  $\tau'$  or  $-1/\tau'$  is in  $\mathcal{R}$ . Now we need to see that no two elements of  $\mathcal{R}$  are equivalent. Let  $\tau, A\tau \in \mathcal{R}$ . wlog let  $\Im A\tau \geq \Im \tau$ , so that  $1 \geq |c\tau + d| \geq |c| \Im \tau \geq |c| \sqrt{3}/2$  and  $c = 0, \pm 1$ . If  $c = 0$  then  $Az = z + b$  and clearly if  $\tau, A\tau \in \mathcal{R}$  then  $b = 0, A = \text{id}$  and  $\tau = A\tau$ . If  $c = \pm 1$  we may take  $c = 1$  and so  $1 \geq |\tau + d|$ . Therefore  $\tau$  is at most at distance one from an integer, and this integer must be  $d = 0, 1$ . If  $d = 0$  then  $|\tau| = 1$  and  $Az = a - \frac{1}{z}$ . It is clear that either  $a = 0, \tau = A\tau = i$  or



$a = -1, \tau = A\tau = e^{2\pi i/3}$ . Finally, if  $d = 1$  then  $\tau = e^{2\pi i/3}$  and  $\Im A\tau = \sqrt{3}/2$  and so  $\tau = A\tau = e^{2\pi i/3}$ .

id ch. Fact: Let  $f$  be a non-constant modular function. Then  $f$  has an *order*  $n \geq 1$  such that each value  $c \in \mathbf{C}_\infty$  is obtained by  $f$  in  $\mathcal{R} \cup \{i\infty\}$  with  $\mathcal{R}$ -multiplicity  $n$ . Here the  $\mathcal{R}$ -multiplicities at  $i, e^{2\pi i/3}$  are to be divided by 2, 3. In particular, they are divisible by 2, 3.

id ch. Explanation: It suffices to show  $0, \infty$  are obtained with the same multiplicity. There's a finite number of zeros and poles in  $\mathcal{R}$  because there isn't an essential singularity at  $\tau = i\infty, q = 0$ . Let us assume there are no zeros or poles on  $\partial\mathcal{R}$ . To get the number of zeros minus the number of poles, we take a sufficiently large  $M$  and integrate  $\frac{f'(\tau)d\tau}{2\pi i f(\tau)}$  over  $\partial(\mathcal{R} \cap \Im \leq M)$ . The value over the vertical parts cancel, and so do the two symmetric parts of the unit circle. [Indeed  $f'(w)/f(w)dw = f'(z)/f(z)dz$  for  $w = z + 1$  or  $w = -1/z$ ] and we're left with only the (signed) order of the zero/pole at  $\tau = i\infty$ . In case there's zeros or poles on the path of integration, we change it. Only if the zero/pole is at  $\tau = i, e^{2\pi i/3}$  the path is a half circle around  $i$  and two 60 degree sectors around  $e^{2\pi i/3}, e^{2\pi i/3} + 1$ .

id ci. Fact: The order of  $J$  is 1. We have  $j(i\infty) = \infty, j(i) = 1, j(e^{2\pi i/3}) = 0$  with (regular) multiplicities 1, 2, 3.

id cj. Fact: We have  $\overline{J(\tau)} = J(-\bar{\tau})$ . Also,  $J$  maps  $-\frac{1}{2} + i[\frac{\sqrt{3}}{2}, \infty), \exp(i[2\pi/3, \pi/2]), i[1, \infty)$  to  $(-\infty, 0], [0, 1], [1, \infty)$ .

id ck. Fact:  $J$  acts as a covering map of degree 1 from  $\mathbf{H} \setminus \text{PSL}_2(\mathbf{Z})\{i, e^{2\pi i/3}\}$  to  $\mathbf{C} \setminus \{0, 1\}$ .

id cl. Fact: An entire function  $f : \mathbf{C} \rightarrow \mathbf{C}$  omitting two values must be constant.

id cl. Explanation: If  $f$  omits  $a, b$  then  $g = \frac{f-a}{b-a}$  omits 0, 1. Since  $\mathbf{C}$  is simply connected,  $g$  may be lifted to a holomorphic  $h : \mathbf{C} \rightarrow \mathbf{H}$  with  $g = J \circ h$ . But then  $h$  must be constant, as must  $g$  and  $f$ .

id cm. Fact: For  $s > 0$  we have

$$\sum_{n \in \mathbf{Z}} e^{-\pi s n^2} = \frac{1}{\sqrt{s}} \sum_{n \in \mathbf{Z}} e^{-\pi n^2/s}$$

id cm. Explanation: We apply Poisson summation formula to  $e^{-\pi s x^2}$ .

id cn. Fact: More generally, if  $a, b, \tau \in \mathbf{C}$  and  $\Im \tau > 0$  then

$$\sum_{n \in \mathbf{Z}} e^{i\pi \tau (n+a)^2} e^{-2\pi i b(n+a)} = \frac{1}{\sqrt{-i\tau}} \sum_{n \in \mathbf{Z}} e^{-i\pi (n+b)^2/\tau} e^{2\pi i n a}$$

id cn. Explanation: We apply Poisson summation to  $e^{-\pi s(x+a)^2} e^{-2\pi i b(x+a)}$  (and analytically continue to the complex domain).

Nomenclature.  $\eta(\tau) = e^{2\pi i \tau/24} \prod_{k=1}^{\infty} (1 - e^{2\pi i k \tau})$ .

id co. Fact:  $\eta$  is holomorphic and non-zero in  $\mathbf{H}$ . It satisfies

i.  $\eta(\tau) = \sum_{n \in \mathbf{Z}} (-1)^n e^{3\pi i \tau (n + \frac{1}{6})^2}$

ii.  $\eta(-1/\tau) = (-i\tau)^{1/2} \eta(\tau)$

$$iii. \Delta(\tau) = (2\pi)^{12}\eta(\tau)^{24}$$

id co. Explanation: *i.* follows from the pentagonal identity. To get *ii.*, we apply [cn] with  $\frac{1}{6}, -\frac{1}{2}, 3\tau$  we get

$$\sum_{n \in \mathbf{Z}} e^{\pi i 3\tau(n+\frac{1}{6})^2} e^{\pi i(n+\frac{1}{6})} = \frac{1}{\sqrt{-3i\tau}} \sum_{n \in \mathbf{Z}} e^{-\frac{\pi i}{3\tau}(n-\frac{1}{2})^2} e^{\frac{\pi i n}{3}}$$

The lhs is  $e^{\pi i/6}\eta(\tau)$ . We split the sum in the rhs to  $n = 3k, 3k+1, 3k+2$ . The sum for  $n = 3k+2$  is a constant times  $\sum_k (-1)^k e^{-\frac{\pi i}{3\tau}(3k+3/2)^2}$ , where the  $k$  summand is minus the  $-k-1$  summand, so this sum is zero. Switching  $k$  to  $-k$  in the first sum, we see that it is  $\eta(-1/\tau)$ , and the second sum is  $e^{\frac{\pi i}{3}}\eta(-1/\tau)$ . For *iii.* we see that  $f(\tau) = \frac{\Delta(\tau)}{(2\pi)^{12}\eta(\tau)^{24}}$  is analytic in  $\mathbf{H} \cup i\infty$ , invariant under  $\text{PSL}_2(\mathbf{Z})$ , has no zeros or poles and equals 1 at  $i\infty$ . Such a modular function must be constant 1.

### 3. Number Theory

id ax. Fact: We have  $\text{lcm}(1, \dots, n) \geq 2^{n-2}$  for all  $n$ .

id ax. Explanation: Consider  $I = \int_0^1 t^m(1-t)^m dt$ . We have  $I = \frac{r}{\text{lcm}(1, \dots, 2m+1)}$  for some positive integer  $r$ , as well as  $I \leq 4^{-m}$ . Thus  $4^m \leq \text{lcm}(1, \dots, 2m+1)$ .

## 4. Combinatorics and Probability

id az. Fact: A monkey is typing letters uniformly at random. The the expected number of letters until ABRA-CADABRA is typed equals  $26 + 26^4 + 26^{11}$ .

id az. Explanation. Imagine a casino where gamblers can bet on the next letter. There are 26 letters and the casino gives fair odds, meaning a gambler can pay the casino  $\$x$  dollars and bet that the next letter will be  $y$ . If the bet was right he will get back  $\$26x$ . Now imagine that before every letter, a new gambler arrives with  $\$1$  and bets all their money that the next 11 letters will be ABRACADABRA (of course this will take several turns, the first putting down  $\$1$  on the first letter being A. If the bet works out, the next bet would be  $\$26$  dollars on the next letter being B, and so on). Starting with  $\$0$ , let  $X_n$  be the cash of the casino after the  $n$  letter was typed. Let  $T$  be the stopping time - when ABRACADABRA is first typed. Then  $X_n$  is a martingale with bounded differences,  $\mathbf{E}T < \infty$  and so  $\mathbf{E}X_T = \mathbf{E}X_0 = \$0$ . Let's consider the cash after time  $T$ . Every unsuccessful gambler lost  $\$1$  to the casino. There are 3 successful gamblers, who put in an initial  $\$1$  and took  $26 + 26^4 + 26^{11}$  from the casino. In total,  $X_T = T - 26 + 26^4 + 26^{11}$ .

Nomenclature. In a *branching process* of microbes, the 0-th generation consists of a single microbe, and each microbe gives birth to a number of children, randomly generated from a fixed distribution  $X$  (taking non-negative integer values) independently of all other microbes. In formula,  $Z_0 = 1$ , and  $Z_{n+1} = A_{1,n} + \dots + A_{Z_n,n}$ , where  $A_{i,j} \sim X$  are independent.

id bb. Fact: Let  $(X, Z_n)_{n=0}^\infty$  be a branching process, and set  $f(\theta) = \mathbf{E}[\theta^X]$ ,  $g_n(\theta) = \mathbf{E}[\theta^{Z_n}]$  for  $\theta \in [0, 1]$ . Then  
*i.*  $g_n = f \circ f \cdots \circ f$  is the  $n$ -th composition of  $f$ .  
*ii.* The *probability of extinction*  $p = \mathbf{P}[Z_m = 0 \text{ for some } m]$  is the smallest non-negative fixed point of  $f$ .

id bg. Fact: If the edges of  $K_n$  are exactly covered by the edges of  $m$  complete bi-partite graphs, then  $m \geq n - 1$ .

id bg. Explanation: Suppose  $E(K_n) = \bigsqcup_{k=1}^m E(L_k, R_k)$  with  $m + 1 < n$ . Fix a later chosen  $N$ , and for each mapping  $\sigma : [n] \rightarrow [N]$  let  $P(\sigma) = (\sum_{L_1} \sigma, \dots, \sum_{L_m} \sigma, \sum_{[n]} \sigma)$ . Then  $P$  has domain of size  $N^n$  and range of size at most  $(nN)^{m+1}$ . We choose  $N$  guaranteeing  $P$  is not injective. Fix  $\sigma_1 \neq \sigma_2$  with  $P(\sigma_1) = P(\sigma_2)$ , and set  $\tau = \sigma_1 - \sigma_2$ . We have  $\sum_{i < j} \tau(i)\tau(j) = \sum_{k=1}^m \sum_{L_k} \tau \sum_{R_k} \tau = 0$ , and  $\sum \tau(i) = 0$ . Thus  $\sum \tau(i)^2 = [\sum \tau(i)]^2 - 2 \sum_{i < j} \tau(i)\tau(j) = 0$ , in contradiction to  $\sigma_1 \neq \sigma_2$ .

id bh. Fact: Let  $f$  be a bounded continuous function on  $[0, \infty)$ ,  $\mathcal{L}[f]$  its Laplace transform, let  $X_1, \dots, X_n$  be a sequence of iid variables  $\sim \text{Exp}(\lambda)$  and  $S_n = X_1 + \dots + X_n$ . We have

*i.*  $\mathbf{E}[f(S_n)] = \frac{(-1)^{n-1} \lambda^n}{(n-1)!} \frac{d^{n-1} \mathcal{L}[f](\lambda)}{d\lambda^{n-1}}.$

*ii.* For all  $y > 0$  we have  $f(y) = \lim_{n \rightarrow \infty} \left( \frac{(-1)^{n-1} \lambda^n}{(n-1)!} \frac{d^{n-1} \mathcal{L}[f](\lambda)}{d\lambda^{n-1}} \right) \Big|_{\lambda=n/y}$ , a formula for inverting  $\mathcal{L}[f]$  back to  $f$ .

id bn. Fact: Let  $S_n$  be a symmetric random walk on  $\mathbf{Z}$ ,  $T = \inf\{n : S_n = 1\}$ . Then  $P(T = 2m - 1) = \left(\frac{1}{2}\right) \binom{2m-1}{m} (-1)^{m+1}$ .

id bn. Explanation: We know that  $T < \infty$  almost surely (though  $\mathbf{E}[T] = \infty$ ). Let  $f(\alpha) = \mathbf{E}[\alpha^T]$ . Then  $f(\alpha) = \frac{1}{2}\alpha + \frac{1}{2}\alpha\mathbf{E}[\alpha^{T'}]$  where  $T'$  is the time to get to 1, starting at  $-1$ . Now  $T'$  is a sum of two independent variables, both distributed as  $T$ . Thus  $\mathbf{E}[\alpha^{T'}] = f(\alpha)^2$ . It follows that  $f(\alpha) = \frac{1 - \sqrt{1 - \alpha^2}}{\alpha}$ . It remains to compare coefficients.

id bo. Fact: Let  $Z_n$  be a Markov chain on  $(E, P)$ . Then the following are equivalent.

- i. For any  $i, j \in E$  we have  $\mathbf{P}(\exists n \geq 1 : j = Z_n \mid Z_0 = i) = 1$ .
- ii. All  $P$ -super-harmonic non-negative functions  $h$  on  $E$  are constant.

id bo. Explanation:  $i. \implies ii.$  Clearly  $h(Z_n)$  is a non-negative super-martingale. Starting at  $Z_0 = i$ , let  $T$  be the stopping time of reaching  $j$ , so  $T$  is almost-surely finite, and thus  $h(j)\mathbf{E}[h(Z_T)] \leq h(Z_0) = h(i)$ .

$ii. \implies i.$  For fixed  $j$  the function  $h_j(i) = \mathbf{P}[\text{hitting } j \mid Z_0 = i]$  is  $P$ -super-harmonic. Indeed,  $\sum_k h_j(k)p_{ik} \leq \sum_{k \neq j} h_j(k)p_{ik} + p_{ij} = h_j(i)$ . Therefore  $h_j \equiv c_j$  is constant and  $p_{ij} = p_{ij}c_j$ . If  $p_{ij} = 0$  for all  $i$  then the function  $\delta_j$  would be  $P$ -super-harmonic and non-constant, and so  $c_j = 1$ .

id bs. Fact: Let  $X$  be a random variable with distribution function  $F$  and characteristic function  $\varphi(\theta) = \mathbf{E}[e^{i\theta X}]$ . Then for all  $a < b$

$$\frac{1}{2\pi i} \int_{-T}^T \frac{e^{-i\theta a} - e^{-i\theta b}}{\theta} \varphi(\theta) d\theta \xrightarrow{T \rightarrow \infty} \mathbf{P}[x \in (a, b)] + \frac{1}{2} \mathbf{P}[X \in \{a, b\}]$$

In particular, the characteristic function of  $X$  determines (the distribution of)  $X$ . Specifically, if  $\int_{\mathbf{R}} |\varphi| < \infty$  then  $X$  has a density function  $f$ , and we have

$$f(x) = \frac{1}{2\pi} \int_{\mathbf{R}} e^{-i\theta x} \varphi(\theta) d\theta \quad \varphi(\theta) = \int_{\mathbf{R}} e^{i\theta x} f(x) dx$$

id bv. Fact: Let  $(X_n)$  be a sequence of i.i.d random variables with mean  $\mu$  and variance  $\sigma^2$ . Then for all  $a \in \mathbf{R}$  we have

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(\frac{1}{n}(X_1 + \dots + X_n) - \mu \leq a \frac{\sigma}{\sqrt{n}}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-\frac{1}{2}x^2} dx$$

id bv. Explanation: We first show the point-wise convergence of the characteristic functions  $\varphi_{G_n}(\theta)$  to the characteristic function of the standard normal. Indeed,  $\varphi_{G_n}(\theta) = \varphi\left(\frac{\theta}{\sqrt{n}\sigma}\right)^n$  where  $\varphi$  is the characteristic function of  $X_i - \mu$ . Recall [??] that  $\varphi(\alpha) = 1 - (\frac{1}{2}\sigma^2 + o(1))\alpha^2$  as  $\alpha \rightarrow 0$ . Thus  $\varphi_{G_n}(\theta) = \left(1 - \frac{\theta^2/2 + o(1)}{n}\right)^n$  converges to  $e^{-\frac{1}{2}\theta^2}$  as  $n$  approaches infinity [bq]. It remains to show the following.

Nomenclature. A sequence  $(F_n)$  of distribution functions *weakly converge* to a distribution function  $F$  if

$$\lim_{n \rightarrow \infty} F_n(x) = F(x)$$

for all points of continuity  $x$  of  $F$ .

id bw. Fact: Let  $F_n$  be a sequence of distribution functions. Assuming the point-wise convergence of the characteristic functions  $\varphi_{F_n}(\theta)$  to a function  $g(\theta)$  continuous at 0. Then  $F_n$  weakly converge to a distribution function  $F$  and  $g(\theta) = \varphi_F(\theta)$  is its characteristic function.

## 5. Abstract Algebra

id bm. Fact: For a subset  $S \subseteq \mathbf{F}_2^d$  let  $\Gamma_S(x, y) = \sum_{s \in S} x^{d-W(s)} y^{W(s)}$  where  $W(s) = |i \in [d] : s_i = 1|$  is the Hamming weight. If  $U$  is a subspace of  $\mathbf{F}_2^d$  and  $W = \{v \in \mathbf{F}_2^d : u^t v = 0 \ \forall u \in U\}$  is the "orthogonal" subspace, then

$$\Gamma_W(x, y) = \frac{1}{|U|} \Gamma_U(x + y, x - y)$$

In particular, the multiset of Hamming weights of  $U$  is sufficient information for that of  $W$ .

id bm. Explanation: Let  $f(s) = x^{d-W(s)} y^{W(s)}$  be a function on  $\mathbf{F}_2^d$  and  $\psi$  be a linear character of  $\mathbf{F}_2^d$ . We have  $\psi = \psi_v : x \mapsto (-1)^{v^t x}$  for some  $v \in \mathbf{F}_2^d$ . Thus  $2^d \hat{f}(\psi) = \sum_{s \in \mathbf{F}_2^d} x^{d-W(s)} y^{W(s)} (-1)^{v^t s} = \prod_{i=1}^d \sum_{s_i=0}^1 x^{1-s_i} y^{s_i} (-1)^{v_i s_i} = \prod_{i=1}^d (x + y)^{1-v_i} (x - y)^{v_i} = (x + y)^{d-W(v)} (x - y)^{W(v)}$ . It remains to utilize Poisson summation formula.

id ba. Fact: If  $A$  is a Noetherian ring then  $A[t]$  is a Noetherian ring.

id ba. Explanation: Fix an ideal  $\mathcal{J} \trianglelefteq A[t]$ . A best effort attempt to generate  $\mathcal{J}$  is as follows: as long as  $f_1(t), \dots, f_{n-1}(t)$  don't generate  $\mathcal{J}$ , let  $f_n(t)$  be the polynomial of minimal degree in  $\mathcal{J}$  and not in  $\langle f_1(t), \dots, f_{n-1}(t) \rangle$ . Note that  $d_n = \deg f_n(t)$  is weakly increasing, and that the leading coefficient of  $f_n(t)$  does not belong to the ideal of  $A$  generated by the leading coefficients of  $f_1(t), \dots, f_{n-1}(t)$ , for otherwise there are  $a_1, \dots, a_{n-1} \in A$  for which  $g(t) = f_n(t) - a_1 f_1(t) t^{d_n-d_1} - \dots - a_{n-1} f_{n-1}(t) t^{d_n-d_{n-1}}$  has degree smaller than  $f_n(t)$  and is also member of  $\mathcal{J} \setminus \langle f_1(t), \dots, f_n(t) \rangle$ , which can't be. It follows that the sequence of ideals of  $A$  generated by the leading coefficients of  $f_1(t), \dots, f_n(t)$  is strictly increasing, and  $A$  is not Noetherian.

id be. Fact: In a vector space, if  $(u_1, \dots, u_i)$  is linearly independent and  $(w_1, \dots, w_s)$  is spanning, then  $i \leq s$ . Moreover, we may replace some  $i$  members of the latter with the members of the former to produce a spanning set.

id be. Explanation: By induction on  $i$ , the base  $i = 0$  being trivial. For the step, we have wlog  $(u_j)_{j < i} \cup (w_k)_{k \geq i}$  spanning. Writing  $u_i$  as a linear combination, a coefficient of one of the  $w$ 's must be non-zero, wlog say that of  $w_i$ . Thus  $w_i$  is spanned by  $(u_j)_{j \leq i} \cup (w_k)_{k > i}$ . This set is spanning, since it contains the remaining members of the previously spanning set.

We also give specific explanations for when  $\mathbf{F}$  is a finite field and when  $\mathbf{F} = \mathbf{Q}$ .

id be. Explanation for finite fields: If  $|\mathbf{F}| = q$  is finite, then the span of  $u_1, \dots, u_i$  has precisely  $q^i$  elements, where as the span of  $w_1, \dots, w_s$  has at most  $q^s$  elements.

id be. Explanation for  $\mathbf{F} = \mathbf{Q}$ : Write  $u_t = \sum_j \alpha_{tj} w_j$  for  $\alpha_{tj} \in \mathbf{Q}$ ,  $t \in [i]$ ,  $j \in [s]$ . Let  $M$  be a common multiple of the denominators,  $f_t = M u_t$  and  $\beta_{tj} = M \alpha_{tj} \in \mathbf{Z}$ . Let  $R$  be a positive integer variable,  $A_R = \{r_1 f_1 + \dots + r_i f_i : r_t \in \mathbf{Z}, |r_t| \leq R\}$  and  $B_R = \{m_1 w_1 + \dots + m_s w_s : m_j \in \mathbf{Z}, |m_j| \leq CR\}$  where  $C = \max_j \sum_t |\beta_{tj}|$ . We have  $A_R \subseteq B_R$ , and with the polynomial inequality  $(2R + 1)^i = |A_R| \leq |B_R| \leq (2CR + 1)^s$  valid for all values of  $R$  we deduce  $i \leq s$ .

id bf. Fact: If  $f : \mathbf{F}_2^d \rightarrow \pm 1$  satisfies  $f(x+y) = f(x)f(y)$  in probability  $\geq 1 - \varepsilon$ , then there exists a character  $\phi : \mathbf{F}_2^d \rightarrow \pm 1$  with  $f = \phi$  in probability  $\geq 1 - \varepsilon$ .

id bf. Explanation: Write  $f = \sum \hat{f}(\psi)\psi$ . We have

$$4^d(1 - 2\varepsilon) \leq \sum_{x,y} f(x+y)f(x)f(y) = \sum_{x,y,\psi_1,\psi_2,\psi_3} \hat{f}(\psi_1)\hat{f}(\psi_2)\hat{f}(\psi_3)\psi_1(x+y)\psi_2(x)\psi_3(y) = 4^d \sum_{\psi} [\hat{f}(\psi)]^3$$

by character orthogonality. Moreover  $1 = \langle f, f \rangle = \sum_{\psi} [\hat{f}(\psi)]^2$ . Noting that the  $\hat{f}(\psi)$  are real, we have  $\max_{\psi} \hat{f}(\psi) \geq \frac{\sum_{\psi} [\hat{f}(\psi)]^3}{\sum_{\psi} [\hat{f}(\psi)]^2} \geq 1 - 2\varepsilon$  and so for  $\phi = \arg \max \hat{f}$  we have  $f = \phi$  in probability  $\geq 1 - \varepsilon$ .

id bl. Fact: Let  $G$  be a finite abelian group. For a subgroup  $H$  let  $H^\perp = \{\phi \in \widehat{G} : \phi|_H \equiv 1\}$ . Then

i.  $H^\perp$  is a subgroup naturally isomorphic to  $\widehat{G/H}$ .

ii.  $\widehat{G}/H^\perp$  is naturally isomorphic to  $\widehat{H}$ .

iii.  $H^{\perp\perp}$  is naturally isomorphic to  $H$ .

iv. We have  $\sum_{\psi \in H^\perp} \psi(g) = \begin{cases} [G:H] & g \in H \\ 0 & g \notin H \end{cases}$

v. *Poisson summation formula*: For  $f : G \rightarrow \mathbf{C}$  we have

$$\sum_{h \in H} f(h) = |H| \sum_{\psi \in H^\perp} \hat{f}(\psi)$$

id bl. Explanation:

iv. This follows from i. and character orthogonality.

v. We may assume  $f$  is a character of  $G$ . If  $f \in H^\perp$  the formula is clear. If  $f \notin H^\perp$  then  $f|_H$  is a non-trivial character on  $H$  so that the lhs vanishes, as does the right.

Nomenclature. A society has voters and politicians. A *voting system* is a mapping, taking in a linear preference of the politicians from each voter, and returning a linear preference of the politicians. The *majority* system works if there's an odd number of voters and two politicians, and ranks on top the politician the majority of voters ranked on top. The majority system is pretty fair, and on the other spectrum we have a *dictatorship*: a system which always returns the preferences of one specific voter, the dictator.

id bp. Fact: Suppose a voting system on three politicians satisfies the following:

i. If all the voters rank one politician higher than another, so will the system.

ii. For any pair of politicians, the preference between the two returned by the system depends only on the preferences of the voters between the two.

iii. If the voters choose their preferences (uniformly) at random, then any one of politician has even odds at ranking higher than any other.

Then the system is a dictatorship.

id bp. Explanation: We first need a model. If for example there's four voters, the three politicians are named  $a, b, c$  and the preferences are  $a > b > c$ ,  $c > a > b$ ,  $b > a > c$ ,  $b > c > a$  then the preferences of the voters between any pair of pair of politicians are as follows:

$$\begin{bmatrix} a : b & a > b & a > b & b > a & b > a \\ b : c & b > c & c > b & b > c & b > c \\ c : a & a > c & c > a & a > c & c > a \end{bmatrix}$$

For each row, the voting system returns a preference of the two politicians. In other words, the system is composed of three functions taking in  $n$  signs (where  $n$  is the number of voters) and returning one sign. In our

example, the preferences are described as the sign vectors

$$\begin{bmatrix} x : & +1 & +1 & -1 & -1 \\ y : & +1 & -1 & +1 & +1 \\ z : & -1 & +1 & -1 & +1 \end{bmatrix}$$

and assumption *ii.* implies that the voting system returns a triple  $(f(x), g(y), h(z))$  describing the final preferences  $(a : b, b : c, c : a)$ . It is important to see that each of  $f, g, h$  is defined on the full space  $\{\pm 1\}^n$  but out of the  $8^n$  triples  $x, y, z$ , only  $6^n$  describe potential preferences of voters, namely for each  $i \in [n]$  the three signs  $x_i, y_i, z_i$  cannot all be the same. We draw at random preferences for the voters (out of a space of  $6^n$  possibilities), described as sign vectors  $x, y, z$ . The three signs  $f(x), g(y), h(z)$  are not all the same, meaning

$$f(x)g(y) + g(y)h(z) + h(z)f(x) = -1$$

is a constant random variable. On the other hand

$$\mathbf{E}[f(x)g(y)] = \sum_{\psi_1, \psi_2 \in \widehat{\{\pm 1\}^n}} \hat{f}(\psi_1) \hat{g}(\psi_2) \mathbf{E}[\psi_1(x) \psi_2(y)]$$

Now a character  $\psi$  corresponds to a set  $S$  via  $\psi(x) = \prod_{s \in S} x_s$ . If  $\psi_1 \neq \psi_2$  then from  $\mathbf{E}[\psi_1(x) \psi_2(y)]$  we may factor out a zero in the form of  $\mathbf{E}x_j$  or  $\mathbf{E}y_j$ . Since  $\mathbf{E}[x_j y_j] = -\frac{1}{3}$  we have  $\mathbf{E}[f(x)g(y)] = \sum_{S \subseteq [n]} \hat{f}(S) \hat{g}(S) \left(-\frac{1}{3}\right)^{|S|}$ . In total we have

$$1 = \sum_{S \subseteq [n]} -\left(-\frac{1}{3}\right)^{|S|} \left(\hat{f}\hat{g} + \hat{g}\hat{h} + \hat{h}\hat{f}\right)(S)$$

Now  $\hat{f}(\emptyset) = \hat{g}(\emptyset) = \hat{h}(\emptyset) = 0$  by assumption *iii.* Moreover  $|r_1 r_2 + r_2 r_3 + r_3 r_1| \leq r_1^2 + r_2^2 + r_3^2$  for real  $r_i$  so

$$1 \leq \sum_{\emptyset \neq S \subseteq [n]} \frac{\hat{f}(S)^2 + \hat{g}(S)^2 + \hat{h}(S)^2}{3^{|S|}} \leq \frac{\langle f, f \rangle + \langle g, g \rangle + \langle h, h \rangle}{3} = 1$$

implying the only non-trivial coefficients are singletons, so  $f, g, h$  are linear functions. Since they only take two values, they're of the form  $f(x) = \pm x_j$ ,  $g(y) = \pm y_k$ ,  $h(z) = \pm z_\ell$ . Assumption *i.* yields positive signs, and it's easy to see  $j = k = \ell$  is a dictator.

id bt. Fact: The set of homomorphisms  $G \rightarrow \mathbf{F}^\times$  is a linear independent set in  $G \rightarrow \mathbf{F}$ . In particular, if  $G$  is finite then  $G$  admits at most  $|G|$  homomorphisms to  $\mathbf{F}^\times$ .

id bt. Explanation: Let  $\phi_1, \dots, \phi_n$  be distinct homomorphisms and suppose  $\sum_i \alpha_i \phi_i \equiv 0$ . We have

$$\sum_i \alpha_i \phi_i(h) \phi_i(g) = \sum_i \alpha_i \phi_i(hg) = 0 = \sum_i \alpha_i \phi_i(g) = \sum_i \alpha_i \phi_n(h) \phi_i(g) =$$

We may inductively assume  $\phi_1, \dots, \phi_{n-1}$  are linearly independent. This yields  $\alpha_i \phi_i(h) = \alpha_i \phi_n(h) \forall i \in [n-1]$ ,  $h$  so  $\alpha_1 = \dots = \alpha_{n-1} = 0 \implies \alpha_n = 0$ .

id bz. Fact: The groups  $\text{GL}(3, 2)$  and  $\text{PSL}(2, 7)$  are isomorphic.

id bz. Explanation: Both groups have order 168. We shall find an isomorphism  $\Psi : \text{PSL}(2, 7) \rightarrow \text{GL}(3, 2)$ . As a 3-dimensional  $\mathbf{F}_2$  vector space we take  $\mathbf{F}_8$ , realized as  $\mathbf{F}_2[t]/\langle t^3 + t + 1 \rangle$  with  $\alpha = \bar{t}$  being a generator for  $\mathbf{F}_8^\times$ . Thus  $\mathbf{F}_7 \cup \{\infty\} \longleftrightarrow \mathbf{F}_8 = \{\alpha^i\}_{i=0}^6 \cup \{0\}$  via  $i \longleftrightarrow \alpha^i$ , writing  $\alpha^\infty = 0$ . Given a Möbius transformation  $h$  of  $\mathbf{F}_7 \cup \{\infty\}$ , set  $T_h = \Psi(h) : \mathbf{F}_8 \rightarrow \mathbf{F}_8$ ,  $\alpha^i \mapsto \alpha^{f(i)} + \alpha^{f(\infty)}$ . Let us consider 3 examples. If  $h(z) = 2z$  then  $T_h(\gamma) = \gamma^2$ . If  $h(z) = z + 1$  then  $T_h(\gamma) = \alpha\gamma$ . If  $h(z) = -1/z$  then  $T_h(b_0 + b_1\alpha + b_2\alpha^2) = b_0 + b_2\alpha + b_1\alpha^2$ . The three Möbius transformations generate  $\text{PSL}(2, 7)$ , their images under  $\Psi$  are additive functions of  $\mathbf{F}_8$ . We have that  $\Psi$  is a homomorphism, since if  $f, g \in \text{PSL}(2, 7)$  are such that  $\Psi_f, \Psi_g$  are additive, then by definition  $\Psi_{f \circ g} = \Psi_f \circ \Psi_g$ . Finally,  $\Psi$  is a bijection, either because  $\text{PSL}(2, 7)$  is simple and  $\Psi$  is non-trivial, or because the above 3 transformations generate  $\text{GL}(3, 2)$ .



id bz. Explanation 2: The group  $\text{PSL}(2, 7)$  acts on size 4 subsets of  $\mathbf{F}_7 \cup \{\infty\}$ . The orbit of  $\{0, 1, 3, \infty\}$  is:

$\{0, 1, 3, \infty\}, \{2, 4, 5, 6\}$   
 $\{1, 2, 4, \infty\}, \{3, 5, 6, 0\}$   
 $\{2, 3, 5, \infty\}, \{4, 6, 0, 1\}$   
 $\{3, 4, 6, \infty\}, \{5, 0, 1, 2\}$   
 $\{4, 5, 0, \infty\}, \{6, 1, 2, 3\}$   
 $\{5, 6, 1, \infty\}, \{0, 2, 3, 4\}$   
 $\{6, 0, 2, \infty\}, \{1, 3, 4, 5\}$

These 14 sets form a block design - every 3 elements of  $\mathbf{F}_7 \cup \{\infty\}$  belong to precisely one set. In fact, the 14 affine planes of  $\mathbf{F}_2^3$  form an isomorphic design. For example we can match

$\infty \longleftrightarrow (0, 0, 0)$   
 $0 \longleftrightarrow (1, 0, 0)$   
 $1 \longleftrightarrow (0, 1, 0)$   
 $2 \longleftrightarrow (0, 0, 1)$   
 $3 \longleftrightarrow (1, 1, 0)$   
 $4 \longleftrightarrow (0, 1, 1)$   
 $5 \longleftrightarrow (1, 1, 1)$   
 $6 \longleftrightarrow (1, 0, 1)$

Now, the group of automorphisms of the design is clearly the affine group  $G = \{v \mapsto Av + b : A \in \text{GL}(3, 2), b \in \mathbf{F}_2^3\}$  and so we have a homomorphism  $\text{PSL}(2, 7) \rightarrow G$ . Composing with  $G \rightarrow \text{GL}(3, 2)$  (via  $Av + b \mapsto A$ ) we have  $\text{PSL}(2, 7) \rightarrow \text{GL}(3, 2)$ . But both groups have the same order (168), the first is simple, and the homomorphism is non-trivial, so it is an isomorphism.

id cd. Fact: Any symmetric polynomial in  $x_1, \dots, x_n$  can be expressed as a polynomial in the esp's  $e_1, \dots, e_n$ .

id cd. Explanation: We give a recursive algorithm. If  $n = 1$  then  $e_1 = x_1$  and we're done. If  $P(x_1, \dots, x_n)$  is our symmetric polynomial, then  $Q(x_1, \dots, x_{n-1}) = P(x_1, \dots, x_{n-1}, 0)$  [namely the constant term of  $P$  when expressed as a polynomial of  $x_n$ ] is a symmetric polynomial of fewer variables, and we may express it as some polynomial  $f$  of the esp's  $e_1, \dots, e_{n-1}$  of  $x_1, \dots, x_n$ . A neat observation is that  $e_i(x_1, \dots, x_{n-1}) = e_i(x_1, \dots, x_{n-1}, 0)$ , which suggests to look at  $D(x_1, \dots, x_n) = P - f(e_1(x_1, \dots, x_n), \dots, e_{n-1}(x_1, \dots, x_n))$ . Now  $D$  is symmetric in all  $n$  variables, and  $D(x_1, \dots, x_{n-1}, 0) \equiv 0$ . This means  $x_1$  divides  $D$ , and so  $x_1 \dots x_n$  divides  $D$ . In other words  $D = e_n R$  for some symmetric polynomial  $R$  of smaller degree than  $P$ .

## 6. Analysis

Nomenclature. A function  $f$  defined on the vertices of a graph will be called *harmonic* if for every vertex  $v$  the average value of  $f(u)$  over the neighbours  $u$  of  $v$  equals  $f(v)$ .

id ae. Fact: Let  $f$  be a real valued harmonic function on the lattice  $\mathbf{Z}^d$ . If  $f$  is non-constant, then it is unbounded.

id ae. Explanation: If  $f$  is non-constant, then wlog the bounded harmonic function  $g(x) = f(x + e_1) - f(x)$  attains a positive value. Let  $S = \sup g > 0$ . Given  $\varepsilon$  we may find  $x_\varepsilon$  for which  $g(x_\varepsilon) > S - \varepsilon$ . Since the average of  $g$  over the  $2d$  neighbours of  $x_\varepsilon$  is larger than  $S - \varepsilon$ , and each value is at most  $S$ , it follows that  $g$  assigns each neighbour a value greater than  $S - 2d\varepsilon$ . Thus  $g(x_\varepsilon + e_1) > S - 2d\varepsilon$ , and generally  $g(x_\varepsilon + ke_1) > S - (2d)^k\varepsilon$ . However, if we pick  $k$  large enough and  $\varepsilon$  small enough, we can make  $f(x + (k+1)e_1) - f(x) = g(x) + g(x + e_1) + \dots + g(x + ke_1) > (k+1)(S - (2d)^k\varepsilon)$  arbitrarily large, meaning  $f$  is unbounded.

id af. Fact: Let  $T$  be a strict contraction of a complete metric space  $M$ . Then in  $M$  there exists a unique fixed point  $m^*$  of  $T$ . Moreover, for any  $m \in M$  we have  $\lim_{n \rightarrow \infty} T^n(m) = m^*$ .

id af. Explanation: A strict contraction clearly cannot have more than one fixed point. Let  $m$  be an arbitrary point, and let  $m_n = T^n(m)$  be its orbit under  $T$ . We have  $d(m_n, m_{n+1}) \leq q^n d(m, Tm)$ , where  $q$  is the contraction constant for  $T$ . Thus  $\sum d(m_n, m_{n+1}) \leq \frac{d(m, Tm)}{1-q}$  is convergent, implying  $m_n$  is Cauchy, and hence convergent. By continuity, the limit of the orbit is a fixed point.

id av. Fact: The Fourier transform is injective on  $\mathcal{L}^1(\mathbf{R})$ .

id av. Explanation: Let  $f \in \mathcal{L}^1, \hat{f} \equiv 0$ . For all  $a, \xi \in \mathbf{R}$  we have

$$\int_{-\infty}^a f(t) e^{i\xi(t-a)} dt = - \int_a^{\infty} f(t) e^{i\xi(t-a)} dt$$

and we denote this value by  $F_a(\xi)$ . The right hand side may be analytically continued to arguments in the closed upper half plane, and the left hand side may be continued to the closed lower half plane. It follows that  $F_a(\xi)$  is an entire function. However, it is everywhere bounded by  $\|f\|_1$ , and so must be constant. By dominated convergence we have  $\lim_{r \rightarrow \infty} F_a(ir) = 0$ , and so  $F_a(\xi) = 0$ . In particular  $\int_{-\infty}^a f(t) dt = F_a(0) = 0$  for all  $a$  and so  $f = 0$  almost everywhere.

id ag. Fact: Let  $f \in \mathcal{C}(\mathbf{T})$  and let  $F_m(t) = \sum_{j=-m}^m \hat{f}(j) e^{2\pi i j t}$  be the Fourier approximations of  $f$ . Then the sequence of averages  $\sigma_n(t) = \frac{F_0(t) + \dots + F_n(t)}{n+1}$  converges uniformly to  $f$ .

id aq. Explanation: We have  $\hat{f}(j) e^{2\pi i j t} = \int_0^1 f(x) e^{2\pi i j(t-x)} dx = \int_{-t}^{1-t} f(u+t) e^{-2\pi i j u} du = \int_0^1 f(u+t) e^{-2\pi i j u} du$

and therefore

$$F_m(t) = \int_0^1 f(u+t)P_m(u)du \quad \sigma_n(t) = \int_0^1 f(u+t)D_n(u)du$$

For

$$P_m(u) = \sum_{j=-m}^m e^{2\pi i j u} \quad D_n(u) = \frac{P_0(u) + \cdots + P_n(u)}{n+1}$$

We have the identities  $\int_0^1 D_n(u) = 1$  and  $(n+1)D_n(u) = \left( \sum_{k=0}^n e^{2\pi i (k-\frac{n}{2})u} \right)^2 = \left( \frac{\sin(\pi(n+1)u)}{\sin(\pi u)} \right)^2 \geq 0$ . (To clarify,  $D_n(u) = n+1$  for  $u \in \mathbf{Z}$ ). We continue

$$\sigma_n(t) - f(t) = \int_0^1 [f(u+t) - f(t)]D_n(u)du \implies |\sigma_n(t) - f(t)| \leq \int_0^1 |f(u+t) - f(t)| D_n(u)du$$

Given  $\varepsilon$  we find  $\delta$  such that  $|x-y| \leq \delta \implies |f(x) - f(y)| \leq \varepsilon$ . The first part of the integral is bounded independent of  $n$  or  $t$ :  $\int_{-\delta}^{\delta} |f(u+t) - f(t)| D_n(u)du \leq \varepsilon \int_0^1 D_n(u)du = \varepsilon$ . Finally, on  $[\delta, 1-\delta]$  we have  $D_n(u) \leq \frac{1}{(n+1)\sin(\pi\delta)^2}$  and so if  $n$  is large enough then  $D_n(u) \leq \varepsilon$  on this segment, and in total  $|\sigma_n(t) - f(t)| \leq \varepsilon + 2\|f\|_{\infty}\varepsilon$  independent of  $t$ .

id bc. Fact: Let  $b, m, n$  be elements of a  $C^*$  algebra with  $m, n$  normal and  $bn = mb$ . Then  $bn^* = m^*b$ .

id bc. Explanation: We have  $bn^k = m^k b$ , and so  $bf(n) = f(m)b$  for analytic functions  $f$ . In particular  $b = \exp(\xi m)b\exp(-\xi n)$  for all  $\xi \in \mathbf{C}$ . Normality yields

$$\exp(i\zeta m^*)b\exp(-i\zeta n^*) = \exp(i(\zeta m^* + \zeta^* m))b\exp(-i(\zeta^* n - \zeta n^*))$$

The rhs is an entire function of  $\zeta$ , and the lhs is bounded by  $\|b\|$ , and so the expression is a constant, namely  $b$ . Finally, the lhs coefficient of  $\zeta$  is  $im^*b - ibn^*$ .

id bj. Fact: For a box  $I \in \mathbf{R}^d$ , not necessarily parallel to the standard axes, write  $\rho(I)$  for the sum of the  $d$  edge lengths of  $I$ . If  $I_1 \subseteq I_2$  then  $\rho(I_1) \leq \rho(I_2)$ .

id bj. Explanation: This is easy for  $d=1, 2$ . For  $d \geq 3$ , we consider the volume of the set of points of distance  $\leq r$  from  $I$ , denoted  $f(r; I)$ , where we take  $I$  as a solid.  $f(r; I) = c_0 r^d + c_1 \rho(I) r^{d-1} + \cdots + \text{Vol}(I)$  is a polynomial in  $I$  of degree  $d$ , and clearly  $f(r; I_1) \leq f(r; I_2)$  for all  $r$ .

id bk. Fact: Let  $I, I_1, I_2, \dots$  be intervals with  $I \subseteq \bigcup_k I_k$ . Then  $\text{len}(I) \leq \sum_k \text{len}(I_k)$ .

id bk. Explanation: We first assume  $I$  is compact and all the  $I_k$  are open. Thus  $I \subseteq I_1 \cup \cdots \cup I_N$  for some  $N$ , and we'll show  $\text{len}(I) \leq \text{len}(I_1) + \cdots + \text{len}(I_N)$  by induction on  $N$ . The base case is trivial. For the step, assume wlog  $m = \min I \in I_1 = (\alpha, \beta)$ . Then  $[\beta, \max I] \subseteq I_2 \cup \cdots \cup I_N \implies \text{len}(I) < \text{len}(I_1) + \text{len}([\beta, \max I]) \leq \text{len}(I_1) + \cdots + \text{len}(I_N)$ . A second step is to remove the openness assumption on the  $I_k$ , but keep the compactness assumption on  $I$ . Let  $J_k$  be the interior of  $I_k$ , and for some fixed  $\varepsilon$  let  $L_k, R_k$  be open intervals of length  $\varepsilon/2^k$  around the left and right endpoints of  $I_k$  (if one of the  $I_k$  is infinite there's nothing to prove). We get that  $\text{len}(I) \leq \sum \text{len}(I_k) + 2\varepsilon$ , and the desired inequality follows in the limit. The final step of removing the compactness assumption on  $I$  is by an approximation of  $I$  using compact intervals.

## Proving the central limit theorem

We embark on a quest to prove the central limit theorem [bv]. Our best tool around independent random variables is interchanging expectation and products. The theorem has to do with sums rather than products, and this suggests raising an exponential.

Nomenclature. The *characteristic function* of a random variable  $X$  is  $\varphi(\theta) = \mathbf{E}[e^{i\theta X}] = \mathbf{E} \cos(\theta X) + i\mathbf{E} \sin(\theta X)$ .

id by. Examples:

$$i. X \sim U[-1, 1] \implies \varphi(\theta) = \frac{\sin(\theta)}{\theta}$$

$$ii. f(x) = \frac{1}{\pi(1+x^2)} \implies \varphi(\theta) = e^{-|\theta|}$$

$$iii. X \sim N(0, 1) \implies \varphi(\theta) = e^{-\frac{1}{2}\theta^2}$$

$$iv. X \sim \text{Exp}(1) \implies \varphi(\theta) = \frac{1}{1-i\theta}$$

$$v. f(x) = \frac{e^{-|x|}}{2} \implies \varphi(\theta) = \frac{1}{1+\theta^2}$$

$$vi. f(x) = (1 - |x|)1_{|x| \leq 1} \implies \varphi(\theta) = 2 \frac{1 - \cos(\theta)}{\theta^2}$$

$$vii. f(x) = \frac{1 - \cos x}{\pi x^2} \implies \varphi(\theta) = (1 - |\theta|)1_{|\theta| \leq 1}$$

id bx. Fact: The characteristic function  $\varphi = \varphi_X$  is a well-defined continuous function  $\mathbf{R} \rightarrow \mathbf{C}$ , bounded in absolute value by 1, and mapping 0 to 1. We have

$$\varphi_{aX+b}(\theta) = e^{i\theta b} \varphi_X(a\theta)$$

and

$$\varphi_{X+Y}(\theta) = \varphi_X(\theta) \varphi_Y(\theta)$$

if  $X, Y$  are independent.

# -1. Solutions

id aa. The probability is  $1/101$ , independent of  $k = 0, \dots, 100$ . To show this, note that we may generate a biased coin toss by generating a uniform random number  $t$  from  $[0, 1]$  and then show heads iff  $t < p$ . Therefore the process can be described by generating  $1 + 100$  independent uniform random numbers  $p = t_0, t_1, \dots, t_{100}$  from  $[0, 1]$  and the number of heads is the index of  $t_0$  after sorting, which is distributed uniformly in  $0, \dots, 100$ .

id ab. First consider the case  $n = 2$ . The two players have numbers in  $0, 1$ , and see the other person's number. To guarantee a correct guess, one will guess they have the same number, and the other will guess they have different numbers. We generalize this to arbitrary  $n$  by assigning the players id's  $0, 1, \dots, n - 1$ , and the player with id  $i$  guessing that the sum of all their numbers is congruent to  $i$  modulo  $n$ . This strategy guarantees exactly one of them will be right, namely the one with id equal to the sum of their numbers modulo  $n$ .

id ad. Let  $e_n$  denote the  $n$ -th coordinate sequence, set  $\alpha_n = f(e_n)$ . We'll show that  $\alpha_{N+1} = \alpha_{N+2} = \dots = 0$  after some value of  $N$ . To do this, we form a sequence  $b = (b_1, b_2, b_3, \dots)$  as follows:  $b_1 = 1$ , and  $b_{k+1}$  is a power of 2 greater than both  $b_k$  and  $2|f(b_1e_1 + \dots + b_ke_k)|$ . Note that  $b_k \mid b_{k+1}$  and  $\lim b = \infty$ . Set  $B = f(b)$ . We have  $B = f(b_1e_1 + \dots + b_ke_k) + b_{k+1}f(0, 1, b_{k+2}/b_{k+1}, b_{k+3}/b_{k+1}, \dots)$ . If  $k$  is sufficiently large, we must have  $B = f(b_1e_1 + \dots + b_ke_k)$ , or otherwise  $|B| \geq b_{k+1} - |f(b_1e_1 + \dots + b_ke_k)| \geq b_{k+1}/2$ . Therefore, if  $k$  is sufficiently large,  $\alpha_k = f(e_k) = 0$ . Now we may consider  $g(x) = f(x) - \sum \alpha_i x_i$  as an additive function satisfying  $g(e_n) = 0$  for all  $n$ , and we wish to show  $g(x) \equiv 0$ . Write  $x_n = 2^n d_n + 3^n r_n$  for some integer sequences  $d, r$ . We have  $g(x) = g(2d_1, 4d_2, 8d_3, \dots) + g(3r_1, 9r_2, 27r_3, \dots)$ , but since  $g(e_n) = 0$ , the left summand is divisible by all powers of two, and the right summand is divisible by all powers of three, so they are both zero, and  $g(x) \equiv 0$ .

id ah. Subtracting a linear factor, we may assume  $f(0) = f(1) = 0$ , with the intention of showing  $f \equiv 0$ . Let  $f_\varepsilon(t) = f(t) - \varepsilon t(1 - t)$ . Then  $\lim_{h \rightarrow 0} \frac{f_\varepsilon(t_0 + h) - 2f_\varepsilon(t_0) + f_\varepsilon(t_0 - h)}{h^2} = \varepsilon$ , implying  $f_\varepsilon$  does not attain its maximum at any inner point  $t_0 \in (0, 1)$ . Therefore  $f_\varepsilon \leq 0$ , and in the limit  $f \leq 0$ . Applying the same to  $-f$  we have  $f \equiv 0$ .

id ai. What demands explanation is that all smooth/continuous homomorphisms are exponentials. Indeed, for  $\phi$  smooth

$$\phi(s + t) = \phi(s)\phi(t) \implies \phi'(s + t) = \phi'(s)\phi(t) \implies \phi'(t) = \phi'(0)\phi(t) \implies \phi(t) = \exp(t\phi'(0))$$

It remains to show that a continuous homomorphism is smooth. Indeed, we have  $\int_x^{x+a} \phi(t)dt = \phi(x) \int_0^a \phi(t)dt$ . We pick a small enough  $a$  so that  $\frac{1}{a} \int_0^a \phi(t)dt = \text{id} + o(1)$  is invertible, yielding

$$\phi(x) = \left( \int_x^{x+a} \phi(t)dt \right) \left( \int_0^a \phi(t)dt \right)^{-1}$$

is smooth.

id aj. Suppose  $\sum a_n^2$  diverges, and form a partition  $0 = N_0 < N_1 < \dots$  for which  $s_k = \sum_{(N_{k-1}, N_k]} a_n^2 > 1$ . The sequence  $b_n = a_n/k s_k$  (where  $n \in (N_{k-1}, N_k]$ ) is in  $\ell^2$ , since  $\sum b_n^2 = \sum_k \frac{1}{k^2 s_k}$ . However,  $\sum a_n b_n = \sum_k 1/k$  diverges.

id aw. The number 100 can only be the top card once, after that it will stay forever at the bottom. Between any two times the number 99 is on top, 100 has to be on top. Between any two times the number 98 is on top, either 100 or 99 has to be on top, ..., between any two times the number 2 is on top a number bigger than 2 has to be on top. Therefore each number except for 1 can only be on top finitely many times.

id aw. Solution 2. If 100 ever becomes the top card, it'll later be stuck at the bottom. If not, the bottom card will never change, and the game will be indifferent to switching the numbers of the bottom card with 100. So in either case we have an induction step.

id aw. Solution 3. Let the score of a deck be the sum of  $2^i$  for the cards  $i$  placed correctly at  $i$ . If the top card  $k$  is not one, the score grows: a  $2^k$  is added and at worst  $1 + 2 + \dots + 2^{k-1}$  is subtracted. Since the score can't grow forever, eventually the top card becomes 1.

id aw. Solution 4. Eventually the state of the deck enters a loop. Once we're in the loop, the maximal number appearing on top has to be 1, otherwise the remaining numbers on top would keep it away from coming back on top.

id bh. Explanation:

i. we have  $\mathbf{E}[f(S_n)] = \int_{x_i \geq 0} f(\sum x_i) \lambda^n e^{-\lambda \sum x_i} dx_1 \dots dx_n$ . Change variables to  $y_i = x_i$  for  $i < n$  and  $t = \sum x_i$ . The Jacobian is 1, the domain  $0 \leq y_i, \sum y_i \leq t$ , and we get  $\lambda^n \int_{t=0}^{\infty} dt f(t) e^{-\lambda t} \int_{y_i \geq 0, \sum y_i \leq t} dy_1 \dots dy_{n-1}$ . We recognise the inner volume as  $\frac{t^{n-1}}{(n-1)!}$ . Finally,  $\frac{d^{n-1} \mathcal{L}[f](\lambda)}{d\lambda^{n-1}} = \int_0^{\infty} (-x)^{n-1} e^{-\lambda x} f(x) dx$ .

ii. We have  $\mathbf{E}[X] = \lambda^{-1}$ ,  $\text{Var} X = \lambda^{-2}$ , so for  $\lambda = n/y$  we have  $\mathbf{E}[S_n] = y$ ,  $\text{Var}(S_n) = y^2/n$ . Chebyshev gives  $\mathbf{P}[|S_n - y| > yn^{-1/4}] \leq n^{-1/2}$ . If  $\varepsilon_n = \max |f(\alpha) - f(\beta)|$  over  $\{|\alpha - \beta| \leq yn^{-1/4}, 0 \leq \alpha, \beta \leq 2y\}$ , then  $\varepsilon_n$  tends to 0 by uniform continuity, and  $|f(y) - \mathbf{E}[f(S_n)]| \leq \mathbf{E}[|f(y) - f(S_n)|] \leq 2 \|f\|_{\infty} n^{-1/2} + \varepsilon_n$  tends to 0, so  $f(y) = \lim \mathbf{E}[f(S_n)]$ .

id br. Let  $Z_k$  be our capital after the  $k$ -th bet. We have  $\mathbf{E}[Z_{k+1}|Z_k] \leq 2pZ_k$ , with equality iff we bet all our money on the  $k+1$ -st round. Thus the strategy maximizing  $\mathbf{E}[Z_n]$  is to always bet everything, and has a value of  $\$(2p)^n$ . This strategy is extremely risky, since only in probability  $p^n$  we don't lose our initial investment. Replacing the utility  $Z$  by  $\log(Z)$  has two effects: it greatly punishes losing all your capital, and returns are diminishing. Now if we wager a proportion  $\alpha$  of our capital  $Z_k$  in the  $k+1$ -st round we'll have  $\mathbf{E}[\log(Z_{k+1})|Z_k, \alpha] = p \log(Z_k(1+\alpha)) + q \log(Z_k(1-\alpha))$  where  $q = 1-p$ . This is maximized for  $\alpha = 2p-1$ . Thus the optimal strategy always wagers a proportion of  $2p-1$  of our capital (which could be guessed, since for  $p = 1/2$  we'd not wager anything and for  $p = 1$  we'd be all in). This strategy has  $\mathbf{E}[\log(Z_n)] = n(p \log p + q \log q + \log 2)$ .

id br. engineer's solution: Suppose we always wager a fixed proportion  $\alpha$  of our capital. We'll have approximately  $pn$  wins and  $qn$  losses, meaning  $Z_n \approx (1+\alpha)^{pn}(1-\alpha)^{qn}$ . Thus  $\log(Z_n) \approx n(p \log(1+\alpha) + q \log(1-\alpha))$  is to be maximized, as before.

id bu. Let  $q_n$  be an enumeration of  $\mathbf{Q}$ . Let  $s_0(n) = n$  and  $s_k(n)$  a sub-sequence of  $s_{k-1}(n)$  for which  $F_{s_k(n)}(q_k)$  converges. Let  $s(n) = s_n(n)$ . Then  $H(q) = \lim_n F_{s(n)}(q)$  exists for all  $q \in \mathbf{Q}$ . We take  $F(x) = \inf_{q > x} H(q)$ . Note that  $H, F$  are weakly-increasing and take values in  $[0, 1]$ . Let us show that  $F$  is right-continuous: If  $x, \varepsilon$  are given we may find  $x < q$  so that  $H(q) \leq F(x) + \varepsilon$ . It follows that for any  $x \leq y < q$  we have  $F(x) \leq F(y) \leq F(x) + \varepsilon$ . For any  $x$  we have

$$\sup_{q < x} H(q) = \sup_{q < x} \lim_n F_{s(n)}(q) \leq \liminf_n F_{s(n)}(x) \leq \limsup_n F_{s(n)}(x) \leq \inf_{x < q'} \lim_n F_{s(n)}(q') = \inf_{x < q'} H(q') = F(x)$$

Now, if  $x$  is a continuity point of  $F$ , we fix rational sequences  $x - \frac{1}{n} < r_n^1 < r_n^2 < x$  so that  $F(x) \leftarrow F(r_n^1) \leq H(r_n^2) \leq \sup_{q < x} H(q)$  which yields  $\lim_n F_{s(n)}(x) = F(x)$ . The last part is immediate.

id cf. The numbers 5 and 504 are related by [ce] where  $L(\tau) = \frac{1}{504} \left( 1 - \frac{27g_3(\tau)}{8\pi^6} \right) = \sum_{\ell} \sigma_5(\ell) q^{\ell} = \sum_{k,n} k^5 q^{kn} = \sum_k \frac{k^5 q^k}{1 - q^k} = R(q)$ . We are almost there, except we want  $1+q^k$  in our denominator, and only odd  $k$ . Here we need

the following:  $\frac{q^k}{1+q^k} = \frac{q^k}{1-q^k} - 2\frac{q^{2k}}{1-q^{2k}}$ . Hence  $F(q) = \sum_{2 \nmid k} \frac{k^5 q^k}{1+q^k} = R(q) - 32R(q^2) - 2(R(q^2) - 32R(q^4)) = R(q) - 34R(q^2) + 64R(q^4)$ . However,  $G_6(i) = 0$  and  $G_6(i/2) = -64G_6(2i)$  and so the  $g_3$  parts cancel and we are left with  $F(\tau = i/2, q = e^{-\pi}) = 31/504$ .

## -2. Notes

The following is a minimal attempt at giving due credit.

id av. The given explanation is attributed to D.J. Newman.

id ax. The given explanation is attributed to M. Nair.

id be. The given explanation for  $\mathbf{F} = \mathbf{Q}$  is attributed to S. Vishwanathan.

id bg. The given explanation is attributed to S. Vishwanathan.

id bm. The Fact is attributed to F.J. MacWilliams.

id bv. The fact is known as the *central limit theorem*.

id bw. The fact is attributed to Lévy.

id bz. The explanation is due to E. Brown and N. Loehr.

id cn. The given explanation for *ii.* is due to Z.Y Kong and L.P Teo.