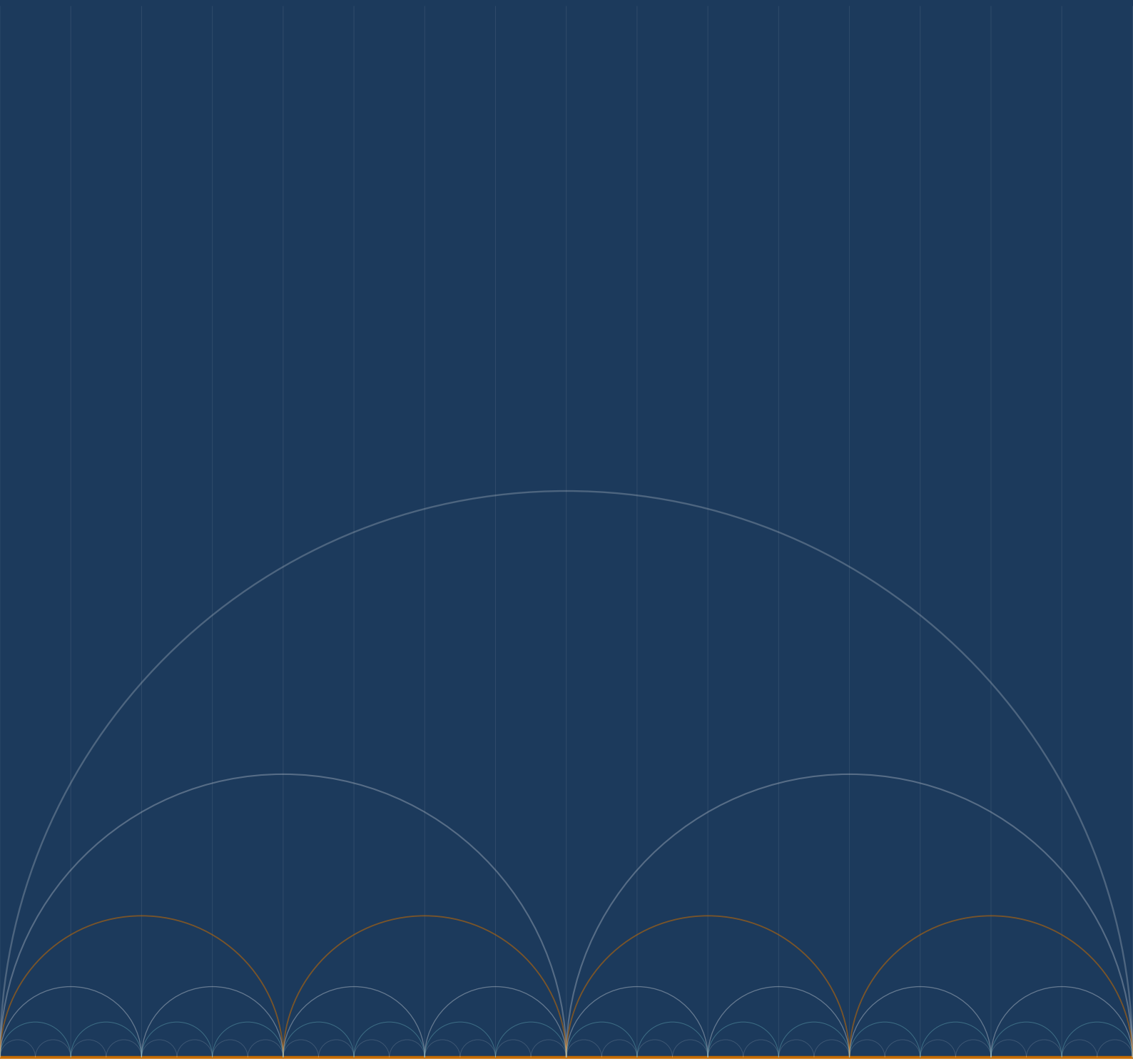
Lecture notes

A Spectral Tour of Hyperbolic Surfaces



Asma Hassannezhad

CUSO Summer School in Mathematics, 2026

Abstract

The first nonzero Laplace eigenvalue of a compact hyperbolic surface — its *spectral gap* — measures how well-connected the surface is. These notes are largely self-contained, with references for further details. They follow a single question across three lectures: how large can the gap be, and how large is it typically? The route passes through McKean's theorem ($\mathrm{spec}(\mathbb{H}) = [\frac{1}{4}, \infty)$), the Cheeger–Buser inequalities and the collar lemma, the sharp $2g - 2$ count of small eigenvalues, Huber's asymptotic ceiling $\frac{1}{4}$, and Mirzakhani's integration theory on moduli space, and ends at the 2025 theorem of Anantharaman and Monk: with respect to the Weil–Petersson probability measure, a random surface of genus $g \to \infty$ satisfies $\lambda_1 \geq \frac{1}{4} - \varepsilon$ with high probability. The parallel story for random regular graphs (Alon–Boppana bound, Alon's conjecture, Friedman's theorem) is developed alongside as both motivation and method. To help you along the way, the notes also include problem sets, appendices with the background you need, proof sketches, and guides to notation and reading.

An early draft of these notes was prepared with the assistance of Claude (Anthropic); the author is responsible for all content.

Contents

Introduction: one number, three lectures **4**

1 Lecture 1: The bottom of the spectrum **6**
1.1 Strings and drums: the spectrum as sound . . . 6
1.2 What the spectral gap measures . . . 8
1.3 First computation: flat tori . . . 8
1.4 Interlude: hearing a spectrum — heat trace, Weyl's law, isospectrality . . . 10
1.5 The hyperbolic plane . . . 11
1.6 The magic number $\frac{1}{4}$ . . . 14
1.7 Compact hyperbolic surfaces . . . 15
1.8 The dumbbell, and the three questions . . . 17

2 Lecture 2: Geometry controls the gap **19**
2.1 The Cheeger constant . . . 19
2.2 Cheeger's inequality . . . 20
2.3 Buser's inequality . . . 21
2.4 Collars: the geometry of short geodesics . . . 22
2.5 Small eigenvalues located: test functions, and Schoen–Wolpert–Yau . . . 24
2.6 How many small eigenvalues? . . . 25
2.7 The ceiling: no gap beats $\frac{1}{4}$ in large genus . . . 26
2.8 Who attains the ceiling? . . . 28

3 Lecture 3: Random hyperbolic surfaces and the optimal gap **30**
3.1 Moduli space as a probability space . . . 30
3.2 Mirzakhani's toolkit: integrating over moduli space . . . 32
3.3 Typical geometry and the first uniform spectral gap . . . 35
3.4 Why $\frac{1}{4}$ should be the truth: Benjamini–Schramm convergence . . . 36
3.5 Interlude: the same story for regular graphs . . . 37
3.6 The trace method on surfaces: from $\frac{3}{16}$ to the optimal gap . . . 39
3.7 Panorama and open problems . . . 41
Epilogue: the story in one page . . . 43

A Geometric background **44**
A.1 Curvature . . . 44
A.2 Geodesics and polar coordinates in $\mathbb{H}$ . . . 44
A.3 Classification of Möbius isometries . . . 44
A.4 Why "curvature -1" = "quotient of $\mathbb{H}$" . . . 45
A.5 Spectral theory used . . . 45
A.6 Hyperbolic trigonometry . . . 45
A.7 The heat kernel of $\mathbb{H}$ . . . 45

B What we did not prove: sketches with references **45**

C Annotated reading guide **48**

D Notation glossary **50**

E Exercise sheets **51**

E.1 Exercise sheet 1 51
E.2 Exercise sheet 2 52
E.3 Exercise sheet 3 55

Introduction: Spectral gap

A compact hyperbolic surface is a closed surface carrying a metric of constant curvature -1 — we will construct these objects in the first lecture. Its Laplace operator has a discrete spectrum

$$0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots \nearrow \infty,$$

and this mini-course is mainly concerned with the single number λ_1, the *spectral gap*. The gap measures how well-connected the surface is: surfaces with a large gap are hard to cut into pieces, have small diameter, and mix heat and random walks fast. One driving question is the following:

> *How large can the spectral gap of a compact hyperbolic surface be, and how large is it for a* ***typical*** *surface of large genus?*

We shall see that there is a hard ceiling: as the genus $g \to \infty$, no surface can beat $\lambda_1 \approx \frac{1}{4}$ (Lecture 2). And the ceiling is essentially attained by *almost every* surface: with respect to the natural (Weil–Petersson) probability measure on the space of genus-g surfaces,

$$\mathbb{P}_g^{\mathrm{WP}}\left(\lambda_1 \geq \tfrac{1}{4} - \varepsilon\right) \xrightarrow[g\to\infty]{} 1 \qquad \text{for every } \varepsilon > 0,$$

a theorem of Anantharaman and Monk [AM25, Mon26] in 2025 crowning a line of work begun by Mirzakhani [Mir13] in 2013. Lecture 3 tells this story, which runs parallel to a famous story about random regular graphs (Alon's 1986 conjecture, proved by Friedman in 2008).

Why we care about the spectral gap?

Here is an answer in four parts.

The applications are real. A uniform spectral gap is the mechanism behind **expander graphs**: bounded-degree networks that remain well connected and whose random walks mix rapidly, with applications to error-correcting codes, derandomisation, hashing, and robust network design [HLW06, CLG09]. The same principle governs reversible Markov chains: the first non-zero eigenvalue gives the L^2 decay rate $e^{-\lambda_1 t}$ and hence controls relaxation in Markov chain Monte Carlo, a basic tool of modern statistics and machine learning [LP17]. Hyperbolic surfaces provide a natural geometric setting for the same phenomenon: λ_1 measures both isoperimetric connectivity and the rate of heat diffusion, while the Selberg trace formula links this dynamical information exactly to the geometry of closed geodesics.

What transfers is the method. Few of you will work on random surfaces; all of you may reuse the methods you'll see in this mini-course: variational principles and test-function craft, the coarea trick, converting spectral problems into counting problems, and — above all — the probabilistic method's lesson, that **typical objects can beat explicit constructions**. Years of ingenious explicit constructions culminated in one probabilistic idea: "choose it at random".

You get to watch mathematics move. The endgame of this story is 2022–2025: two teams independently reaching $\frac{3}{16}$, the barrier diagnosed, the graph analogy imported, the

optimal theorem proved — with the neighbouring open problem (Selberg's conjecture) still sitting there.

It is a crossroads. Analysis, geometry, probability, number theory and mathematical physics meet at this one number. You will also see a recurring practice in mathematical reasoning: identifying the quantity that controls a problem and testing carefully how far an analogy can be pushed. The graphs↔surfaces correspondence provides a particularly clear example.

Learning goals

These are deliberately ambitious; there is no expectation of covering every item in full in three lectures. The main things to take away:

1. define the Laplace spectrum of a compact surface via quadratic forms and min–max, compute it for circles and flat tori, and explain what the spectral gap measures (Poincaré, mixing, connectivity);
2. compute concretely in the hyperbolic plane (lengths, areas, geodesics, isometries) and reproduce the proof that $\operatorname{spec}(\mathbb{H}) = [\frac{1}{4}, \infty)$;
3. prove Cheeger's inequality (coarea + Cauchy–Schwarz), and state Buser's inequality;
4. explain why short *separating* geodesic systems — and only those — produce small eigenvalues (collar lemma, Schoen–Wolpert–Yau);
5. explain why $\frac{1}{4}$ is the asymptotic ceiling (ball growth → diameter → Cheng → Huber);
6. describe the Weil–Petersson model and use Mirzakhani's integration formula in a computation;
7. Narrate the $0.002 \to \frac{3}{16} \to \frac{2}{9} \to \frac{1}{4} - \varepsilon$ history alongside its graph twin (Alon–Boppana, Alon's conjecture, Friedman). The proofs themselves are quoted, not given: for those, §3.6 points to the literature.

How to read these notes

Each lecture is one section; the material marked *Quick exercise* should be done in the margin as you read (answers are one-liners); the labelled Exercises 1.1–3.2* are collected in Appendix E: sheet 1 (after Lecture 1: 1.1–1.3), sheet 2 (after Lecture 2: 2.1, 2.2, 2.3), and a self-study sheet on the Lecture-3 material (3.1, 3.2*). Four appendices follow: Appendix A, the geometric background used but not developed here; B, a paragraph on each result we quote without proof, with a reference; C, an annotated reading guide by lecture; and D, a notation glossary.

1 Lecture 1: The bottom of the spectrum

1.1 Strings and drums: the spectrum as sound

Let us start in dimension one, where everything is computable, and extract the features that will survive to hyperbolic surfaces.

Example 1.1 (The string)**.** Consider $-u'' = \lambda u$ on $[0, L]$ with Dirichlet conditions $u(0) = u(L) = 0$. The general solution of the ODE for $\lambda > 0$ is $u = A\sin(\sqrt{\lambda}x) + B\cos(\sqrt{\lambda}x)$; the boundary conditions force $B = 0$ and $\sin(\sqrt{\lambda}L) = 0$, so

$$\lambda_k = \left(\frac{\pi k}{L}\right)^2, \qquad u_k(x) = \sin\left(\frac{\pi k x}{L}\right), \qquad k = 1, 2, 3, \dots$$

($\lambda \le 0$ gives only $u \equiv 0$; check.) These are the fundamental frequencies of a vibrating string: the *sound* of $[0, L]$ is its spectrum, and a long string has a deep fundamental tone $\lambda_1 = \pi^2/L^2$.

Example 1.2 (The circle)**.** On $\mathbb{R}/L\mathbb{Z}$ — a *closed* space, no boundary — the operator $-\frac{d^2}{dx^2}$ has the L-periodic eigenfunctions $e^{2\pi i k x/L}$, $k \in \mathbb{Z}$, and

$$\operatorname{spec}(\mathbb{R}/L\mathbb{Z}) = \left\{\left(\frac{2\pi k}{L}\right)^2 : k \ge 0\right\},$$

each nonzero eigenvalue with multiplicity 2. Completeness of the system is the theory of Fourier series.

Two features of Example 1.2 persist for *all* closed spaces and should be memorised now:

- The lowest eigenvalue of a closed space is 0, with the *constants* as eigenfunctions. It carries no information. The interesting quantity is the first *nonzero* eigenvalue — the **spectral gap**.
- Eigenvalues scale like $(\text{length})^{-2}$: doubling all lengths divides every eigenvalue by 4. Big spaces have deep bass notes. Keep this dimensional analysis running in the background all through the course; it explains, for instance, why bounds like Cheng's $\lambda_1 \le \frac{1}{4} + C/\operatorname{diam}^2$ (Lecture 2) have the shape they do.

The Laplacian on a surface

Let X be a compact surface equipped with a metric — for us always a *conformal* metric, introduced properly in §1.5; for now it is enough that X carries a gradient ∇, an area element $\mathrm{d}A$, and hence a Dirichlet energy $\int_X |\nabla f|^2 \,\mathrm{d}A$. Rather than writing a coordinate formula, define the Laplacian by the **quadratic-form recipe**:

$$\int_X f\,\Delta f \,\mathrm{d}A = \int_X |\nabla f|^2 \,\mathrm{d}A \qquad \text{for all smooth } f, \tag{1}$$

i.e. Δ is the (Friedrichs) self-adjoint operator associated with the closed form $q(f) = \int |\nabla f|^2$ on $H^1(X) \subset L^2(X)$. With our sign convention Δ is a *positive* operator (probabilists and geometers often use the opposite sign).

Theorem 1.3 (Spectrum of a compact surface)**.** *Let X be a compact connected surface. Then Δ has discrete spectrum*

$$0 = \lambda_0 < \lambda_1 \le \lambda_2 \le \cdots \to \infty,$$

with finite multiplicities and an orthonormal basis of $L^2(X)$ of smooth eigenfunctions $(\varphi_j)_{j\ge 0}$. The eigenvalue $\lambda_0 = 0$ has the constants as eigenspace, with multiplicity one.

Idea of proof. This is verbatim the argument for a bounded Euclidean domain. The embedding $H^1(X) \hookrightarrow L^2(X)$ is compact (Rellich–Kondrachov; compactness of X replaces boundedness of the domain), so the resolvent $(\Delta + 1)^{-1}$ is a compact self-adjoint operator, and the spectral theorem for compact operators gives the eigenvalue sequence. Smoothness of eigenfunctions is elliptic regularity. If $\Delta f = 0$ then $\int |\nabla f|^2 = 0$ by (1), so f is constant; connectedness makes the multiplicity one. □

The variational principle

The single most used tool of the course:

Theorem 1.4 (Min–max; variational characterisation of the gap)**.** *For every smooth $f \not\equiv 0$ define the* Rayleigh quotient $\mathcal{R}(f) = \int_X |\nabla f|^2 \,\mathrm{d}A \,\big/ \int_X f^2 \,\mathrm{d}A$. *Then*

$$\lambda_1(X) \;=\; \min\Big\{ \mathcal{R}(f) \;:\; f \neq 0,\; \int_X f \,\mathrm{d}A = 0 \Big\},$$

attained exactly at λ_1-eigenfunctions. More generally

$$\lambda_k = \min_{\dim V = k+1} \; \max_{f \in V \setminus 0} \mathcal{R}(f)$$

over subspaces $V \subset H^1(X)$.

Proof. Expand $f = \sum_j c_j \varphi_j$ in the eigenbasis. The constraint $\int f = 0$ says $c_0 = 0$, and by (1),

$$\mathcal{R}(f) = \frac{\sum_{j\ge 1} \lambda_j c_j^2}{\sum_{j \ge 1} c_j^2} \;\ge\; \lambda_1,$$

with equality iff $c_j = 0$ whenever $\lambda_j > \lambda_1$. For the general statement: given V of dimension $k+1$, it meets the codimension-k space $\{c_0 = \cdots = c_{k-1} = 0\}$ nontrivially, and on that intersection $\mathcal{R} \ge \lambda_k$; conversely $V = \operatorname{span}(\varphi_0, \dots, \varphi_k)$ achieves $\max \mathcal{R} = \lambda_k$. □

We record the principle in a box; Lectures 2 and 3 are variations on it:

Upper bounds are cheap: exhibit *one* mean-zero test function of small energy; then $\lambda_1 \le \mathcal{R}(f)$ — you only have to build one thing.

Lower bounds are expensive: prove that *every* mean-zero function pays energy — an inequality valid for all f.

1.2 What the spectral gap measures

Three readings of "λ_1 is large", each opening onto a different field.

Proposition 1.5 (Poincaré inequality)**.** *For all $f \in H^1(X)$, with $\bar{f} = \frac{1}{\operatorname{Area} X}\int f$:*

$$\int_X (f-\bar{f})^2 \,\mathrm{d}A \;\le\; \frac{1}{\lambda_1}\int_X |\nabla f|^2 \,\mathrm{d}A,$$

and $1/\lambda_1$ is the best constant.

Proof. $f - \bar{f}$ has mean zero; apply Theorem 1.4. Sharpness: $f = \varphi_1$. □

A large gap means the surface does not tolerate cheap non-constant behaviour: any function that varies must pay energy. This is the form in which λ_1 enters PDE and geometry.

Proposition 1.6 (Mixing)**.** *The heat semigroup $e^{-t\Delta}$ (solving $\partial_t u = -\Delta u$) satisfies*

$$\left\| e^{-t\Delta} f - \bar{f} \right\|_{L^2} \;\le\; e^{-\lambda_1 t}\, \|f - \bar{f}\|_{L^2}.$$

Proof. In the eigenbasis, $e^{-t\Delta} f = \sum_j c_j e^{-\lambda_j t}\varphi_j$; the $j = 0$ term is $\bar{f}$ and every other mode decays at least like $e^{-\lambda_1 t}$. □

So λ_1 is the *rate of equilibration* of heat, and $1/\lambda_1$ the relaxation time of Brownian motion on X. A small gap means that somewhere, heat gets stuck — and the reader should already suspect *where*: behind a bottleneck. That suspicion is the third reading:

(Connectivity — a theorem in Lecture 2, via Cheeger and Buser)
λ_1 is small *precisely when* a short curve cuts X into two large pieces.

Surfaces with a large gap are hard to cut, have small diameter, and mix fast: they are **continuous expanders**. The same trio (Poincaré / mixing / cuts) defines expander *graphs*, and this analogy is central to Lecture 3.

1.3 First computation: flat tori

Example 1.7 (Rectangular tori)**.** Let $T_{a,b} = \mathbb{R}^2/(a\mathbb{Z}\oplus b\mathbb{Z})$ with the Euclidean metric, $a \ge b > 0$. Functions on $T_{a,b}$ are (a,b)-biperiodic functions on $\mathbb{R}^2$; the exponentials

$$e_{m,n}(x,y) = e^{2\pi i(mx/a + ny/b)}, \qquad (m,n) \in \mathbb{Z}^2,$$

are eigenfunctions of $\Delta = -(\partial_x^2 + \partial_y^2)$ with eigenvalues

$$\lambda_{m,n} = 4\pi^2\left(\frac{m^2}{a^2} + \frac{n^2}{b^2}\right),$$

and form a Hilbert basis of L^2 (double Fourier series). Since a self-adjoint operator with compact resolvent has no spectrum beyond a complete eigensystem, this is the whole spectrum, and (recall $a \geq b$)

$$\lambda_1(T_{a,b}) = \frac{4\pi^2}{a^2}, \qquad \text{eigenfunctions } \cos\frac{2\pi x}{a},\ \sin\frac{2\pi x}{a}.$$

The gap does not see the area alone; it sees the *shape*. Normalise $ab = 1$ and let $a \to \infty$: the area is constant but

$$\lambda_1(T_{a,1/a}) = \frac{4\pi^2}{a^2} \longrightarrow 0.$$

The long thin torus is our first **bottleneck**. Look at the eigenfunction $\cos(2\pi x/a)$: it is essentially $+1$ on one half of the torus and -1 on the other, transitioning across the two short circles $\{x = a/4\}, \{x = 3a/4\}$ of length $b = 1/a$ each. All three readings of §1.2 degenerate coherently:

- *Poincaré fails cheaply*: the ± 1 function has tiny energy but unit variance;
- *mixing slows*: heat started on one half needs time $\asymp a^2$ to cross;
- *connectivity fails*: two circles of total length $2/a$ cut the torus into halves of area $\frac{1}{2}$ each — a short fence around a lot of land.

Quick exercise 1.8. Show $\operatorname{diam}(T_{a,b}) \geq a/2$, and check that the collapse $\lambda_1 \asymp 1/a^2$ is consistent with the general upper bound $\lambda_1 \leq C/\operatorname{diam}^2$ valid (with the right constant) on any closed surface. Exercise 1.1 takes this example further: general lattices, hearing the shape, and the cut-open cylinder.

Geometry controls the gap. The rest of the course makes this precise for a class of surfaces whose geometry is much richer than that of tori.

Why hyperbolic surfaces?

Why not stay with flat tori? Three reasons, in increasing order of importance.

Topology forces us beyond. A closed orientable surface of genus $g \geq 2$ — the generic case — carries no flat metric (Gauss–Bonnet, §1.7, makes this precise: total curvature must equal $2\pi\chi = 2\pi(2-2g) < 0$). It always carries metrics of constant curvature -1. Hyperbolic surfaces are not an exotic choice; they are where most surfaces live.

A rigid skeleton. Negative curvature endows a surface with a canonical infinite family of curves: each free homotopy class contains a *unique* closed geodesic, with a well-defined length. This "length spectrum" gives hyperbolic geometry a crystalline, rigid character

that flat geometry (where geodesics come in parallel families) lacks entirely — and it is the geometric side of the trace formula.

The richest spectral theory. Selberg's trace formula, arithmetic examples, connections to quantum chaos — and, our destination, a complete answer to the typical-gap question, something not available in any other geometry.

So we must build the hyperbolic plane.

1.4 Interlude: hearing a spectrum — heat trace, Weyl's law, isospectrality

Before leaving the flat world, three classical facts that calibrate what a spectrum *can* know about a space; they will quietly organise Lecture 3.

Theorem 1.9 (Weyl's law; heat trace)**.** *On a closed surface X, the eigenvalue counting function $N(\lambda) = \#\{j : \lambda_j \le \lambda\}$ satisfies*

$$N(\lambda) \;\sim\; \frac{\mathrm{Area}(X)}{4\pi}\,\lambda \qquad (\lambda \to \infty),$$

equivalently the heat trace obeys $\sum_{j\ge 0} e^{-\lambda_j t} \sim \frac{\mathrm{Area}(X)}{4\pi t}$ as $t \downarrow 0$, with a full asymptotic expansion whose next coefficient is $\frac{1}{12\pi}\int_X K \,\mathrm{d}A = \frac{\chi(X)}{6}$.

Quick exercise 1.10. Check Weyl's law by hand on the square torus $T_{L,L}$: $N(\lambda)$ counts lattice points of $\frac{2\pi}{L}\mathbb{Z}^2$ in a disk of radius $\sqrt{\lambda}$, which is $\sim$ area of the disk over the covolume $= \frac{L^2}{4\pi}\lambda$.

Consequences worth internalising:

- *The spectrum knows the area* — and, by the χ-term, the genus. So on hyperbolic surfaces (where area $= 2\pi(2g-2)$ anyway) the spectrum determines the topology.
- *But not everything*: there exist pairs of non-isometric surfaces (and famously, of planar domains: "one cannot hear the shape of a drum", Gordon–Webb–Wolpert 1992) with identical spectra. Isospectral hyperbolic surfaces exist in every large genus (Vignéras; Sunada; cf. Buser [Bus92, Ch. 11–12]). The spectrum is a powerful but incomplete fingerprint.
- *On hyperbolic surfaces the right pairing is with the length spectrum*: Huber's theorem states that two compact hyperbolic surfaces have the same Laplace spectrum if and only if they have the same length spectrum. The identity behind this equivalence is the Selberg trace formula — the engine of Lecture 3. Weyl's law itself is its simplest output.

1.5 The hyperbolic plane

1.5.1 Conformal metrics

A *conformal metric* on a domain $\Omega \subset \mathbb{R}^2$ is a prescription

$$ds^2 = \phi(x,y)\,(dx^2 + dy^2), \qquad \phi > 0 \text{ smooth}:$$

the "ruler" at the point (x,y) rescales Euclidean lengths by the factor $\sqrt{\phi(x,y)}$. All the geometry we need flows mechanically from this one idea:

$$\text{length:}\quad L(\gamma) = \int_a^b \sqrt{\phi(\gamma(t))}\,|\gamma'(t)|\,\mathrm{d}t, \qquad d(p,q) = \inf_\gamma L(\gamma), \tag{2}$$

$$\text{area:}\quad \operatorname{Area}(E) = \int_E \phi\,\mathrm{d}x\,\mathrm{d}y, \qquad \mathrm{d}A = \phi\,\mathrm{d}x\,\mathrm{d}y, \tag{3}$$

$$\text{gradient:}\quad |\nabla f|^2 = \frac{f_x^2 + f_y^2}{\phi}, \tag{4}$$

$$\text{Laplacian:}\quad \Delta f = -\frac{f_{xx} + f_{yy}}{\phi}, \tag{5}$$

$$\text{divergence:}\quad \operatorname{div} X = \frac{1}{\phi}\Big(\partial_x(\phi X^1) + \partial_y(\phi X^2)\Big), \qquad \Delta f = -\operatorname{div}\nabla f. \tag{6}$$

Geodesics are locally shortest curves. Formulas (4)–(5) deserve a derivation rather than decree. The metric pairing of two vectors v, w at a point is $\langle v, w\rangle_\phi = \phi\,(v \cdot w)$; the gradient is defined by $\langle \nabla f, v\rangle_\phi = df(v) = \nabla^{\text{euc}} f \cdot v$, which forces $\nabla f = \phi^{-1}\nabla^{\text{euc}} f$ and hence (4). The Laplacian is then defined by the quadratic-form recipe (1); let us check (5) satisfies it:

$$\int f\cdot\Big(-\frac{f_{xx}+f_{yy}}{\phi}\Big)\,\phi\,\mathrm{d}x\,\mathrm{d}y = -\int f\,(f_{xx}+f_{yy})\,\mathrm{d}x\,\mathrm{d}y = \int (f_x^2+f_y^2)\,\mathrm{d}x\,\mathrm{d}y = \int \frac{f_x^2+f_y^2}{\phi}\,\phi\,\mathrm{d}x\,\mathrm{d}y.$$

The middle step is Euclidean integration by parts; the outer terms are $\int f\Delta f\,\mathrm{d}A$ and $\int |\nabla f|^2\,\mathrm{d}A$.

Formula (6) is fixed by the same requirement, that the divergence theorem hold: writing ν for the outward unit normal (unit *in the metric*, so $\nu = \phi^{-1/2} n^{\text{euc}}$) and $\mathrm{d}s$ for arclength,

$$\int_\Omega \operatorname{div} X\,\mathrm{d}A = \oint_{\partial\Omega} \langle X, \nu\rangle_\phi\,\mathrm{d}s. \tag{7}$$

This is not a new theorem: both sides equal the Euclidean flux of ϕX, since $\operatorname{div} X\,\mathrm{d}A = \big(\partial_x(\phi X^1) + \partial_y(\phi X^2)\big)\,\mathrm{d}x\,\mathrm{d}y$ and $\langle X,\nu\rangle_\phi\,\mathrm{d}s = \phi\,(X\cdot n^{\text{euc}})\,\mathrm{d}s^{\text{euc}}$, so (7) *is* the divergence theorem of multivariable calculus applied to the vector field ϕX. Note the sign in (6): with our positive Δ, a function with $\Delta u = -1$ has $\operatorname{div}\nabla u = +1$. Notice now what happened in the chain above:

Proposition 1.11 (Conformal invariance of energy — the dimension-2 coincidence)**.** *For every conformal factor ϕ,*

$$\int |\nabla f|^2\,\mathrm{d}A = \iint (f_x^2 + f_y^2)\,\mathrm{d}x\,\mathrm{d}y. \tag{8}$$

The Dirichlet energy does not see the metric at all.

The two ϕ's — one from the gradient, one from the area element — cancel exactly in dimension 2 (in dimension n the same computation leaves $\phi^{(n-2)/2}$). Two consequences we use repeatedly: energies can always be computed in the flat picture, and only the L^2-norm $\int f^2\phi\,\mathrm{d}x\,\mathrm{d}y$ retains the geometry.

Quick exercise 1.12. Deduce from (8) that on a compact surface, $\lambda_1 \cdot \text{Area}$ is a *scale-invariant* quantity for conformal rescalings $\phi \to c\phi$ — recovering the $(\text{length})^{-2}$ law of §1.1.

1.5.2 What curvature does

One genuinely geometric invariant is needed. For a conformal metric written as $ds^2 = e^{2u}(dx^2 + dy^2)$, define the **Gaussian curvature**

$$K = -e^{-2u}\,(u_{xx} + u_{yy}). \tag{9}$$

Its operational meaning (we take it as the definition of what curvature *does*):

$$\text{Area}\,B(p,r) = \pi r^2\Big(1 - \frac{K(p)}{12}r^2 + o(r^2)\Big) \qquad (r \to 0)\,:$$

negative curvature = *more area* in small disks than Euclidean geometry allows. Three checks:

- $u \equiv 0$ (Euclidean plane): $K \equiv 0$.
- Round sphere of radius 1 via stereographic projection: $\phi = 4/(1+x^2+y^2)^2$; a short computation gives $K \equiv +1$.
- *The upper half-plane with $\phi = y^{-2}$, i.e. $u = -\log y$:* $u_{xx} + u_{yy} = 1/y^2$, so $K = -y^2 \cdot y^{-2} = -1$.

The last line is the object of the course.

Definition 1.13 (Hyperbolic plane)**.**

$$\mathbb{H} = \{z = x + iy :\ y > 0\}, \qquad ds^2 = \frac{dx^2 + dy^2}{y^2}.$$

It has constant curvature $K \equiv -1$: like $\mathbb{R}^2$ and the round sphere, it looks the same at every point and in every direction — the third and last member of the constant-curvature family.

1.5.3 First computations in $\mathbb{H}$

Vertical distances. For $0 < y_1 \le y_2$, the vertical segment from iy_1 to iy_2 has length $\int_{y_1}^{y_2} dy/y = \log(y_2/y_1)$, and no path does better: for any $\gamma = (x(t), y(t))$,

$$L(\gamma) = \int \frac{\sqrt{\dot{x}^2 + \dot{y}^2}}{y}\,\mathrm{d}t \;\ge\; \int \frac{|\dot{y}|}{y}\,\mathrm{d}t \;\ge\; \Big|\int \frac{\dot{y}}{y}\,\mathrm{d}t\Big| = \log\frac{y_2}{y_1}, \tag{10}$$

with equality iff $\dot{x} \equiv 0$ and $\dot{y}$ has constant sign. Hence $d(iy_1, iy_2) = \log(y_2/y_1)$, vertical lines are geodesics, and — letting $y_1 \to 0$ — *the boundary* $\{y = 0\}$ *is infinitely far away.* $\mathbb{H}$ is complete: an entire, inescapable world, despite being drawn inside a half-plane.

Isometries. For a real matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $ad - bc = 1$, consider the Möbius map $T(z) = \frac{az+b}{cz+d}$. Two identities (Exercise 1.2 asks you to verify them):

$$\operatorname{Im} T(z) = \frac{\operatorname{Im} z}{|cz+d|^2}, \qquad |T'(z)| = \frac{1}{|cz+d|^2}. \tag{11}$$

A holomorphic map scales infinitesimal Euclidean lengths by $|T'|$; the hyperbolic length of a tangent vector v at z is $|v|/\operatorname{Im} z$; so the image vector has hyperbolic length

$$\frac{|T'(z)|\,|v|}{\operatorname{Im} T(z)} = \frac{|v|\,|cz+d|^{-2}}{\operatorname{Im} z \cdot |cz+d|^{-2}} = \frac{|v|}{\operatorname{Im} z}\,:$$

unchanged. Thus every element of $\mathrm{PSL}_2(\mathbb{R})$ (matrices modulo $\pm I$) acts on $\mathbb{H}$ by isometries; these are in fact *all* orientation-preserving isometries. The action is transitive on points and even on unit tangent vectors: $z \mapsto z+b$ and $z \mapsto \lambda z$ ($\lambda > 0$) already move i anywhere, and the stabiliser of i rotates directions. $\mathbb{H}$ *is homogeneous and isotropic — it has no special points, just as the definition "constant curvature" promised.*

Geodesics, all of them. Isometries map geodesics to geodesics; Möbius maps send the vertical lines to vertical lines and half-circles orthogonal to $\mathbb{R}$; a count of endpoints on $\partial\mathbb{H} = \mathbb{R} \cup \{\infty\}$ shows every such arc arises. Conclusion:

The geodesics of $\mathbb{H}$ *are the vertical half-lines and the half-circles meeting* $\mathbb{R}$ *orthogonally.*

Through every point, in every direction, exactly one geodesic — and any two points are joined by a unique one.

Polar coordinates and exponential growth. Fix $p \in \mathbb{H}$ and let r be the hyperbolic distance to p, θ the direction at p. In these geodesic polar coordinates the metric takes the form

$$ds^2 = dr^2 + \sinh^2 r\, d\theta^2. \tag{12}$$

(Sketch: by homogeneity take $p = i$; by isotropy the metric must be $dr^2 + G(r)d\theta^2$ (homogeneity says there is no preferred point; isotropy says there is no preferred direction at a point.); computing along the vertical geodesic $r \mapsto ie^r$ with (11) identifies $G = \sinh^2 r$; the factor $\sinh r$ is exactly the rate at which nearby geodesics through p spread — in curvature -1, Jacobi fields grow like sinh.) Consequently

$$\operatorname{Area} B(p,r) = \int_0^{2\pi}\!\!\int_0^r \sinh\rho\,\mathrm{d}\rho\,\mathrm{d}\theta = 2\pi(\cosh r - 1) = 4\pi\sinh^2(r/2) \;\sim\; \pi e^r. \tag{13}$$

Inequality (13) underlies much of what follows. A hyperbolic disk of radius 10 has area $\approx 6.9 \times 10^4$, against $100\pi \approx 314$ in the plane: there is *exponentially much room at infinity* in $\mathbb{H}$. This is behind the spectral bound $\frac{1}{4}$ (next subsection), the logarithmic diameter bound (Lecture 2), the exponential count of closed geodesics, and the barrier at $\frac{3}{16}$ in the trace method (Lecture 3).

1.6 The magic number $\frac{1}{4}$

On $\mathbb{H}$, by (5), $\Delta = -y^2(\partial_x^2 + \partial_y^2)$, self-adjoint on $L^2(\mathbb{H}, \frac{dx\,dy}{y^2})$. Since $\mathbb{H}$ is not compact, Theorem 1.3 does not apply: the spectrum is continuous. The surprise is where it starts.

Theorem 1.14 (McKean [McK70])**.** $\operatorname{spec}\big(\Delta, L^2(\mathbb{H})\big) = \big[\frac{1}{4}, \infty\big)$.

Contrast: $\operatorname{spec}(\Delta, L^2(\mathbb{R}^2)) = [0, \infty)$ (Fourier transform). The bottom of the spectrum of the hyperbolic plane is *strictly positive* — a genuinely hyperbolic phenomenon, and the origin of every $\frac{1}{4}$ in this course.

Proof that $\operatorname{spec} \subset [\frac{1}{4}, \infty)$. By conformal invariance (8), for $f \in C_c^\infty(\mathbb{H})$,

$$\int_{\mathbb{H}} |\nabla f|^2 \,\mathrm{d}A = \iint (f_x^2 + f_y^2)\,\mathrm{d}x\,\mathrm{d}y \;\geq\; \iint f_y^2 \,\mathrm{d}x\,\mathrm{d}y.$$

Now a purely one-dimensional fact, applied on each vertical line:

Hardy's inequality. For $u \in C_c^\infty(0, \infty)$:

$$\int_0^\infty u'(y)^2 \,\mathrm{d}y \;\geq\; \frac{1}{4}\int_0^\infty \frac{u(y)^2}{y^2}\,\mathrm{d}y. \tag{14}$$

Proof: integrate by parts and apply Cauchy–Schwarz:

$$\int_0^\infty \frac{u^2}{y^2}\,\mathrm{d}y = \int_0^\infty u^2\, d\Big(-\frac{1}{y}\Big) = \int_0^\infty \frac{2uu'}{y}\,\mathrm{d}y \leq 2\Big(\int u'^2\Big)^{1/2}\Big(\int \frac{u^2}{y^2}\Big)^{1/2};$$

divide and square. (Compact support kills the boundary terms — an essential hypothesis.)

Applying (14) to $u = f(x, \cdot)$ for each fixed x and integrating in x:

$$\int_{\mathbb{H}} |\nabla f|^2 \,\mathrm{d}A \;\geq\; \frac{1}{4}\iint \frac{f^2}{y^2}\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{4}\,\|f\|^2_{L^2(\mathbb{H})},$$

i.e. $\Delta - \frac{1}{4} \geq 0$ on a form core, hence on $L^2(\mathbb{H})$. □

Why $\frac{1}{4}$ *is sharp: the functions* y^s. Compute on power functions: $\Delta y^s = -y^2\,\partial_y^2 y^s = s(1-s)\,y^s$. Setting $s = \frac{1}{2} + ir$, $r \in \mathbb{R}$:

$$\Delta\, y^{1/2+ir} = \Big(\frac{1}{4} + r^2\Big) y^{1/2+ir} :$$

generalised eigenfunctions sweeping exactly $[\frac{1}{4}, \infty)$. They are not in L^2, but localising them produces quasimodes: with $u_n(y) = y^{1/2}\varphi(\log y/n)$ and a slow cutoff in x, one computes $\mathcal{R} \to \frac{1}{4}$ (Exercise 1.3(c*) carries this out in full). Hence every $\lambda \geq \frac{1}{4}$ lies in the spectrum, and no smaller λ can, by the first half. □

Remark 1.15 (Another route to $\frac{1}{4}$: the ground-state transform). For any smooth $h > 0$, writing $f = \phi h$, expanding $|\nabla f|^2$ and integrating by parts gives $\int |\nabla f|^2 = \int \frac{\Delta h}{h} f^2 + \int h^2 |\nabla \phi|^2 \;\geq\; \int \frac{\Delta h}{h}\, f^2$ (with our positive Δ). In geodesic polar coordinates (12), the choice $h = e^{-r/2}$ gives $\frac{\Delta h}{h} = \frac{1}{2}\coth r - \frac{1}{4} > \frac{1}{4}$, reproving McKean's lower

bound; the radial test functions $e^{-r/2}\chi_R$ prove sharpness with no two-scale cutoff. More generally $h = e^{-\alpha r}$ yields $\alpha \coth r - \alpha^2 \geq \alpha - \alpha^2$, maximised at $\alpha = \frac{1}{2}$: the parametrisation $s(1-s)$ again, seen from the radial side. We omit the full details, which include a small regularity point at $r = 0$; the same transform reappears in Exercise 2.2(d).

Remark 1.16 (The parametrisation to memorise). Write, once and for all,

$$\lambda = s(1-s), \qquad s = \tfrac{1}{2} + ir.$$

Then: $\lambda \geq \frac{1}{4} \iff r \in \mathbb{R}$ (*tempered* spectrum, the spectrum of $\mathbb{H}$); $\lambda \in (0, \frac{1}{4}) \iff r$ purely imaginary, $r \in i(0, \frac{1}{2})$, i.e. $s \in (0, \frac{1}{2})$ — or equivalently $s \in (\frac{1}{2}, 1)$, since s and $1-s$ give the same λ (*exceptional* eigenvalues — impossible on $\mathbb{H}$, possible on compact quotients, and the source of the difficulty in Lecture 3: in the trace formula an imaginary r_1 makes test-function contributions *exponentially large*). The number $\frac{1}{4}$ is the symmetry point of $s \mapsto s(1-s)$.

1.7 Compact hyperbolic surfaces

1.7.1 Quotients

Definition 1.17. A *compact hyperbolic surface* is a quotient $X = \Gamma\backslash\mathbb{H}$, where $\Gamma \subset \mathrm{PSL}_2(\mathbb{R})$ is a subgroup that is *discrete*, acts *freely* (no fixed points in $\mathbb{H}$), and *cocompactly* (compact quotient). Equivalently (Appendix A): a closed surface carrying a metric of constant curvature -1.

Functions on X are exactly the Γ-periodic functions on $\mathbb{H}$; the metric, area, gradient and Laplacian all descend, because they are $\mathrm{PSL}_2(\mathbb{R})$-invariant. A *fundamental domain* is a region of $\mathbb{H}$ containing one point of each orbit; X is obtained by gluing its boundary via elements of Γ, and $\mathrm{Area}(X)$ is the area of the domain.

Example 1.18 (A cylinder — the simplest quotient, and a preview of collars)**.** Take $\Gamma = \langle z \mapsto e^{\ell} z \rangle$, the cyclic group generated by one dilation ($\ell > 0$). A fundamental domain is the annular region $\{1 \leq |z| < e^{\ell}\} \cap \mathbb{H}$. The quotient $C_\ell = \Gamma\backslash\mathbb{H}$ is an infinite cylinder containing exactly one simple closed geodesic — the image of the imaginary axis, of length ℓ (the segment from i to ie^{ℓ}, by (10)). In Fermi coordinates (ρ = signed distance to that geodesic, $t \in \mathbb{R}/\mathbb{Z}$ proportional to arclength along it) the metric becomes

$$ds^2 = d\rho^2 + \ell^2 \cosh^2\rho \; dt^2. \tag{15}$$

(The verification is a pleasant computation with (11); the useful identity is $\sinh(\mathrm{dist}(x+iy,\ i\mathbb{R}_{>0})) = |x|/y$.) The cylinder flares exponentially away from its waist — (13) again. Formula (15) will be, verbatim, the model of *collars* around short geodesics in Lecture 2; the entire quantitative theory of pinched surfaces happens inside it. *Not compact*, of course: compact quotients need bigger groups.

Example 1.19 (The Bolza surface: gluing an octagon)**.** In $\mathbb{H}$, consider regular octagons centred at i (in the disk model of Appendix A they are honest regular octagons centred at 0). A tiny regular octagon is nearly Euclidean: interior angles near $3\pi/4$. As the octagon grows, its angles decrease monotonically to 0 (its vertices approach the boundary at infinity). By continuity there is exactly one size with all angles $\pi/4$. Glue opposite sides by (uniquely determined) isometries. All eight vertices are identified to a *single* point, where the total angle is $8 \cdot \frac{\pi}{4} = 2\pi$: the glued object is a smooth closed hyperbolic surface. Its topology: the standard octagon-gluing of the genus-2 surface. This is the *Bolza surface*, the most symmetric surface of genus 2; remember it — it returns at the very end of the course as the conjectural champion of the spectral gap in genus 2.
Gauss–Bonnet for polygons: a geodesic n-gon with angles α_i has area $(n-2)\pi - \sum \alpha_i$: the octagon has area $6\pi - 8 \cdot \frac{\pi}{4} = 4\pi$ — matching $2\pi(2g-2)$ for $g = 2$ below.

1.7.2 The basic facts

We record the structural facts about compact hyperbolic surfaces that the rest of the course relies on.

Theorem 1.20 (Structure of compact hyperbolic surfaces)**.** *Let $X = \Gamma\backslash\mathbb{H}$ be a compact hyperbolic surface. Then:*

*(i) (**Topology and area**) X is a closed orientable surface of some genus $g \geq 2$, and*

$$\mathrm{Area}(X) = -2\pi\chi(X) = 2\pi(2g-2).$$

(Gauss–Bonnet: total curvature $= 2\pi\chi$; with $K \equiv -1$, area $= -2\pi\chi$.) Topology fixes the area: large genus $\iff$ large surface. *No rescaling is available — hyperbolic geometry has an absolute length scale.*

*(ii) (**Spectrum**)* $\mathrm{spec}(\Delta_X)$ *is discrete:* $0 = \lambda_0 < \lambda_1 \leq \cdots \to \infty$ *(Theorem 1.3 applies).*

*(iii) (**Closed geodesics**) Every non-contractible free homotopy class of loops on X contains a* unique *closed geodesic, the length minimiser of its class. These correspond to conjugacy classes of Γ, as follows. Every $\gamma \in \Gamma\setminus\{1\}$ is* hyperbolic *(Appendix A): it fixes two points of $\partial\mathbb{H}$, and the geodesic joining them — its* axis *— is preserved by γ. Conjugating those fixed points to 0 and ∞ turns γ into $z \mapsto e^{\ell}z$, whose axis is the imaginary half-line; by (10), $d(iy, ie^{\ell}y) = \log e^{\ell} = \ell$, so γ translates its axis by ℓ. Since $z \mapsto e^{\ell}z$ is $\left(\begin{smallmatrix} e^{\ell/2} & 0 \\ 0 & e^{-\ell/2}\end{smallmatrix}\right)$, of trace $2\cosh(\ell/2)$, and the trace is a conjugacy invariant,*

$$\ell(\gamma) = 2\,\mathrm{arccosh}\left(|\mathrm{tr}\,\gamma|/2\right).$$

The axis is γ-invariant, so its image in $X = \Gamma\backslash\mathbb{H}$ is a closed curve: the closed geodesic of the class of γ, of length $\ell(\gamma)$.

(iv) ***(Exponential proliferation; Huber's prime geodesic theorem)*** *The number of primitive closed geodesics of length $\leq L$ satisfies*

$$\#\{\gamma : \ell_\gamma \leq L\} \sim \frac{e^L}{L} \qquad (L \to \infty),$$

in perfect analogy with the prime number theorem, $\#\{p \leq x\} \sim x/\log x$, with e^L playing the role of x and geodesics that of primes.

We take (i) (Gauss–Bonnet) and the classification of Möbius elements as background (the latter is in Appendix A); (iii) is then a short argument (minimise length in a homotopy class; negative curvature prevents both non-existence and non-uniqueness); (iv) is deep — it historically *follows from* the Selberg trace formula of Lecture 3, and its exponential growth rate is yet another manifestation of (13).

The set $\{\ell_\gamma\}$, with multiplicities, is the **length spectrum** of X. Two spectra now live on every hyperbolic surface — Laplace eigenvalues and geodesic lengths — and the deepest theme of the subject is that *they determine each other* (Selberg). The exponential growth in (iv) will be both our tool and our obstacle in Lecture 3.

Quick exercise 1.21. On the cylinder of Example 1.18, verify (iii) by hand: the conjugacy class of the generator corresponds to the waist geodesic, and the class of its k-th power to the same curve traversed k times, length $k\ell$. (Answer: $\Gamma = \langle g \rangle$ is abelian, so every conjugacy class is a single g^k. All the g^k with $k \neq 0$ share the same axis, the imaginary half-line, and $g^k : z \mapsto e^{k\ell} z$ translates it by $|k|\ell$ — as the trace formula confirms, $2\operatorname{arccosh}(|\operatorname{tr} g^k|/2) = 2\operatorname{arccosh}(\cosh(k\ell/2)) = |k|\ell$. Its image in C_ℓ is the waist, traversed $|k|$ times.)

1.8 The dumbbell, and the three questions

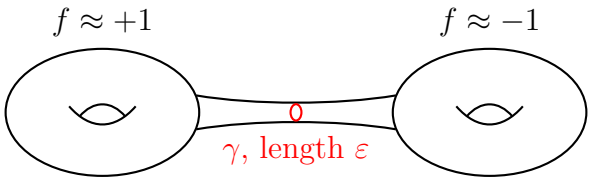


Figure 1: The genus-2 dumbbell. A separating geodesic γ of length ε pinches the surface into two halves; the test function $f = \pm 1$ on the two sides pays energy only in the neck, forcing $\lambda_1 \to 0$ as $\varepsilon \to 0$.

Hyperbolic surfaces with a tiny spectral gap are easy to manufacture (Figure 1). Take genus 2, and arrange for a closed geodesic γ that *separates* X into two halves to have small length ε. (Why is this possible? Because cuff lengths are freely prescribable parameters of hyperbolic surfaces — this is the Fenchel–Nielsen coordinate system, previewed here and constructed in Lecture 3. The dumbbell is a legitimate citizen of the space of genus-2 surfaces.) Now test:

$f = +1$ on one half, $f = -1$ on the other, interpolating across the thin neck.

The energy of f is paid only in the neck and is small; $\int f^2 \approx \mathrm{Area}(X) = 4\pi$; if the halves have equal area, $\int f = 0$ by symmetry. By Theorem 1.4, $\lambda_1(X) \le \mathcal{R}(f) \to 0$ as $\varepsilon \to 0$. Lecture 2 makes this quantitative — the truth is $\lambda_1 \asymp \varepsilon$, and the *right* test function is constructed in §2.5, and Exercise 2.2 explores what it beats and how it generalises — and proves a converse: this is essentially the *only* way to make the gap small.

The dumbbell is the hyperbolic analogue of the long thin torus: the Poincaré inequality is weak, mixing is slow, and connectivity fails — it fails each of the properties discussed in §1.2.

Meanwhile, in the other direction: the Bolza surface of Example 1.19 has $\lambda_1 \approx 3.84$ — far *above* the $\frac{1}{4}$ of McKean's theorem. So in fixed genus the gap can be substantial, and small genus feels no ceiling. The tension between these observations drives the whole course:

Question. **Q1.** Exactly which geometric features force λ_1 to be small? *(Lecture 2: short separating multicurves — and nothing else.)*
Q2. How large can λ_1 be in genus g, as $g \to \infty$? *(Lecture 2: at most $\frac{1}{4} + o(1)$ — the hyperbolic plane's bottom is an asymptotic ceiling.)*
Q3. Are surfaces with near-maximal gap rare or generic? *(Lecture 3: generic — typical surfaces achieve $\frac{1}{4} - \varepsilon$.)*

What to remember from Lecture 1

1. The gap λ_1: variational (Theorem 1.4); measures Poincaré / mixing / connectivity. Upper bounds = one test function; lower bounds = an inequality for all f.
2. Conformal metrics: formulas (2)–(5); energy is conformally invariant (8).
3. $\mathbb{H}$: geodesics, Möbius isometries, and above all Area $B(p, r) \sim \pi e^r$.
4. $\mathrm{spec}(\mathbb{H}) = [\frac{1}{4}, \infty)$, by Hardy's inequality; parametrise $\lambda = s(1-s)$, $s = \frac{1}{2} + ir$.
5. Compact hyperbolic surfaces $= \Gamma\backslash\mathbb{H}$; genus $g \ge 2$; Area $= 2\pi(2g-2)$; unique geodesic per homotopy class; e^L/L of them.
6. Dumbbells: short separating geodesics give small gaps; Bolza: fixed genus allows big gaps.

Exercise session 1 follows this lecture (Appendix E.1): **A1** flat tori — general lattices and hearing the shape; **A2** upper half-plane workout; **A3** Hardy's inequality with general exponents.

2 Lecture 2: Geometry controls the gap

Lecture 1 ended with three questions. This lecture answers Q1 (which geometry makes λ_1 small: short *separating* systems of geodesics, and nothing else) and Q2 (how large can λ_1 be: at most $\frac{1}{4} + o(1)$ as $g \to \infty$), and ends by setting up Q3.

2.1 The Cheeger constant

We first need a *number* that quantifies "worst bottleneck".

Definition 2.1 (Cheeger constant)**.** For a closed surface X,

$$h(X) = \inf_{\alpha} \frac{\ell(\alpha)}{\min\{\mathrm{Area}(X_1), \mathrm{Area}(X_2)\}},$$

the infimum over finite unions α of disjoint smooth closed curves that cut X into two pieces: $X \setminus \alpha = X_1 \sqcup X_2$.

Read it as: *cheapest fence, per unit of land enclosed on the smaller side.*

- The min in the denominator is essential (why?)
- α may be disconnected — think of the optimal cut of the long thin torus. One circle does not separate a torus!
- By an observation of Yau [Yau75] (see also [Bus92, Lemma 8.3.6]) one may restrict to cuts with both sides connected; and on a hyperbolic surface one may tighten each curve of α to the geodesic in its free homotopy class, shortening $\ell(\alpha)$ — so the infimum "wants" to be realised on systems of simple closed geodesics. (See Mirzakhani's Lemma in §3.3)

Example 2.2. The genus-2 dumbbell with waist length ε and halves of area 2π each has $h(X_\varepsilon) \le \varepsilon/2\pi$.

Remark 2.3 (Even connectivity has a ceiling: linear isoperimetry)*.* How large can h be for hyperbolic surfaces? In $\mathbb{H}$ itself, the isoperimetric inequality is *linear*: every region Ω satisfies $\ell(\partial\Omega) \ge \mathrm{Area}(\Omega)$ (Yau [Yau75], who proved $(n-1)\,\mathrm{vol}(\Omega) \le \mathrm{vol}(\partial\Omega)$ in $\mathbb{H}^n$; check on disks: $\ell(\partial B_r) = 2\pi \sinh r > 2\pi(\cosh r - 1) = \mathrm{Area}\, B_r$, with ratio $\to 1$). Consequently the Cheeger constant of any hyperbolic surface satisfies $h(X) \le 1 + o(1)$ as $g \to \infty$ (Cheng's comparison, in the form $h^2 \le 1 + O(1/\operatorname{diam}^2)$; [Che75]). So the natural scale for h on large hyperbolic surfaces is the interval $(0, 1)$, and Mirzakhani's 0.099 in Lecture 3 should be read against the ceiling 1.

Quick exercise 2.4. In the half-plane, set $u = -\log y$ and compute directly from $ds^2 = \frac{dx^2+dy^2}{y^2}$ that $|\nabla u| \equiv 1$ and $\operatorname{div} \nabla u \equiv 1$. Why does such a u give the linear isoperimetric inequality? (Answer: $\mathrm{Area}(\Omega) = \int_\Omega \operatorname{div} \nabla u \le \oint_{\partial\Omega} |\nabla u| = \ell(\partial\Omega)$ by Cauchy–Schwarz — level sets of u here are horocycles. The full argument is sketched

in Appendix B, *Linear isoperimetry in* $\mathbb{H}$.)

2.2 Cheeger's inequality

Theorem 2.5 (Cheeger 1970 [Che70])**.** *For every closed surface* $\lambda_1(X) \geq \dfrac{h(X)^2}{4}$.

This is remarkable: from *one* isoperimetric number, a lower bound on the gap — the "expensive" direction of the variational principle, obtained by one formula from geometric measure theory.

Coarea formula. For $\varphi \geq 0$ smooth on a surface, $\displaystyle\int_X |\nabla\varphi| \,\mathrm{d}A = \int_0^\infty \ell\big(\{\varphi = t\}\big)\,\mathrm{d}t$. *Meaning: the total variation of* φ *is the accumulated length of its level curves. One-dimensional case: level sets are* points*, so length is replaced by counting, and the formula reads* $\int |\varphi'|\,\mathrm{d}x = \int \#\{\varphi = t\}\,\mathrm{d}t$ *— total rise and fall, counted with multiplicity. (For monotone increasing* φ *on* $[a,b]$ *both sides equal* $\varphi(b) - \varphi(a)$*.)*

Note why coarea is *the* tool here: the Cheeger constant speaks of curve lengths and areas of the two sides; the coarea formula converts integrals of gradients into precisely these quantities, applied to all level sets simultaneously.

Proof of Theorem 2.5. Step 1: restrict to the positivity set. Let u be a λ_1-eigenfunction. Since $\int u = 0$, both $\{u > 0\}$ and $\{u < 0\}$ are nonempty; replacing u by $-u$ if necessary, assume $\Omega := \{u > 0\}$ has $\mathrm{Area}(\Omega) \leq \frac{1}{2}\mathrm{Area}(X)$. On Ω we have $u > 0$, $u = 0$ on $\partial\Omega$, and $\Delta u = \lambda_1 u$; multiplying by u and integrating by parts *on* Ω — the boundary term vanishes because $u|_{\partial\Omega} = 0$ — gives

$$\lambda_1 = \frac{\int_\Omega |\nabla u|^2}{\int_\Omega u^2}. \tag{16}$$

This last step needs justifying: the nodal set $\{u = 0\}$ need *not* be a smooth (or Lipschitz) curve — it is singular where u and ∇u vanish together, and there several arcs cross, so neither the integration by parts nor $\ell(\partial\Omega)$ is meaningful a priori. Two remedies. Either run the argument on $\{u > \delta\}$ for a regular value δ (Sard, as in Step 2) and let $\delta \downarrow 0$: there the boundary is smooth and the boundary term is signed, so (16) holds with $\leq$ — the direction Step 3 needs. Or use that u is real-analytic, the metric being so: its nodal set is then finitely many analytic arcs meeting at finitely many points, of zero length and harmless to both computations.

Step 2: coarea on $\varphi = (u^+)^2$*.* Set $u^+ = \max(u, 0)$ and $\varphi = (u^+)^2$, so that $\varphi = u^2$ on Ω and $\varphi \equiv 0$ off it. By Sard's theorem, a.e. $t > 0$ is a regular value of φ, so $\{\varphi = t\} = \{u = \sqrt{t}\}$ is a finite union of disjoint smooth circles, cutting X into $\{\varphi > t\}$ and its complement. Since φ vanishes outside Ω we have $\mathrm{Area}(\{\varphi > t\}) \leq \mathrm{Area}(\Omega) \leq \frac{1}{2}\mathrm{Area}(X)$ and the definition of h applies with $\min = \mathrm{Area}(\{\varphi > t\})$:

$$\ell(\{\varphi = t\}) \;\geq\; h \cdot \mathrm{Area}(\{\varphi > t\}) \qquad \text{for a.e. } t.$$

Integrate over t, using coarea on the left and the layer-cake formula on the right:

$$\int_\Omega |\nabla\varphi|\,\mathrm{d}A \;=\; \int_0^\infty \ell(\{\varphi = t\})\,\mathrm{d}t \;\geq\; h\int_0^\infty \mathrm{Area}(\{\varphi > t\})\,\mathrm{d}t \;=\; h\int_\Omega \varphi \;=\; h\int_\Omega u^2.$$

Step 3: Cauchy–Schwarz. $|\nabla\varphi| = 2u^+|\nabla u|$, so

$$\int_\Omega |\nabla\varphi| \;\leq\; 2\Big(\int_\Omega |\nabla u|^2\Big)^{1/2}\Big(\int_\Omega u^2\Big)^{1/2} = 2\sqrt{\lambda_1}\int_\Omega u^2.$$

Comparing with Step 2: $h \leq 2\sqrt{\lambda_1}$. □

Anatomy of the proof: coarea is where the geometry enters, through the definition of h; Cauchy–Schwarz is where sharpness is lost. It returns $h \leq 2\sqrt{\lambda_1}$— the square root is a real loss (the dumbbell has $h \asymp \varepsilon$ but $\lambda_1 \asymp \varepsilon$, not ε^2 — see §2.5), and it will have a price tag in Lecture 3: Mirzakhani's 0.002 instead of 0.25. Exercise 2.1 has you transplant every step to graphs.

2.3 Buser's inequality

Cheeger's inequality cannot be reversed in general — but it can under a curvature bound:

Theorem 2.6 (Buser 1982 [Bus82, Bus92])**.** *For a closed surface with curvature ≥ -1 (in particular any hyperbolic surface),*

$$\lambda_1(X) \;\leq\; 2\,h(X) + 10\,h(X)^2.$$

Idea of proof. Take a cut α nearly realising h. The natural test function is ± 1 on the two sides, interpolated linearly across a tubular neighbourhood of α of width δ: its Rayleigh quotient is roughly $\frac{\mathrm{Area(tube)}}{\delta^2 \min \mathrm{Area}(X_i)}$. To convert this into a bound in terms of $h = \frac{\ell(\alpha)}{\min \mathrm{Area}(X_i)}$ one must control the *area of the tube* by $\delta \cdot \ell(\alpha)$ up to constants — and that is a curvature statement: with $K \geq -1$, tubes fatten at most like sinh, i.e. controllably. Optimising δ gives the stated bound. (Full proof: [Bus82], or [Bus92, Ch. 8]; a heat-kernel proof by Ledoux [Led94] is also classical.) □

Remark 2.7 (Curvature is necessary). Buser [Bus79] constructs surfaces (of very negative, non-constant curvature) with $h \to 0$ but λ_1 bounded below: the cut is short, but its neighbourhood flares so violently that crossing it is expensive for *every* function, not just the linear profile. So the reverse Cheeger inequality is a theorem about curvature, not just topology. For our hyperbolic surfaces ($K \equiv -1$) both inequalities hold, and:

$$\frac{h^2}{4} \;\leq\; \lambda_1 \;\leq\; 2h + 10h^2 \qquad\Longrightarrow\qquad \lambda_1 \to 0 \iff h \to 0.$$

"Gap = connectivity" is now a theorem.

Remark 2.8 (The graph twins). Both inequalities have exact discrete analogues — for a d-regular graph with adjacency matrix A and Cheeger constant $h_G = \min_{|S|\le n/2} \frac{|\partial S|}{|S|}$:

$$\frac{h_G^2}{2d} \;\le\; d - \mu_2 \;\le\; 2h_G$$

(Alon–Milman [AM85], Dodziuk [Dod84]). This discrete Cheeger theory is the foundation of expander graphs, and the reader should hold on to the parallel: it becomes systematic in Lecture 3.

Interlude: expanders, in one page

The Cheeger–Buser circle of ideas is best appreciated against its discrete original. A sequence of d-regular graphs G_n with $|V(G_n)| \to \infty$ is an *expander family* if $\inf_n h_{G_n} > 0$ — equivalently (discrete Cheeger–Buser) $\inf_n(d - \mu_2(G_n)) > 0$: sparse graphs ($dn/2$ edges only), yet uniformly impossible to cut cheaply. Three facts frame the analogy we will need:

- *Existence is a probabilistic fact*: random d-regular graphs are expanders whp (Pinsker [Pin73]) — long before explicit constructions caught up. (Foreshadowing: Mirzakhani 2013 is exactly the surface analogue.)
- *Explicit optimality came from arithmetic*: the best possible spectral gap — the *Ramanujan* bound — was reached by Lubotzky–Phillips–Sarnak [LPS88], by number theory. (Foreshadowing: Selberg's $\frac{3}{16}$.)
- *The optimal gap of random graphs was hard*: Alon's conjecture (1986) took until Friedman (2008). (Foreshadowing: Anantharaman–Monk 2025.)

By Cheeger–Buser, a sequence of hyperbolic surfaces with $g \to \infty$ and $\inf \lambda_1 > 0$ is precisely a *continuous expander family*. Whether such sequences exist at all was, historically, first settled through the arithmetic route (Selberg).

2.4 Collars: the geometry of short geodesics

Which cuts can be short? On a hyperbolic surface a short essential curve can be tightened to a short closed *geodesic*, so we must understand neighbourhoods of short geodesics. The answer is a gem of hyperbolic trigonometry, and the single most used geometric fact in the subject.

Theorem 2.9 (Collar lemma; see [Bus92, Ch. 4])**.** *Let γ be a simple closed geodesic of length ℓ on a compact hyperbolic surface X. Then the neighbourhood*

$$\mathcal{C}(\gamma) = \{p \in X : d(p,\gamma) \le w(\ell)\}, \qquad \boxed{\sinh w(\ell) = \frac{1}{\sinh(\ell/2)}}$$

is an embedded cylinder, isometric to $[-w(\ell), w(\ell)]_\rho \times (\mathbb{R}/\mathbb{Z})_t$ with the metric (15):

$$ds^2 = d\rho^2 + \ell^2 \cosh^2\rho \; dt^2.$$

Moreover:

(i) collars around disjoint simple closed geodesics are disjoint;

(ii) every closed geodesic of length $< 2\operatorname{arcsinh}(1) \approx 1.76$ *is simple, and any two such are disjoint.*

Where this comes from. Locally the picture is Example 1.18: γ lifts to a geodesic $A \subset \mathbb{H}$, translated by an element $g \in \Gamma$ of length ℓ, and $\langle g \rangle \backslash \mathbb{H}$ is exactly the cylinder C_ℓ. Being a collar means surviving the further quotient down to X: the w-tube around γ embeds precisely when the w-tube around A meets no other lift $h(A)$, $h \in \Gamma \setminus \langle g \rangle$.

Keeping the lifts apart is one inequality, proved by cutting the configuration into right-angled hexagons:

if two closed geodesics of lengths ℓ_1, ℓ_2 *cross, then* $\sinh(\ell_1/2)\sinh(\ell_2/2) > 1$.

Both halves of the theorem come out of it. Applied to a geodesic and itself, a self-crossing forces $\sinh(\ell/2) > 1$, that is $\ell > 2\operatorname{arcsinh} 1$ — which is (ii). Applied to A and its translates, the same identities put the tubes in contact exactly when $\sinh w \sinh(\ell/2) = 1$ — which is the stated $w(\ell)$. See [Bus92, Thm. 4.1.1]. □

Read the formulas as $\ell \to 0$:

$$w(\ell) \approx \log \frac{4}{\ell} \to \infty, \qquad \operatorname{Area}(\mathcal{C}) = \int_0^1 \int_{-w}^{w} \ell \cosh \rho \, \mathrm{d}\rho \, \mathrm{d}t = 2\ell \sinh w(\ell) = \frac{2\ell}{\sinh(\ell/2)} \to 4.$$

A short geodesic sits at the waist of an ever longer tube of essentially fixed area: pinching is never free — it always drags a mandatory cylinder along. Also note the consequence of (i)+(ii): a surface of genus g carries at most $3g-3$ geodesics of length $< 2\operatorname{arcsinh} 1$ (disjoint simple curves on a genus-g surface number at most $3g-3$); *degeneracy is finite and localised.*

Corollary 2.10 (Thick–thin decomposition)**.** *Fix* $\mu < \operatorname{arcsinh} 1$. *Every compact hyperbolic surface decomposes as* $X = X_{\text{thin}} \sqcup X_{\text{thick}}$, *where* $X_{\text{thin}} = \{p : \operatorname{inj}(p) < \mu\}$ *is a disjoint union of subsets of collars of geodesics shorter than* 2μ, *and every point of* X_{thick} *is the centre of an embedded hyperbolic disk of radius* μ.

Proof. If $\operatorname{inj}(p) < \mu$ there is a geodesic loop at p of some length $L < 2\mu$; being homotopically nontrivial it is freely homotopic to a closed geodesic γ with $\ell = \ell(\gamma) \leq L < 2\operatorname{arcsinh} 1$, so γ is simple by Theorem 2.9(ii). The quick exercise below gives $d(p,\gamma) < w(\ell)$, i.e. $p \in \mathcal{C}(\gamma)$, and these collars are disjoint by (i). Contrapositively, a point lying outside all of them satisfies $\operatorname{inj}(p) \geq \mu$, and is therefore the centre of an embedded disk of radius μ. □

Quick exercise 2.11. Let g be a hyperbolic isometry of $\mathbb{H}$ with translation length ℓ and axis A. Taking $A = i\mathbb{R}_{>0}$, verify

$$\sinh\left(\tfrac{1}{2} d(x, gx)\right) = \cosh\Big(d(x, A)\Big) \sinh(\ell/2).$$

Deduce that a geodesic loop of length $L < 2\mu$ at p forces $d(p,\gamma) < w(\ell)$, and locate

where the hypothesis $\mu < \operatorname{arcsinh} 1$ is used. (Note $\cosh w(\ell) = \coth(\ell/2)$; the point is that $\sinh(L/2) < \sinh \mu \le 1 \le \cosh(\ell/2)$.)

This is the surface case of the Margulis lemma; for the general nonpositively curved statement see Ballmann–Gromov–Schroeder [BGS85]. In Lecture 3, Benjamini–Schramm convergence will say that for a random surface the thin part occupies a negligible fraction of the area.

2.5 Small eigenvalues located: test functions, and Schoen–Wolpert–Yau

2.5.1 The optimal test function across a collar

How cheaply can a function cross a collar from -1 to $+1$? By (15), for $f = f(\rho)$,

$$E(f) = \int_{\mathcal{C}} |\nabla f|^2 \, \mathrm{d}A = \int_{-w}^{w} f'(\rho)^2 \, \ell \cosh \rho \, \mathrm{d}\rho.$$

Minimise subject to $f(\pm w) = \pm 1$: the Euler–Lagrange equation $(\cosh \rho \, f')' = 0$ gives $f'(\rho) = c/\cosh \rho$ — *the transition concentrates at the thin waist, where crossing is cheapest.* With the boundary conditions,

$$E_{\min} = \frac{4\ell}{I_w}, \qquad I_w = \int_{-w}^{w} \frac{\mathrm{d}\rho}{\cosh \rho} = 2 \arctan(\sinh w) \xrightarrow[\ell \to 0]{} \pi, \qquad \text{so} \quad E_{\min} \approx \frac{4\ell}{\pi}. \tag{17}$$

The antiderivative of sech is $\arctan \circ \sinh$; the limit is then immediate, since $\sinh w(\ell) = 1/\sinh(\ell/2) \to \infty$ and $\arctan \to \frac{\pi}{2}$. So the π is an arctangent, not a coincidence. Two morals. First, the cost is *linear* in ℓ — the naïve linear-in-ρ profile gives only $E \asymp 1/(\log \frac{1}{\ell})^2$, substantially worse; optimising the profile matters (Exercise 2.2 explores this). Second, the constant π is sharp and will match the true asymptotics of λ_1 for pinched surfaces.

Proposition 2.12 (Pinching bound)**.** *Let α be a union of disjoint simple closed geodesics of total length $L \le 1$ separating X into $X_1 \sqcup X_2$. Then*

$$\lambda_1(X) \;\le\; \frac{4L}{\pi} \cdot \frac{1}{\min\{\operatorname{Area}(X_1), \operatorname{Area}(X_2)\} - 4N} \, (1 + o(1)),$$

where N is the number of curves in α (each collar has area < 4); in particular for the genus-2 dumbbell with waist $\ell \to 0$,

$$\lambda_1(X_\ell) \;\le\; \frac{\ell}{\pi(\pi - 1)} \, (1 + o(1)) \;=\; O(\ell).$$

Proof. Take $F = \pm 1$ on the two components off the collars, the optimal profile (17) inside each collar. Then $E(F) \le 4L/I_w$ (energies add over collars, and $\sum 4\ell_i/I_{w_i} \le 4L/\min I \to 4L/\pi$), $\int F^2 \ge \operatorname{Area}(X) - \sum \operatorname{Area}(\mathcal{C}_i)$, and after subtracting the mean the variance is at least $\frac{4A_1'A_2'}{A_1'+A_2'} \ge \min(A_1', A_2')$, where A_i' are the off-collar areas (the variance is $\min_c \int_X (F-c)^2$; discarding the collars leaves $A_1'(1-c)^2 + A_2'(1+c)^2$, minimised at

$c = \frac{A_1' - A_2'}{A_1' + A_2'}$). Apply Theorem 1.4. For the dumbbell: $A_1' = A_2' \to 2\pi - 2$, mean zero by symmetry, and $\mathcal{R}(F) \to \frac{4\ell/\pi}{4\pi - 4} = \frac{\ell}{\pi(\pi-1)}$. □

2.5.2 The crucial asymmetry, and the Schoen–Wolpert–Yau theorem

Everything above required α to *separate*. If a short geodesic does not disconnect X, there is no cut, no Cheeger candidate, no ± 1 test function — and indeed no small eigenvalue: a surface with one short *nonseparating* geodesic can keep λ_1 bounded below — made precise below on the model cylinder. The spectrum is insensitive to short geodesics; what matters is short *separating systems*. Schoen, Wolpert and Yau turned this into a sharp two-sided statement:

Definition 2.13. For $1 \le i \le 2g - 3$, let $L_i(X)$ be the minimal total length of a union of disjoint simple closed geodesics cutting X into $i + 1$ connected pieces.

Theorem 2.14 (Schoen–Wolpert–Yau 1980 [SWY80])**.** *For each g there are constants $0 < \alpha(g) < \beta(g)$ such that for every closed hyperbolic surface of genus g and every $1 \le i \le 2g - 3$,*

$$\alpha(g)\, L_i(X) \;\le\; \lambda_i(X) \;\le\; \beta(g)\, L_i(X).$$

The upper bound is Proposition 2.12 applied to an optimal system for L_i (with an $(i+1)$-dimensional space of piecewise-constant test functions and min–max). The lower bound is the hard converse: *any* eigenfunction with small eigenvalue must essentially organise itself as locally-constant blocks separated by short geodesic systems. Q1 is thereby answered completely: *each λ_i is, up to constants depending only on the genus, the length L_i of an optimal separating multigeodesic.*

On the model cylinder C_ℓ this can be made precise: for every $\ell > 0$,

$$\inf \operatorname{spec}\left(\Delta, L^2(C_\ell)\right) = \frac{1}{4}:$$

an arbitrarily short waist does not push the bottom of the spectrum below $\frac{1}{4}$. The proof is Exercise 2.2(d). The contrast with Proposition 2.12 is sharp: on a compact surface a short *separating* geodesic makes λ_1 small, but on the cylinder both sides of the cut have infinite area, so no admissible ± 1 test function exists.

2.6 How many small eigenvalues?

A maximal system of disjoint simple closed geodesics has $3g - 3$ elements and cuts X into $2g - 2$ pairs of pants (three-holed spheres; see §3.1). Pinching *all* of them produces $2g - 2$ nearly-isolated components; the space of functions constant on each has dimension $2g - 2$; by min–max, $2g - 2$ eigenvalues (counting λ_0) must be small:

Theorem 2.15 (see [Bus92, Ch. 8]; part (ii) is Otal–Rosas [OR09])**.**

(i) For every $\varepsilon > 0$ and $g \ge 2$ there exist closed hyperbolic surfaces of genus g with $2g - 2$ eigenvalues (counting $\lambda_0 = 0$) in $[0, \varepsilon]$.

(ii) Every closed hyperbolic surface of genus g satisfies $\lambda_{2g-2} > \frac{1}{4}$: at most $2g-2$ eigenvalues lie in $[0, \frac{1}{4}]$. By (i), the count $2g-2$ is sharp.

(iii) No topological bound holds above $\frac{1}{4}$: for every $\varepsilon > 0$ and $g \geq 2$ there are genus-g surfaces with arbitrarily many eigenvalues below $\frac{1}{4} + \varepsilon$.

Part (ii) was conjectured by Buser after proving (i), and resisted for three decades; the Otal–Rosas proof analyses the topology of nodal sets of eigenfunctions (each eigenvalue $\leq \frac{1}{4}$ forces an incompressible piece of the surface, and there is room for only $2g-2$ of those). Part (iii): Randol-type constructions with many short *nonseparating* geodesics — these produce eigenvalues just above $\frac{1}{4}$ (from the continuous spectrum of the limiting cylinders), unboundedly many.

Remark 2.16 (Courant's theorem, for calibration). A classical counterpart from general spectral geometry (see [LMP23, Ch. 4]): the k-th eigenfunction of any closed surface has at most $k+1$ nodal domains (Courant). Buser's construction (i) saturates a geometric version of this budget — $2g-2$ nearly-disconnected pants, one sign block each — while Otal–Rosas (ii) is the deeper, hyperbolic-specific statement that below $\frac{1}{4}$ the *topology* (not just the index k) caps the count. The two bounds have entirely different natures: Courant is soft analysis, valid everywhere; Otal–Rosas uses hyperbolic geometry irreducibly.

Assemble what we now know about the eigenvalues below and above $\frac{1}{4}$ (Figure 2):

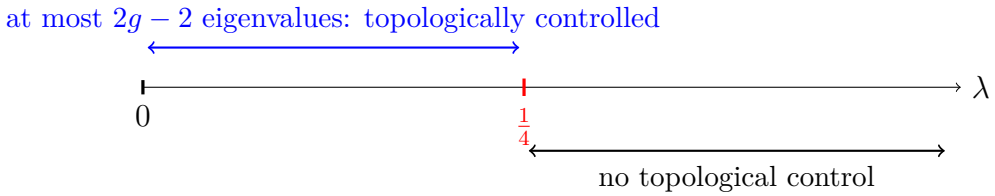


Figure 2: Below $\frac{1}{4}$ the number of eigenvalues is capped by the topology at $2g-2$, sharply; above $\frac{1}{4}$ no such bound holds.

The number $\frac{1}{4}$ — the bottom of $\mathrm{spec}(\mathbb{H})$ — separates the tame window from the wild part of the spectrum. Third appearance of $\frac{1}{4}$; time to explain the pattern.

2.7 The ceiling: no gap beats $\frac{1}{4}$ in large genus

The heuristic first. Large genus means large area ($4\pi(g-1)$, Gauss–Bonnet), hence a *large* surface; a large hyperbolic surface locally resembles its universal cover $\mathbb{H}$; and its spectrum should be pushed toward $\mathrm{spec}(\mathbb{H}) = [\frac{1}{4}, \infty)$. Two quantitative inputs make this a theorem.

Lemma 2.17 (Diameter grows)**.** *Every closed hyperbolic surface of genus g has* $\mathrm{diam}(X) \geq \log\big(4(g-1)\big)$.

Proof. A metric ball $B(p,r) \subset X$ is the image under $\mathbb{H} \to X$ of a ball of the same radius, so $\operatorname{Area} B(p,r) \le \operatorname{Area} B_{\mathbb{H}}(r) = 4\pi \sinh^2(r/2) \le \pi e^r$. Take $r = \operatorname{diam}(X)$: then the ball is all of X, so $\pi e^{\operatorname{diam}} \ge \operatorname{Area}(X) = 4\pi(g-1)$. □

The proof combines two facts from Lecture 1: exponential volume growth (13) and Gauss–Bonnet. Note how weak the conclusion is: the area grows linearly in g, while the diameter is forced to grow only *logarithmically*. Exponential volume growth is exactly what permits this — a ball of radius r already has area $\asymp e^r$, so radius log(area) suffices to cover X — and the order is sharp: random surfaces achieve $\operatorname{diam} = O(\log g)$ (Lecture 3).

The second input needs a small computation with independent interest: how cheap can a compactly supported function on $\mathbb{H}$ be? By McKean, never cheaper than $\frac{1}{4}$; the point is the rate.

Lemma 2.18 (Dirichlet gap of a hyperbolic disk)**.** *Let $B_R \subset \mathbb{H}$ be a disk of radius $R \ge 2$. There is a smooth $f \ne 0$ vanishing outside B_R with*

$$\mathcal{R}(f) \;\le\; \frac{1}{4} + \frac{C}{R^2} \qquad (C \text{ universal}).$$

Proof. Work in polar coordinates (12) at the centre: $\mathrm{d}A = \sinh r \,\mathrm{d}r\,\mathrm{d}\theta$, and for radial f, $|\nabla f|^2 = f'(r)^2$. The McKean-sharp profile is $e^{-r/2}$ (the radial analogue of $y^{1/2}$: it equidistributes mass against $\sinh r \approx \frac{1}{2}e^r$), so take $f(r) = e^{-r/2}\chi(r)$ with χ a cutoff equal to 1 on $[0, R/2]$ and vanishing at R. Write $\mathrm{d}\mu = e^{-r}\sinh r\,\mathrm{d}r$, so that $\int f^2\,\mathrm{d}A = 2\pi\int_0^R \chi^2\,\mathrm{d}\mu$; since $e^{-r}\sinh r \to \frac{1}{2}$, this is $\asymp R$. From $f' = e^{-r/2}(\chi' - \frac{1}{2}\chi)$,

$$\int |\nabla f|^2\,\mathrm{d}A = 2\pi\int_0^R \left(\tfrac{1}{4}\chi^2 - \chi\chi' + \chi'^2\right)\mathrm{d}\mu.$$

The middle term integrates to $O(1)$ (it is $-\frac{1}{2}\int(\chi^2)'\,\mathrm{d}\mu$, and μ has bounded density), so everything turns on $\int \chi'^2\,\mathrm{d}\mu$. A linear cutoff has $|\chi'| \asymp \frac{1}{R}$ on a set of μ-mass $\asymp R$, contributing $\asymp \frac{1}{R}$ and giving only $\mathcal{R}(f) \le \frac{1}{4} + O(1/R)$. Choosing instead the cosine profile $\chi(r) = \cos\left(\frac{\pi(r-R/2)}{R}\right)$ on $[R/2, R]$ makes $\int \chi'^2\,\mathrm{d}\mu = O(1/R)$ against a denominator $\asymp R$, giving the stated $\frac{1}{4} + C/R^2$ — the usual $(\text{length})^{-2}$ law. (Exercise 1.3(c*) is the same computation in the half-plane picture.) □

Theorem 2.19 (Diameter bound for the gap; Cheng [Che75])**.** *There is a universal constant C such that every closed hyperbolic surface satisfies*

$$\lambda_1(X) \;\le\; \frac{1}{4} + \frac{C}{\operatorname{diam}(X)^2}.$$

Proof. Let p, q realise the diameter d and let $B_1 = B(p, d/2)$, $B_2 = B(q, d/2)$: they are disjoint. The input is Cheng's comparison of Dirichlet eigenvalues *of balls*: on a surface with $K \ge -1$, the first Dirichlet eigenvalue of a ball of radius r is at most that of the hyperbolic ball of radius r, which is $\le \frac{1}{4} + C/r^2$ by Lemma 2.18. (See Appendix B, *Cheng's comparison, sharp form.*) With $r = d/2$ this gives f_1, f_2 with disjoint supports

and $\mathcal{R}(f_i) \le \frac{1}{4} + C/d^2$. Their span is 2-dimensional, so min–max (Theorem 1.4) yields $\lambda_1 \le \max_i \mathcal{R}(f_i)$. □

Corollary 2.20 (Huber 1974 [Hub74])**.**

$$\sup_{X \in \mathcal{M}_g} \lambda_1(X) \;\le\; \frac{1}{4} + \frac{C}{(\log g)^2}, \qquad \textit{so} \qquad \limsup_{g\to\infty} \sup_{X \in \mathcal{M}_g} \lambda_1(X) \;\le\; \frac{1}{4}.$$

In large genus, *nothing beats* $\frac{1}{4}$: the bottom of the spectrum of the universal cover is a hard asymptotic ceiling. In fixed genus the ceiling is invisible (Bolza: $\lambda_1 \approx 3.84$) — a surface must be large before it "feels" its universal cover. Q2 is answered — *provided the ceiling is attained.* Is it?

2.8 Who attains the ceiling?

Conjecture 2.21 (Buser 1984)**.** *There exist closed hyperbolic surfaces X_g with $g \to \infty$ and $\lambda_1(X_g) \to \frac{1}{4}$.*

This modest-sounding statement resisted for four decades. Random-looking gluings produce bottlenecks; symmetric constructions are rigid and scarce. For most of that time, the only serious candidates came from *number theory.*

Example 2.22 (Arithmetic surfaces)**.** The modular group $\mathrm{PSL}_2(\mathbb{Z}) \subset \mathrm{PSL}_2(\mathbb{R})$ is discrete; its *congruence subgroups*

$$\Gamma(N) = \{M \in \mathrm{PSL}_2(\mathbb{Z}) : M \equiv I \pmod N\}$$

give quotients $X(N) = \Gamma(N)\backslash\mathbb{H}$ — non-compact but of finite area, with genus growing in N; compact cousins come from quaternion algebras (e.g. the Bolza surface is arithmetic). Selberg proved in 1965, using the arithmetic structure:

$$\lambda_1(X(N)) \;\ge\; \frac{3}{16} \qquad \text{for all } N,$$

and conjectured the optimal $\lambda_1(X(N)) \ge \frac{1}{4}$ — **Selberg's eigenvalue conjecture**, one of the great open problems at the interface of spectral theory and number theory. The current record is $\lambda_1 \ge \frac{975}{4096} \approx 0.238$ (Kim–Sarnak 2003, via functoriality of automorphic forms). See Sarnak's notes [Sar03] and Selberg [Sel56]. Remember the strange constant $\frac{3}{16}$: it will reappear in Lecture 3 with a completely different justification — and the coincidence has a reason.

So for almost sixty years: a sparse, rigid, arithmetic family with uniform gap $\ge \frac{3}{16}$, conjecturally $\frac{1}{4}$ — and nothing else. The modern resolution of Buser's conjecture came from the opposite direction:

Theorem 2.23 (Hide–Magee 2023 [HM23])**.** *Buser's conjecture is true: suitable* random covers *of a fixed surface produce compact hyperbolic surfaces with* $g \to \infty$ *and* $\lambda_1 \to \frac{1}{4}$.

The contrast is worth recording. Forty years of explicit constructions did not reach $\frac{1}{4}$, while choosing the gluing *at random* does. This suggests that optimality is not exceptional but *generic*, which is precisely Q3:

Q3. What is the spectral gap of a typical hyperbolic surface?

To pose the question honestly we need a probability measure on the set of all genus-g surfaces. That space — moduli space — is the first character of Lecture 3.

What to remember from Lecture 2

1. Cheeger $\lambda_1 \geq h^2/4$ (coarea + Cauchy–Schwarz) and Buser $\lambda_1 \leq 2h + 10h^2$ (needs $K \geq -1$): gap = connectivity.
2. Collar lemma: a length-ℓ simple geodesic carries an embedded cylinder (15) of half-width $\operatorname{arcsinh}(1/\sinh(\ell/2)) \approx \log \frac{4}{\ell}$ and area $\to 4$; collars are disjoint; thick–thin decomposition.
3. Crossing a collar costs energy $\approx 4\ell/\pi$ (sech profile!); short *separating* systems give $\lambda_1 = O(L)$; nonseparating short geodesics are harmless. SWY: $\lambda_i \asymp L_i$.
4. Window $[0, \frac{1}{4}]$: at most $2g-2$ eigenvalues (Otal–Rosas), sharply (Buser); no control above $\frac{1}{4}$.
5. The ceiling: $\lambda_1 \leq \frac{1}{4} + C/(\log g)^2$ (ball growth $\to$ diameter $\to$ Cheng). Attained in the limit — by *random* covers (Hide–Magee).

Exercise session 2 follows this lecture (Appendix E.2): **A4** Cheeger's inequality transplanted to graphs; **B1** inside the collar; **B2** non-backtracking walks.

3 Lecture 3: Random hyperbolic surfaces and the optimal gap

Today we answer Q3 with the main theorem of the course:

$$\forall \varepsilon > 0 : \qquad \mathbb{P}_g^{\mathrm{WP}}\left(\lambda_1(X) \;\geq\; \tfrac{1}{4} - \varepsilon\right) \xrightarrow[g\to\infty]{} 1 \qquad \text{(Anantharaman–Monk 2025).}$$

The route: build the probability space (§3.1); learn to integrate over it (§3.2); get a first uniform gap via Cheeger (§3.3); understand why $\frac{1}{4}$ should be the truth (§3.4); learn the trace method in its simplest setting, random graphs (§3.5); and run it on surfaces (§3.6).

3.1 Moduli space as a probability space

"Random surface" requires a *set* of all genus-g surfaces and a *canonical* measure. Both exist, and the construction is beautiful in its own right. Three steps: building blocks, coordinates, measure.

Step 1: pairs of pants, the building blocks of surfaces. A *pair of pants* is a sphere with three disjoint open disks removed, endowed with a hyperbolic metric in which the three boundary curves are geodesics.

> **Fact (hyperbolic trigonometry; [Bus92, Ch. 3]).** For every triple $(\ell_1, \ell_2, \ell_3) \in \mathbb{R}^3_{>0}$ there exists a hyperbolic pair of pants with boundary lengths ℓ_1, ℓ_2, ℓ_3, and it is *unique* up to isometry.
> *Why: cut the pants along its three* ***seams*** *— the geodesic arcs joining the boundary circles pairwise, each meeting both perpendicularly — so that each half is a right-angled hexagon (Figure 3), and right-angled hexagons are rigid with three alternating side-lengths freely prescribable. Existence and uniqueness for arbitrary lengths is deeply hyperbolic — no flat analogue exists.*

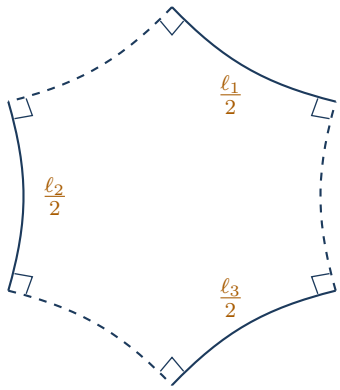


Figure 3: Half a pair of pants: a right-angled hyperbolic hexagon. Prescribing $\frac{\ell_1}{2}, \frac{\ell_2}{2}, \frac{\ell_3}{2}$ determines the hexagon, hence the pants.

Pants are the *building blocks*: rigid pieces with three dials (ℓ_1, ℓ_2, ℓ_3) and no further moduli.

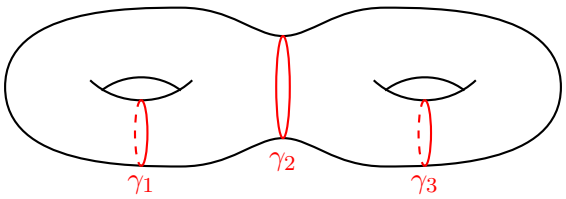


Figure 4: A pants decomposition of a genus-2 surface: $3g-3=3$ curves, $2g-2=2$ pairs of pants.

Step 2: Fenchel–Nielsen coordinates. Fix a genus $g \geq 2$. Cutting X along $3g-3$ disjoint simple closed geodesics decomposes it into $2g-2$ pairs of pants, as in Figure 4 (count: each pants has $\chi=-1$ and $\chi(X)=2-2g$; each of the $3g-3$ curves bounds two pants sides). Conversely, rebuild X from the pants by re-gluing, and record for each cuff i:

$$\ell_i > 0 \quad \text{(common boundary length)}, \qquad \tau_i \in \mathbb{R} \quad (\textit{twist}\text{: relative rotation before gluing}).$$

A *marking* of X is the isotopy class of a homeomorphism $\phi : \Sigma_g \to X$ from a fixed reference surface: it pins down not just the cuffs γ_i but how *every* curve of Σ_g is drawn on X. This is why τ_i is an honest real number and not merely an angle mod ℓ_i. Twisting by a full turn $\tau_i \mapsto \tau_i + \ell_i$ changes the marking by the Dehn twist along γ_i, which fixes every γ_j up to isotopy: the surface and its pants decomposition are unchanged. What moves are the curves *crossing* γ_i, which come back spiralling once around it — and no isotopy undoes that. Bookkeeping with markings gives:

> **Theorem (Fenchel–Nielsen).** The map $(\ell_i, \tau_i)_{i=1}^{3g-3}$ is a bijection (indeed a diffeomorphism)
>
> $$\mathcal{T}_g \;\cong\; (\mathbb{R}_{>0} \times \mathbb{R})^{3g-3},$$
>
> where $\mathcal{T}_g$ — *Teichmüller space* — is the space of *marked* genus-g hyperbolic surfaces.

This is why we *can* pinch: lengths are free parameters (set one $\ell_i = \varepsilon$: the dumbbell of Lecture 1 was a legitimate point of $\mathcal{T}_g$ all along), and why $\dim = 6g-6$.

To pass from marked to honest surfaces, forget the marking: the *mapping class group* MCG_g (self-homeomorphisms of the surface up to isotopy — a countable discrete group, generated by Dehn twists) acts on $\mathcal{T}_g$ by changing markings, and the **moduli space** of genus-g hyperbolic surfaces is

$$\mathcal{M}_g = \mathcal{T}_g / \mathrm{MCG}_g.$$

Step 3: the Weil–Petersson measure. On $\mathcal{T}_g$, in Fenchel–Nielsen coordinates, take Lebesgue measure — more precisely the symplectic volume of

$$\omega_{\mathrm{WP}} = \sum_{i=1}^{3g-3} d\ell_i \wedge d\tau_i, \qquad d\mathrm{Vol}_{\mathrm{WP}} = \prod_{i=1}^{3g-3} d\ell_i \, d\tau_i.$$

A priori this depends on the chosen pants decomposition. The remarkable fact:

Theorem (Wolpert 1982 [Wol82]). ω_{WP} is independent of the pants decomposition used to define it, and invariant under MCG_g. Hence $d\mathrm{Vol}_{\mathrm{WP}}$ descends to $\mathcal{M}_g$.

(Where the name comes from: Weil equipped $\mathcal{T}_g$ with a metric built out of an inner product of Petersson's. Its definition mentions no pants decomposition, so the measure it induces is manifestly independent of one — and Wolpert's theorem is the computation that this intrinsic measure is exactly the elementary $\prod d\ell_i\, d\tau_i$ above. We never need the metric itself.) Moreover the total volume

$$V_g \coloneqq \mathrm{Vol}_{\mathrm{WP}}(\mathcal{M}_g)$$

is *finite* — not obvious, since the ℓ_i range over $(0,\infty)$; finiteness uses Bers' theorem that every surface admits a pants decomposition with all cuffs shorter than a constant B_g (Bers' constant, $\leq Cg$; [Bus92, Ch. 5]), so a fundamental domain has bounded lengths, and the twist directions contribute $\prod \ell_i \leq B_g^{3g-3}$ per unit cell. Normalising:

Definition 3.1 (The WP random surface)**.** $\mathbb{P}_g^{\mathrm{WP}} \coloneqq \mathrm{Vol}_{\mathrm{WP}}/V_g$, a probability measure on $\mathcal{M}_g$. "**With high probability (whp)**" means: with $\mathbb{P}_g^{\mathrm{WP}}$-probability $\to 1$ as $g \to \infty$. The model was popularised by Mirzakhani [Mir13, Mir10]; it is the continuous analogue of the uniform random d-regular graph.

Remark 3.2 (Why this model?). It requires no arbitrary choices (Wolpert), it is compatible with hyperbolic geometry (lengths and twists are the coordinates), and — decisive for us — it is *computable*: expectations of geodesic counts reduce to volume computations, by Mirzakhani's formula, to which we now turn.

3.2 Mirzakhani's toolkit: integrating over moduli space

Every spectral argument to come needs averages, over $\mathcal{M}_g$, of quantities assembled from geodesic lengths. Pick any simple closed curve γ on the topological surface underlying X (say a nonseparating one) and let

$$\mathrm{Orb}(\gamma) = \{\phi(\gamma) : \phi \in \mathrm{MCG}_g\}$$

be its orbit under the mapping class group. There is a single MCG_g-orbit of nonseparating simple closed curves, so $\mathrm{Orb}(\gamma)$ does not depend on which such γ we started from — it is simply *the set of all closed geodesics on X of that type*. (Separating curves are more refined: cutting along one splits the genus as $h + (g-h)$, and curves with different h lie in different orbits; "topological type" below always means an orbit, so a separating curve's type records h too — this is why the theorem just below treats separating curves of type $(h, g-h)$ as their own case.) Now let $f : (0,\infty) \to \mathbb{R}$ be any (measurable, suitably decaying) function — a weight applied to lengths, ours to choose — and define the associated *geometric function*

$$F(X) = \sum_{\gamma' \in \mathrm{Orb}(\gamma)} f(\ell_{\gamma'}(X)),$$

where $\ell_{\gamma'}(X)$ denotes the length of the *unique* closed geodesic freely homotopic to γ' on the surface X — unique by Theorem 1.20(iii). Note the shift of viewpoint this forces: γ' is a *topological* curve, fixed once and for all, while X ranges over $\mathcal{M}_g$. So $\ell_{\gamma'}$ is a function *on moduli space*, and therefore so is F — which is what makes it something we can integrate. It is a sum with one term for every geodesic of that type on X. It is a genuine *count* when f is an indicator: take $f = \mathbf{1}_{[0,L]}$ and $F(X)$ is the number of geodesics of that type shorter than L — how every counting statement below will be proved. Other choices of f weight the geodesics instead, and we will need those too.) The obstacle: $\mathcal{M}_g$ is a quotient by the huge group MCG_g, unpleasant to integrate over. Mirzakhani's insight: *the orbit is not the obstacle but the solution* — summing over it **unfolds** the quotient.

Theorem 3.3 (Mirzakhani's integration formula [Mir07a])**.** *Let γ be a nonseparating simple closed curve and $F(X) = \sum_{\gamma' \sim \gamma} f(\ell_{\gamma'}(X))$ (sum over the* MCG*-orbit, i.e. all nonseparating simple closed geodesics). Then*

$$\int_{\mathcal{M}_g} F(X)\,\mathrm{d}X \;=\; \frac{1}{2}\int_0^\infty t\, f(t)\, V_{g-1,2}(t,t)\,\mathrm{d}t,$$

where $V_{g,n}(b_1,\dots,b_n)$ denotes the WP volume of the moduli space $\mathcal{M}_{g,n}(b)$ of genus-g surfaces with n boundary geodesics of prescribed lengths $b_1,\dots,b_n$. For a separating curve of type $(h, g-h)$ the right-hand side is $\frac{1}{2}\int_0^\infty t f(t) V_{h,1}(t) V_{g-h,1}(t)\,\mathrm{d}t$ (with an extra $\frac{1}{2}$ when $h = g-h$ by symmetry); analogous formulas hold for arbitrary multicurves. (Constant conventions — the $\frac{1}{2}$'s — vary across the literature and never affect orders of magnitude.)

The unfolding. Two actions of MCG_g are in play: on Teichmüller space $\mathcal{T}_g$, by changing the marking, with quotient $\mathcal{M}_g$; and on curves, where $\mathrm{Orb}(\gamma)$ is one orbit. Let $\mathrm{Stab}(\gamma) \subset \mathrm{MCG}_g$ be the mapping classes fixing γ up to isotopy. Then $\phi \mapsto \phi(\gamma)$ is a bijection from $\mathrm{MCG}_g/\mathrm{Stab}(\gamma)$ onto $\mathrm{Orb}(\gamma)$: two mapping classes carry γ to the same curve exactly when they differ by an element of $\mathrm{Stab}(\gamma)$.

Summing over the orbit therefore trades the quotient by all of MCG_g for the quotient by $\mathrm{Stab}(\gamma)$ alone,

$$\int_{\mathcal{M}_g} F(X)\,\mathrm{d}X \;=\; \int_{\mathcal{T}_g/\mathrm{Stab}(\gamma)} f\big(\ell_\gamma(X)\big)\,\mathrm{d}X,$$

because the translates of a fundamental domain for MCG_g, one for each element of the orbit, tile a fundamental domain for $\mathrm{Stab}(\gamma)$. The bigger space $\mathcal{T}_g/\mathrm{Stab}(\gamma)$ is the space of surfaces carrying *one* distinguished curve of type γ.

It remains to recognise that space. Cutting along the distinguished curve leaves a connected surface of genus $g-1$ with two boundary circles of the same length $t = \ell_\gamma$, and conversely such a surface plus a twist reglues. Take a pants decomposition whose first cuff is γ; by Wolpert the measure is $\prod d\ell_i\, d\tau_i$, which factors as $d\ell_\gamma\, d\tau_\gamma$ times the coordinates of the cut surface. The twist is read modulo the Dehn twist along γ, which shifts it by ℓ_γ, so it runs over a circle of circumference t and integrates out to a factor t. The remaining $3g-4$ pairs (ℓ_i, τ_i) are exactly Fenchel–Nielsen coordinates for the cut surface; so integrating them out gives its moduli volume $V_{g-1,2}(t,t)$. Hence $\int_0^\infty t\, f(t)\, V_{g-1,2}(t,t)\,\mathrm{d}t$.

The $\frac{1}{2}$ is the class in $\mathrm{Stab}(\gamma)$ exchanging the two sides of γ, equivalently the two boundary circles of the cut surface. □

Geodesic counting on a random surface is thereby converted into volume asymptotics. Which Mirzakhani then computed:

Theorem 3.4 (Volume polynomials and asymptotics [Mir07a, Mir07b, Mir13])**.**

(i) $V_{g,n}(b_1,\dots,b_n)$ *is a* polynomial *in* $b_1^2,\dots,b_n^2$ *of total degree* $3g-3+n$*, with positive coefficients; e.g.*

$$V_{0,3}(b_1,b_2,b_3)=1, \qquad V_{1,1}(b)=\frac{b^2+4\pi^2}{24}.$$

(ii) As $g\to\infty$*, for fixed* n*:*

$$\frac{V_{g,n+1}}{2g\,V_{g,n}}\to 4\pi^2, \qquad \frac{V_{g,n}}{V_{g-1,n+2}}\to 1, \qquad V_g \;\sim\; \frac{\alpha}{\sqrt{g}}\,(2g-3)!\,(4\pi^2)^{2g-3}\ \ (\exists\,\alpha>0),$$

and, crucially, the polynomials are controlled by their constant terms:

$$V_{g,n}(b_1,\dots,b_n)\;\le\; V_{g,n}\prod_{i=1}^{n}\frac{\sinh(b_i/2)}{b_i/2}.$$

Note the exponent in the sinh bound: boundary lengths enter with weight at most $e^{b/2}$ — *half* the exponential growth rate e^b of geodesic counts. The same exponent $\frac{1}{2}$ returns in §3.6 as the square-root gain.

Example 3.5 (Random surfaces have no short geodesics)**.** Let $N_\varepsilon(X)$ be the number of nonseparating simple closed geodesics of length $\le\varepsilon\le 1$. By Theorems 3.3 and 3.4,

$$\mathbb{E}_g^{\mathrm{WP}}[N_\varepsilon]=\frac{1}{2V_g}\int_0^\varepsilon t\,V_{g-1,2}(t,t)\,\mathrm{d}t \;\le\; \frac{V_{g-1,2}}{2V_g}\Big(\frac{\sinh(\varepsilon/2)}{\varepsilon/2}\Big)^2\frac{\varepsilon^2}{2}$$
$$=\;\big(1+o(1)\big)\,\frac{\varepsilon^2}{4}\big(1+O(\varepsilon^2)\big).$$

By Markov's inequality, $\mathbb{P}_g^{\mathrm{WP}}(N_\varepsilon\ge 1)\le\mathbb{E}_g^{\mathrm{WP}}[N_\varepsilon]=O(\varepsilon^2)$. This concerns nonseparating simple curves only, so two remarks are needed before it says anything about the systole. A closed geodesic of length ≤ 1 is automatically *simple*, since $1<2\,\mathrm{arcsinh}\,1\approx 1.76$ (Theorem 2.9(ii)). And the *separating* case, run through the same machine with $V_{h,1}V_{g-h,1}$ in place of $V_{g-1,2}$, is smaller still — part (ii) makes those products tiny (this is the computation of the next paragraph). Hence $\mathbb{P}_g^{\mathrm{WP}}(\mathrm{sys}(X)\le\varepsilon)=O(\varepsilon^2)$; Mirzakhani proves the matching lower bound, so $\mathbb{P}_g^{\mathrm{WP}}(\mathrm{sys}\le\varepsilon)\asymp\varepsilon^2$. Interpretation: whp there are no *very* short geodesics, but at every fixed scale ε short geodesics do appear with positive probability $\asymp\varepsilon^2$ — the model tolerates thin necks, it just makes them rare. (Where does ε^2 come from? One factor ε from the length range, one from the twist range $\asymp\ell$: the WP measure of the "thin part" of $\mathcal{M}_g$ is $\asymp\varepsilon^2V_g$.) *Exercise 3.1 runs the same machine for* separating

curves, deriving the $\log g$ *bound quoted below.*
For *separating* geodesics the same machine uses $V_{h,1}V_{g-h,1}/V_g$, and part (ii) makes such products tiny: summing over h,

$$\mathbb{P}_g^{\mathrm{WP}}\Big(\exists\,\text{separating geodesic of length} \le a\log g\Big) = O\Big((\log g)^3\, g^{a/2-1}\Big) \longrightarrow 0 \quad (a<2)\,:$$

whp the shortest separating geodesic is longer than $(2-\varepsilon)\log g$. *Random surfaces have no cheap Cheeger cuts* — the door to the spectral gap.

Remark 3.6 (The length spectrum becomes Poisson). The toolkit computes all moments, not just expectations. The payoff (Mirzakhani–Petri [MP19]): for fixed $0<a<b$, the number of primitive closed geodesics of a WP-random surface with length in $[a,b]$ converges in distribution, as $g\to\infty$, to a Poisson random variable with mean

$$\int_a^b \frac{e^t+e^{-t}-2}{2t}\,\mathrm{d}t \;=\; \int_a^b \frac{\cosh t-1}{t}\,\mathrm{d}t.$$

Short geodesics behave like independent rare events, exactly as short cycles do in a random regular graph. So the dictionary of §3.5 is not confined to the spectral gap. The density $\frac{\cosh t-1}{t}$ is $\approx \frac{t}{2}$ for small t — short geodesics are rare, matching Example 3.5 — and $\sim \frac{e^t}{2t}$ for large t, the growth rate of the prime geodesic theorem (iv).

3.3 Typical geometry and the first uniform spectral gap

Pushing the same method further gives much more [Mir13]: whp as $g\to\infty$,

$$\mathrm{sys}(X) \;\ge\; g^{-o(1)}, \qquad \mathrm{diam}(X) = O(\log g), \qquad \exists\, p:\ \mathrm{inj}(p) \;\ge\; \tfrac{\log g}{6},$$

and no separating multicurve is short. The typical surface is *round, thick, and hard to cut*: the exact opposite of a dumbbell. Given Lecture 2 ("gap = no short separating cuts"), a uniform gap should follow. It does:

Theorem 3.7 (Mirzakhani 2013 [Mir13])**.** *For every* $C < \dfrac{\ln 2}{2\pi+\ln 2} \approx 0.099$*:* $\mathbb{P}_g^{\mathrm{WP}}\big(h(X)\le C\big)\to 0$*. Hence, by Cheeger's inequality,*

$$\mathbb{P}_g^{\mathrm{WP}}\Big(\lambda_1(X) > \tfrac14\Big(\tfrac{\ln 2}{2\pi+\ln 2}\Big)^2 \approx 0.0024\Big) \longrightarrow 1.$$

Architecture of the proof. Step 1 (geodesisation of cuts). Define the *geodesic* Cheeger constant $H(X)$ by restricting the infimum in Definition 2.1 to *separating multigeodesics*: unions of *disjoint* simple closed geodesics cutting X into two connected pieces. If the smaller piece has Euler characteristic $-i$, its area is $2\pi i$ by Gauss–Bonnet, so the ratio being minimised is $\ell(\alpha)/2\pi i$. Using the classification of isoperimetric minimisers on hyperbolic surfaces (curves of constant curvature: circles, geodesics, or equidistants of geodesics), Mirzakhani proves

$$\frac{H(X)}{H(X)+1} \;\le\; h(X) \;\le\; H(X)\,:$$

an arbitrary competitor cut can be traded for a geodesic system at a controlled loss. This is where the specific constant comes from: her bound $H \geq \frac{\ln 2}{2\pi}$ whp translates to $h \geq \frac{\ln 2}{2\pi+\ln 2}$.

Step 2 (first-moment method). It remains to rule out short separating geodesic systems. For each topological type of multicurve cutting off Euler characteristic i, the integration formula bounds the expected number of such systems of total length $\leq \frac{\ln 2}{2\pi} \cdot 2\pi i = i \ln 2$; the volume ratios of Theorem 3.4(ii) show these expectations are summably small *unless* the length exceeds $i \ln 2$ — the entropy of the count ($\approx e^L$-many candidate types with total boundary L) is beaten by the measure cost of short boundaries.

Markov turns each of these expectations into a probability. Two summations remain, and they are of different kinds: first over the topological types sharing a given i — there are many, and this is exactly where the entropy estimate above is spent — and then a union bound over the finitely many values $i \leq \frac{1}{2}|\chi(X)| = g - 1$, the smaller side carrying at most half the total Euler characteristic. What survives is: whp *no* separating multigeodesic α at all, whatever its number of components, has *total* length $\ell(\alpha) \leq \frac{\ln 2}{2\pi} \cdot 2\pi i$, where $-i$ is the Euler characteristic of its smaller side. Since $2\pi i$ is that side's area, this says precisely $H(X) > \frac{\ln 2}{2\pi}$ whp. Step 1 converts it, $h \geq \frac{H}{H+1} > \frac{\ln 2}{2\pi+\ln 2}$, and Cheeger's inequality $\lambda_1 \geq h^2/4$ finishes. □

Random surfaces are uniform expanders — the first theorem of its kind (2013). But note the sobering arithmetic: $0.0024 \lll 0.25$. The Cheeger-constant estimate (≈ 0.099) is essentially honest; the damage is done by *squaring* it in Cheeger's inequality — the loss we flagged in Lecture 2. To reach $\frac{1}{4}$, isoperimetry must be abandoned for genuinely spectral tools: one must hear the eigenvalues directly. But first, why should $\frac{1}{4}$ (and not, say, 0.19) be the truth?

3.4 Why $\frac{1}{4}$ should be the truth: Benjamini–Schramm convergence

Definition 3.8 (BS convergence, adapted to surfaces)**.** A sequence of random surfaces *Benjamini–Schramm converges* to $\mathbb{H}$ if for every fixed $r > 0$,

$$\mathbb{E}_g^{\mathrm{WP}}\left[\frac{\mathrm{Area}\{p \in X : \ \mathrm{inj}(p) < r\}}{\mathrm{Area}(X)}\right] \xrightarrow[g\to\infty]{} 0 :$$

around all but a vanishing fraction of points, X is isometric to an r-ball of $\mathbb{H}$.

For WP-random surfaces this holds: $\{p : \mathrm{inj}(p) < r\}$ is contained in the collars of geodesics shorter than $2r$ (thick–thin!), whose expected number is $O_r(1)$ by the computation of Example 3.5 extended to all short geodesics — while $\mathrm{Area}(X) = 4\pi(g-1) \to \infty$. *The typical surface, seen from a uniformly random point, is indistinguishable from the hyperbolic plane.* Its spectrum follows suit:

Theorem 3.9 (Monk 2022 [Mon22]). *Whp, for fixed* $0 \le a < b$,

$$\frac{\#\{j : \lambda_j(X) \in [a,b]\}}{\mathrm{Area}(X)} \longrightarrow \int_a^b \nu_{\mathbb{H}}(\lambda)\,\mathrm{d}\lambda, \qquad \nu_{\mathbb{H}}(\lambda) = \frac{1}{4\pi}\tanh\left(\pi\sqrt{\lambda - \tfrac{1}{4}}\right)\mathbf{1}_{\lambda \ge \frac{1}{4}},$$

with explicit error terms: the macroscopic spectral distribution converges to the spectral measure of the hyperbolic plane, supported on $[\frac{1}{4}, \infty)$. *In particular, whp the number of eigenvalues below* $\frac{1}{4} - \varepsilon$ *is* $O(g^{1-c(\varepsilon)}) = o(g)$.

(The density $\nu_{\mathbb{H}}$ is the Plancherel measure of $\mathbb{H}$; sanity checks: $\nu_{\mathbb{H}} \to \frac{1}{4\pi}$ as $\lambda \to \infty$, recovering Weyl's law $N(\lambda) \sim \frac{\mathrm{Area}}{4\pi}\lambda$; and $\nu_{\mathbb{H}}$ vanishes below $\frac{1}{4}$, as McKean requires. Proof: Selberg's trace formula — §3.6 — with test functions of slowly growing support, plus BS convergence to evaluate the geometric side; a gentle first use of the machine.)

So *on average* the spectrum wants to live above $\frac{1}{4}$. But BS convergence is a statement in density: it cannot rule out a sparse family of low outliers — even a single λ_1 near 0. And Lecture 2 says such outliers correspond to real geometric events (short separating systems) of small but positive probability. The remaining task is precise:

"$o(g)$ exceptional eigenvalues" $\longrightarrow$ "*no* exceptional eigenvalues, whp".

This exact problem was solved once before, in a parallel universe.

3.5 Interlude: the same story for regular graphs

We take the graph side first because it is the simplest place to see the method, not because it came first: Selberg's trace formula dates from 1956, and its graph counterpart — Ihara's zeta function, 1966 — was built on the model of Selberg's. The borrowing runs from surfaces to graphs.

Let G be a d-regular graph on n vertices ($d \ge 3$ fixed), A its adjacency matrix, with spectrum

$$d = \mu_1 > \mu_2 \ge \cdots \ge \mu_n \ge -d, \qquad \mu_+ = \max(\mu_2, |\mu_n|);$$

the spectral gap is $d - \mu_+$, and the same trio holds (Poincaré for functions on vertices; mixing of the random walk; discrete Cheeger cuts, Remark after Theorem 2.6). The role of the universal cover $\mathbb{H}$ is played by the infinite d-regular tree T_d, whose ℓ^2-spectrum is $[-2\sqrt{d-1}, 2\sqrt{d-1}]$ (Kesten): *the number* $2\sqrt{d-1}$ *is the graph counterpart of* $\frac{1}{4}$. The story then runs in exact parallel:

- **Ceiling** (Alon–Boppana; twin of Huber): $\mu_+ \ge 2\sqrt{d-1} - o(1)$ as $n \to \infty$.
- **Arithmetic optimists** (twin of Selberg): *Ramanujan graphs*, with $\mu_+ \le 2\sqrt{d-1}$ exactly, constructed by Lubotzky–Phillips–Sarnak [LPS88] from quaternion arithmetic — here the arithmetic side *won*.
- **Typicality** (Alon's conjecture 1986; twin of our goal): a uniformly random d-regular graph satisfies $\mu_+ \le 2\sqrt{d-1}+\varepsilon$ whp. Proved by Friedman [Fri08]; shorter proofs by Bordenave [Bor20] and, via strong convergence, by Chen–Garza-Vargas–Tropp–van Handel [CGVTvH24].

3.5.1 The trace method

Everything rests on one identity. Powers of A count walks: $(A^k)_{vw} = \#\{\text{walks } v \to w \text{ of length } k\}$, so

$$\operatorname{Tr}(A^{2k}) = \sum_{j=1}^{n} \mu_j^{2k} = \#\{\text{closed walks of length } 2k \text{ in } G\}. \tag{18}$$

Left side: spectrum. Right side: combinatorics. Two immediate uses:

Lower bounds: the Alon–Boppana theorem, in full (Exercise 2.3 develops its non-backtracking refinement).

Theorem 3.10 (Alon–Boppana)**.** *For every sequence of connected d-regular graphs with $n \to \infty$, $\mu_+ \geq 2\sqrt{d-1} - o(1)$.*

Proof. The universal cover of a connected d-regular (simple) graph G is the tree T_d, with covering map $\pi : T_d \to G$; walks in G lift uniquely once a starting lift is chosen, so π injects closed walks of length $2k$ at a root o into closed walks at $\pi(o)$. Closed tree-walks at o number at least $C_k(d-1)^k$: encode such a walk by its distance-from-root profile, a path of ± 1 steps starting and ending at 0 and never going negative — a *Dyck path*, and these are counted by the Catalan number $C_k = \frac{1}{k+1}\binom{2k}{k}$ — and at each of the k descents into a new child choose among $\geq d-1$ children. Since the central binomial coefficient is the largest of $2k+1$ summands totalling 4^k, $C_k \geq \frac{4^k}{(2k+1)(k+1)}$. Summing over the n vertices and comparing with (18):

$$n\, C_k(d-1)^k \;\leq\; \operatorname{Tr}(A^{2k}) \;\leq\; d^{2k} + n\,\mu_+^{2k} \qquad \Longrightarrow \qquad \mu_+^{2k} \;\geq\; C_k(d-1)^k - \frac{d^{2k}}{n}.$$

Choose $k = \lfloor \delta \log n \rfloor$ with $0 < \delta < \frac{1}{2\log d}$: then $d^{2k}/n \to 0$ while $C_k(d-1)^k \to \infty$, so

$$\mu_+ \;\geq\; \Big(C_k(d-1)^k(1-o(1))\Big)^{1/2k} = C_k^{1/2k}\sqrt{d-1}\,(1-o(1)) \;\longrightarrow\; 2\sqrt{d-1},$$

since polynomial factors die under the $\frac{1}{2k}$-th root: $C_k^{1/2k} \to 2$. □

A purely combinatorial count has produced a spectral lower bound, with the universal cover setting the benchmark. Huber's ceiling (Corollary 2.20) is the same mechanism run on surfaces — with Cheng's comparison in place of the walk count.

Upper bounds whp (Friedman). Now average (18) over random G: if

$$\mathbb{E}\Big[\#\{\text{closed walks of length } 2k\}\Big] = \underbrace{n \cdot (\text{tree term})}_{\approx n\,p(k)(d-1)^k} + \text{error},$$

with error small enough, then $\mathbb{E}[\mu_+^{2k}] \lesssim p(k)(d-1)^k$ and, with $2k \gg \log n$, Markov beats the union over eigenvalues. *How small must the error be?* Friedman's definition captures it:

Definition 3.11 (Ramanujan functions [Fri08])**.** $f : \mathbb{N} \to \mathbb{C}$ is *d-Ramanujan* if for

some polynomial p and constant c,

$$|f(k) - p(k)(d-1)^k| \;\le\; c\,k^c\,(d-1)^{k/2} \qquad \text{for all } k.$$

Note the halved exponent: main term $(d-1)^k$, error $(d-1)^{k/2}$ — a **square-root gain**. The exponent is not a convention. If every eigenvalue lies in the tree spectrum $|\mu| \le 2\sqrt{d-1}$, then the non-backtracking counts have size $\mathrm{poly}(\ell)\,(d-1)^{\ell/2}$ and no more (Exercise 2.3(d*)), while a single eigenvalue outside contributes exponentially more; the counts to which the definition is applied are the non-backtracking ones, not $\operatorname{Tr} A^{2k}$ itself. If the expected counts are Ramanujan, the random graph is nearly Ramanujan: $\mu_+ \le 2\sqrt{d-1} + \varepsilon$ whp. A weaker error yields a sub-optimal gap; a stronger one is false. *The square-root gain is exactly what optimal spectral gaps require.* Proving it for random regular graphs is the hard content of Friedman's 100-page memoir: the obstacle is the *tangles*, rare subgraphs that fail to be tree-like — a neighbourhood containing more than one cycle — which would otherwise dominate the average and must be excised before averaging.

The dictionary, assembled for reference, is Table 1.

Graphs	**Hyperbolic surfaces**
d-regular graph, n vertices	genus-g surface, area $4\pi(g-1)$
uniform random d-regular graph	WP random surface
universal cover: the tree T_d	universal cover: $\mathbb{H}$
spectral edge $2\sqrt{d-1}$ (Kesten)	bottom $\frac{1}{4}$ of spec($\mathbb{H}$) (McKean)
Alon–Boppana ceiling	Huber ceiling (Cor. 2.20)
Ramanujan graphs (LPS [LPS88])	Selberg's $\frac{3}{16}$; his $\frac{1}{4}$ conjecture
closed walks; $\operatorname{Tr} A^{2k}$	closed geodesics; Selberg trace formula
non-backtracking walk counts	primitive geodesic counts
tangles (rare non-tree-like subgraphs)	surfaces with atypical short geodesics
Alon conjecture / Friedman [Fri08]	Anantharaman–Monk [AM25]

Table 1: The graphs $\leftrightarrow$ surfaces dictionary.

3.6 The trace method on surfaces: from $\frac{3}{16}$ to the optimal gap

3.6.1 The Selberg trace formula

What replaces (18)? Write eigenvalues as $\lambda_j = \frac{1}{4} + r_j^2$ (so $r_j \in \mathbb{R} \cup i[-\frac{1}{2}, \frac{1}{2}]$, imaginary for exceptional eigenvalues).

Selberg trace formula [Sel56]; [Bus92, Ch. 9]. Let h be even, holomorphic on $|\operatorname{Im} r| \le \frac{1}{2} + \delta$, decaying suitably, with Fourier transform $\widehat{h}$. Then

$$\sum_{j\ge 0} h(r_j) = \frac{\operatorname{Area}(X)}{4\pi} \int_{-\infty}^{\infty} h(r)\, r \tanh(\pi r)\, \mathrm{d}r \;+\; \sum_{\gamma} \sum_{k\ge 1} \frac{\ell_\gamma\, \widehat{h}(k\ell_\gamma)}{2\sinh(k\ell_\gamma/2)},$$

the outer sum over primitive closed geodesics γ, the inner over their iterates.

An *exact identity* between the Laplace spectrum and the length spectrum — the perfected analogue of (18) (closed walks $\leftrightarrow$ closed geodesics; the "tree term" $\leftrightarrow$ the Area term, which is exactly the $\mathbb{H}$-Plancherel contribution of Theorem 3.9). Proof idea: compute Tr $h(\sqrt{\Delta - 1/4})$ two ways — spectrally, and by expanding its kernel over the group Γ, organising by conjugacy classes (which are geodesics, Theorem 1.20(iii)). We use it as a black box; Weyl's law, the prime geodesic theorem, and Theorem 3.9 are among its standard offspring.

How the formula detects small eigenvalues. Take $\widehat{h}$ supported in $[-L, L]$. An exceptional eigenvalue $\lambda_1 = \frac{1}{4} - \beta^2$ has $r_1 = i\beta$ imaginary, so

$$h(i\beta) = \int_{-L}^{L} \widehat{h}(x)\, e^{\beta x}\, \mathrm{d}x \;\approx\; e^{\beta L}\,:$$

exponentially large, while tempered eigenvalues contribute boundedly. Meanwhile the geometric side involves geodesics of length $\leq L$ with weights $\approx e^{-\ell/2}\widehat{h}(\ell)$. Conclusion: **a small eigenvalue must be paid for by an excess of closed geodesics**. No excess (on average) $\Rightarrow$ no small eigenvalue (whp). The strategy is now fully parallel to graphs: average the geometric side over $\mathcal{M}_g$ with Mirzakhani's formula; show the average has no excess; union-bound the exceptional eigenvalue away.

3.6.2 Round 1: the $\frac{3}{16}$ theorem

Theorem 3.12 (Wu–Xue [WX22]; Lipnowski–Wright [LW24] (2022, independent))**.** *Whp,* $\lambda_1 \;\geq\; \frac{3}{16} - \varepsilon$.

Why $\frac{3}{16}$, and why the method stalls there. The available counts are of *simple* (and nearly-simple) geodesics — those are what Mirzakhani's formula integrates — and the volume asymptotics control the averaged geometric side up to errors of relative size $\approx e^L/g$ (each handle-loss costs $1/g$, each unit of length costs e; the sinh bound). With $\widehat{h}$ supported in $[-L, L]$, $L = 4\log g$, affordable error $\approx g\,\mathrm{poly}(\log g)$: an eigenvalue $\frac{1}{4} - \beta^2$ is detected only if

$$e^{\beta L} = g^{4\beta} \;\gg\; g \qquad\Longleftrightarrow\qquad \beta > \tfrac{1}{4} \qquad\Longleftrightarrow\qquad \lambda_1 < \tfrac{1}{4} - \tfrac{1}{16} = \tfrac{3}{16}.$$

(Exercise 3.2 organises this bookkeeping into a general "error dial".) It is a coincidence with a reason that Selberg's arithmetic bound is also $\frac{3}{16}$: in both cases a "trivial" term of size x fights an error of size $\sqrt{x}$ — the exponent $\frac{3}{4} = 1 - \frac{1}{4}$ in Weil's bounds plays the role of our $g^{4\beta}$ vs g threshold.

3.6.3 Round 2: Friedman–Ramanujan functions and the optimal gap

To go beyond, one must count *all* closed geodesics (not just simple ones) with test functions of unbounded support, and prove the averaged counts have errors with the square-root gain — the exact transplant of Definition 3.11:

Definition 3.13 (Friedman–Ramanujan functions, Anantharaman–Monk [AM23, AM25])**.** A function $F(L)$ (an averaged geodesic count) is *Friedman–Ramanujan* if

$$F(L) = \underbrace{p(L)\, e^{L}}_{\text{main term}} + O\Big(e^{L/2}\,\mathrm{poly}(L)\Big) \qquad \text{for all } L,$$

with p polynomial.

Their two-part work proves this property for the relevant counts on WP-random surfaces — via an asymptotic expansion of the volume polynomials in powers of $1/g$, new coordinates on Teichmüller space adapted to a fixed closed geodesic, and cancellation phenomena replacing Friedman's tangle analysis. With it, every $\beta > 0$ becomes detectable (the barrier dissolves: no threshold survives when errors are $e^{L/2}$ for all L), and:

Theorem 3.14 (Anantharaman–Monk 2025 [AM25, Mon26])**.** *For every $\varepsilon > 0$,*

$$\mathbb{P}_g^{\mathrm{WP}}\Big(\lambda_1(X) \;\geq\; \tfrac{1}{4} - \varepsilon\Big) \xrightarrow[g\to\infty]{} 1.$$

Recently, Hide–Macera–Thomas [HMT25] give an independent proof via the polynomial method for strong convergence of random matrices (Chen–Garza-Vargas–Tropp–van Handel [CGVTvH24]), and obtain the quantitative strengthening $\lambda_1 \geq \frac{1}{4} - O(g^{-c})$ whp, for some $c > 0$.

Combined with Huber's ceiling (Corollary 2.20):

The spectral gap of a typical hyperbolic surface of large genus is $\frac{1}{4} + o(1)$.
Typical surfaces are optimal expanders. Q3 is answered.

The route to Theorem 3.14 is summarised in Table 2.

Bound (whp, $g \to \infty$)	Who	Method
$\lambda_1 > 0.002$	Mirzakhani 2013 [Mir13]	Cheeger constant
$\lambda_1 \geq \frac{3}{16} - \varepsilon$	Wu–Xue; Lipnowski–Wright [WX22, LW24]	trace method, simple geodesics
$\lambda_1 \geq \frac{2}{9} - \varepsilon$	Anantharaman–Monk [AM23]	trace method + cancellations
$\lambda_1 \geq \frac{1}{4} - \varepsilon$	Anantharaman–Monk [AM25]	Friedman–Ramanujan functions
$\lambda_1 \geq \frac{1}{4} - O(g^{-c})$	Hide–Macera–Thomas [HMT25]	strong convergence (polynomial method)

Table 2: Successive lower bounds on the spectral gap of a Weil–Petersson random surface.

3.7 Panorama and open problems

- **Existence vs. typicality.** Hide–Magee [HM23] (Buser's conjecture, random covers of a non-compact surface + compactification) shows the ceiling is attained by

some sequence; Theorem 3.14 shows near-attainment is *typical* in the WP model; Hide–Macera–Thomas [HMT25] reprove typicality with a polynomial rate via strong convergence. Neither existence nor typicality implies the other; the random-cover model has its own parallel literature (Magee–Naud–Puder reached $\frac{3}{16}$ there first).

- **How fast?** Theorem 3.14 gives $\frac{1}{4} - \varepsilon$ for each fixed ε, with no rate; [HMT25] sharpens this to $\frac{1}{4} - O(g^{-c})$ for *some* $c > 0$. The optimal c — indeed the true size of $\frac{1}{4} - \lambda_1$ — is open, and the asymmetry is worth noticing: whp we now know

$$\tfrac{1}{4} - O(g^{-c}) \;\le\; \lambda_1(X) \;\le\; \tfrac{1}{4} + \tfrac{C}{(\log g)^2},$$

  the upper bound being Corollary 2.20. The polynomially small error below $\frac{1}{4}$ is far smaller than the logarithmically small slack above it: after [HMT25] the binding constraint is the *ceiling*, not the floor.

- **Fixed genus.** What is $\max_{X \in \mathcal{M}_2} \lambda_1$? Conjecturally the Bolza surface (Example 1.19), with $\lambda_1 = 3.838887\ldots$ computed to many digits with certified numerics by Strohmaier–Uski [SU13]; rigorous upper bounds via linear-programming/spectral-bootstrap methods are now within 2×10^{-5} of this value. Note $3.84 \ggg \frac{1}{4}$: the ceiling is an asymptotic phenomenon, and fixed-genus optimisation is a different (open) world.

- **Selberg's conjecture** ($\lambda_1 \ge \frac{1}{4}$ for congruence surfaces) remains open at $\frac{975}{4096}$. Randomness says nothing about specific arithmetic surfaces; conversely arithmetic methods do not see typical ones. The two worlds inform, but do not resolve, each other.

The tour ends where research begins. The natural next read, written as lecture notes at exactly this level: Monk's 2026 survey [Mon26].

What to remember from Lecture 3

1. WP random surfaces: FN coordinates (ℓ_i, τ_i), Lebesgue measure, canonical by Wolpert; "whp" as $g \to \infty$.
2. Mirzakhani: integration formula (unfolding) + volume asymptotics = computable geodesic counts. Sample: $\mathbb{P}_g^{\mathrm{WP}}(\mathrm{sys} \le \varepsilon) \asymp \varepsilon^2$; no separating geodesics below $\sim 2 \log g$.
3. Mirzakhani 2013: whp $h \gtrsim 0.099$, so $\lambda_1 > 0.002$ — expanders, but Cheeger's square is lossy.
4. Benjamini–Schramm: locally $X \to \mathbb{H}$; spectral distribution $\to$ Plancherel of $\mathbb{H}$ (Monk); only $o(g)$ exceptional eigenvalues — but density statements can't kill sparse outliers.
5. Trace method: $\mathrm{Tr}\, A^{2k}$ / Selberg trace formula; small eigenvalue = excess of geodesics; errors e^L/g give the $\frac{3}{16}$ barrier; square-root errors (Friedman–Ramanujan) give the optimal $\frac{1}{4} - \varepsilon$ whp.
6. The dictionary: tree $\leftrightarrow \mathbb{H}$; $2\sqrt{d-1} \leftrightarrow \frac{1}{4}$; Alon–Boppana $\leftrightarrow$ Huber; LPS $\leftrightarrow$ Selberg; Friedman $\leftrightarrow$ Anantharaman–Monk.

Epilogue: the story in one page

A closed surface hums at frequencies $0 = \lambda_0 < \lambda_1 \leq \cdots$, and the first nonzero one — the spectral gap — measures how well-connected the surface is: it is simultaneously the optimal Poincaré constant, the mixing rate of heat, and (Cheeger–Buser) the strength of the weakest bottleneck.

Hyperbolic surfaces, the constant-curvature homes of all genus-≥ 2 topology, have a distinguished frequency: $\frac{1}{4}$, the bottom of the spectrum of their common universal cover $\mathbb{H}$ — a sharp Hardy constant, ultimately an expression of the exponential room $\operatorname{Area} B_r \sim \pi e^r$. The gap of a hyperbolic surface can be made as small as one likes, but only in one way: pinch a *separating* system of geodesics, each of which drags along an explicit collar whose crossing cost is linear in its waist. Below $\frac{1}{4}$ the number of eigenvalues is capped by topology at $2g-2$, sharply; and above, as the genus grows, nothing survives: $\lambda_1 \leq \frac{1}{4} + C/(\log g)^2$, because a big surface locally is $\mathbb{H}$.

Whether the ceiling $\frac{1}{4}$ is ever reached was Buser's conjecture; the deterministic world (arithmetic surfaces, Selberg's $\frac{3}{16}$ and his still-open $\frac{1}{4}$) besieged it for decades, and randomness took it: random covers reach $\frac{1}{4}$ (Hide–Magee), and — the main theorem of this course — with respect to the Weil–Petersson measure, whose geodesic statistics Mirzakhani made computable, a *typical* surface of large genus has $\lambda_1 \geq \frac{1}{4} - \varepsilon$. The proof is a trace method: a small eigenvalue would force a visible excess of closed geodesics; averaged over moduli space, geodesic counts show no such excess, *provided* one can prove error terms with a square-root gain — the Friedman–Ramanujan property, transplanted from Friedman's proof of Alon's conjecture for random regular graphs, whose story ours mirrors line by line.

One number, λ_1; one ceiling, $\frac{1}{4}$; and the typical surface touches it.

> **Optional exercises** on this lecture (Appendix E.3, self-study): **B3** separating geodesics are long — deriving Mirzakhani's bound; **B4*** the error-exponent dial.

A Geometric background

This appendix collects the geometric background that was used — with a proof where they are short, and a reference where they are not.

A.1 Curvature

For $ds^2 = e^{2u}(dx^2 + dy^2)$: $K = -e^{-2u}(u_{xx} + u_{yy})$, and its meaning is the ball-area expansion Area $B(p, r) = \pi r^2(1 - \frac{K}{12}r^2 + o(r^2))$. Catalogue of the three constant-curvature geometries in conformal coordinates:

Euclidean plane:	$\phi = 1,$	$K = 0;$
round sphere (radius 1, stereographic):	$\phi = \dfrac{4}{(1+x^2+y^2)^2},$	$K = +1;$
hyperbolic plane (half-plane model):	$\phi = \dfrac{1}{y^2},$	$K = -1;$
hyperbolic plane (disk model, on $\lvert z\rvert < 1$):	$\phi = \dfrac{4}{(1-x^2-y^2)^2},$	$K = -1.$

The disk model is isometric to the half-plane model via the Möbius map $z \mapsto \frac{z-i}{z+i}$; it is the better picture for symmetric objects like the Bolza octagon (Example 1.19). Each verification of K is a two-line computation with the formula; do one.

A.2 Geodesics and polar coordinates in $\mathbb{H}$

Vertical lines are geodesics, with $d(iy_1, iy_2) = \log(y_2/y_1)$: proved in (10). Möbius maps (11) are isometries acting transitively on unit tangent directions; they preserve the family {vertical lines, circles orthogonal to $\mathbb{R}$}, and every unit direction at every point is tangent to exactly one member: hence that family is exactly the set of geodesics. Geodesic polar coordinates: around any p, the metric is $dr^2 + G(r)\,d\theta^2$; the function $\sqrt{G}$ satisfies the Jacobi equation $(\sqrt{G})'' = -K\sqrt{G} = \sqrt{G}$ with $\sqrt{G}(0) = 0$, $(\sqrt{G})'(0) = 1$, so $\sqrt{G} = \sinh r$:

$$ds^2 = dr^2 + \sinh^2 r\, d\theta^2, \qquad \text{Area } B(p, r) = 2\pi(\cosh r - 1) = 4\pi \sinh^2(r/2).$$

(Compare: flat $\sqrt{G} = r$; sphere $\sqrt{G} = \sin r$. The three geometries in one ODE.)

A.3 Classification of Möbius isometries

An element $T = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PSL}_2(\mathbb{R})$, $T \neq \mathrm{id}$, is one of:

- **elliptic** ($|\mathrm{tr}\,| < 2$): one fixed point in $\mathbb{H}$; a rotation around it;
- **parabolic** ($|\mathrm{tr}\,| = 2$): one fixed point on $\partial\mathbb{H}$; conjugate to $z \mapsto z + 1$;
- **hyperbolic** ($|\mathrm{tr}\,| > 2$): two fixed points on $\partial\mathbb{H}$; conjugate to $z \mapsto e^{\ell} z$ with $\cosh(\ell/2) = |\mathrm{tr}\,|/2$; translates the geodesic joining the fixed points (its *axis*) by distance ℓ.

If Γ is discrete, torsion-free and cocompact, every $T \in \Gamma \setminus \{1\}$ is hyperbolic: elliptics are excluded by freeness (fixed points), parabolics by compactness (they would create a cusp escaping to infinity). This yields Theorem 1.20(iii): conjugacy classes $\leftrightarrow$ closed geodesics, $\ell(T) = 2\,\mathrm{arccosh}(|\mathrm{tr}\,T|/2)$.

A.4 Why "curvature -1" = "quotient of $\mathbb{H}$"

A closed surface with $K \equiv -1$ has a complete, simply connected universal cover of curvature -1, which is isometric to $\mathbb{H}$ (Cartan–Hadamard plus the uniqueness of the constant-curvature simply connected surface — a Jacobi-equation argument as above, globalised). Deck transformations are isometries acting freely and properly discontinuously: a subgroup $\Gamma \subset \mathrm{PSL}_2(\mathbb{R})$ as in Definition 1.17. Conversely, quotients inherit $K = -1$.

A.5 Spectral theory used

The Laplacian of a compact surface is defined via its quadratic form (1) (Friedrichs extension); discreteness of the spectrum is Rellich compactness of $H^1 \hookrightarrow L^2$ plus the spectral theorem for compact self-adjoint operators (Theorem 1.3); min–max is Theorem 1.4. Elliptic regularity makes eigenfunctions smooth (real-analytic, in fact, for our real-analytic metrics — used once, in Cheeger's proof, for the regularity of nodal sets, alongside Sard's theorem). The reference at exactly the right level is Levitin–Mangoubi–Polterovich [LMP23].

A.6 Hyperbolic trigonometry

Right-angled hexagons: given alternate sides $a, b, c > 0$, there is a unique right-angled geodesic hexagon with these alternate side lengths; gluing two copies along the remaining three sides yields the pair of pants with prescribed boundary lengths (§3.1). The collar width $w(\ell) = \operatorname{arcsinh}(1/\sinh(\ell/2))$ and the simplicity of short geodesics (Theorem 2.9) come from the same hexagon identities. All of this is Chapters 1–4 of Buser [Bus92], readable independently of the rest of the book.

A.7 The heat kernel of $\mathbb{H}$

The heat kernel of the hyperbolic plane is explicitly known: $e^{-t\Delta}$ has kernel depending only on the distance $d = d(z, w)$,

$$p_t(d) = \frac{\sqrt{2}\, e^{-t/4}}{(4\pi t)^{3/2}} \int_d^\infty \frac{s\, e^{-s^2/4t}}{\sqrt{\cosh s - \cosh d}}\, \mathrm{d}s.$$

Two features to notice without verifying the formula. First, the prefactor $e^{-t/4}$: the spectral gap of $\mathbb{H}$ is visible in the long-time decay of heat — McKean's theorem in the language of heat flow. Second, kernels on quotients are obtained by summing over the group, $p_t^X(z, w) = \sum_{T \in \Gamma} p_t(d(z, Tw))$: this *automorphic summation* is the mechanism behind the trace formula. The heat kernel is also the standard route to the Weyl law (Theorem 1.9) and to Buser-type inequalities (Ledoux's proof).

B What we did not prove: sketches with references

For each result the notes quoted without proof: a paragraph on the idea, and a reference for the details.

Isothermal coordinates (every metric is locally conformal). Local coordinates making a surface metric $\phi(dx^2 + dy^2)$ exist; finding them amounts to solving a Beltrami equation $\partial_{\bar{z}} w = \mu \partial_z w$. So nothing is lost by using only conformal metrics, as §1.5 does throughout. *See* Chern (*Proc. AMS* 6, 1955); do Carmo sets it as an exercise.

Uniqueness of the constant-curvature simply connected surface. In geodesic polar coordinates, the metric is $dr^2 + G(r)d\theta^2$ with $(\sqrt{G})'' = -K\sqrt{G}$: curvature is an ODE along rays. $K \equiv -1$ forces $\sqrt{G} = \sinh r$, i.e. the metric of $\mathbb{H}$, around every point; completeness and simple connectivity globalise the identification (Cartan–Hadamard). *See* do Carmo, Ch. 8; [Bus92, Ch. 1].

The collar lemma. Lift to the cylinder cover C_ℓ (Example 1.18). If the w-tube around the waist fails to embed in X, two lifts of the tube meet, producing a geodesic arc returning to γ within distance $2w$; hyperbolic trigonometry (right-angled hexagons/pentagons) forces such an arc to be long when ℓ is short — quantitatively, embedding survives until $\sinh w \sinh(\ell/2) = 1$. Disjointness of collars: the same estimate applied to two geodesics. *See* [Bus92, Thm. 4.1.1, 4.1.6].

Buser's inequality $\lambda_1 \le 2h + 10h^2$. Given a near-optimal Cheeger cut, one cannot just interpolate linearly (the tube might be big); Buser's trick is to interpolate with the *distance function* to the cut and use the curvature lower bound through the Heintze–Karcher tube-volume comparison: with $K \ge -1$, $\mathrm{Area}\{p : d(p,\alpha) < \delta\} \le \ell(\alpha)\sinh(\delta) \cdot C$. Optimise δ. An alternative proof runs through the heat semigroup [Led94]. *See* [Bus82]; [Bus92, Ch. 8.1]; [Led94].

Schoen–Wolpert–Yau (the hard direction). Let u be an eigenfunction with λ small. Cover X by thick pieces and collars; on each thick piece, a local Poincaré inequality (with uniform constant, by compactness of the space of thick pieces) forces u to be nearly constant; the total energy budget λ then localises the variation of u into the collars, and the collar computation (17) says each crossing costs $\asymp \ell$. Organising the near-constant values into a graph and cutting along the cheapest collars produces a separating system of total length $\lesssim \lambda$. *See* [SWY80]; refined in [DPRS87].

Otal–Rosas ($\lambda_{2g-2} > \frac{1}{4}$). Consider the eigenfunctions of the first $2g-1$ eigenvalues; a topological argument on nodal sets shows some nontrivial linear combination would have to have a nodal domain on which the eigenvalue equation with value $\le \frac{1}{4}$ contradicts a comparison with $\mathrm{spec}(\mathbb{H})$ on incompressible subsurfaces (each such domain carries $\ge$ curvature-forced spectral mass; only $2g-2$ "independent" incompressible pieces fit in genus g). *See* [OR09]; see also Ballmann–Matthiesen–Mondal for generalisations.

Cheng's comparison, sharp form. Cheng proves $\lambda_1(X) \le \lambda_1^{\mathrm{Dir}}(B_{\mathbb{H}}(d/2))$ for any closed X with $K \ge -1$ and diameter d, by a clean comparison of first Dirichlet eigenvalues of balls (Barta's principle: transplant the comparison ball's ground state). The hyperbolic Dirichlet ball eigenvalue is $\frac{1}{4} + O(1/R^2)$ (Lemma 2.18 gives $\le$; a matching $\ge$ holds). *See* [Che75].

The Selberg trace formula. For a nice test function, the operator $h(\sqrt{\Delta - \frac{1}{4}})$ on $X = \Gamma\backslash\mathbb{H}$ has a kernel obtained by summing the corresponding $\mathbb{H}$-kernel $k(z,w)$ over Γ: $K_X(z,w) = \sum_{T\in\Gamma} k(z,Tw)$. Compute the trace $\int_X K_X(z,z)$ two ways: spectrally ($\sum h(r_j)$); and by grouping Γ into conjugacy classes — the identity gives the Area term (this is exactly the $\mathbb{H}$-Plancherel formula), each hyperbolic class unfolds to an integral over the cylinder of its axis, computable in coordinates (15) and producing the weight $\frac{\ell_{\gamma_0}}{2\sinh(k\ell/2)}\widehat{h}(k\ell)$. *See* [Bus92, Ch. 9] (compact case, complete and readable); Sarnak [Sar03] for the noncompact/arithmetic case.

Selberg's $\frac{3}{16}$. For congruence surfaces, eigenfunctions have Fourier expansions whose coefficients are Kloosterman sums; Weil's Riemann-hypothesis-for-curves bound $|S(m,n;c)| \le d(c)\sqrt{c}\,(m,n,c)^{1/2}$ gives square-root cancellation, which through the Petersson/Kuznetsov machinery keeps exceptional eigenvalues above $\frac{3}{16}$. The parallel with our story — square-root cancellation buying spectral gap — is exact and historically prior. *See* [Sel56]; Sarnak [Sar03], §2.

Mirzakhani's recursion and volume asymptotics. McShane's identity expresses the constant function 1 on $\mathcal{M}_{1,1}$ as a sum over simple geodesics of explicit functions of their lengths; Mirzakhani generalised it to all (g,n) and integrated it with her unfolding formula, obtaining a recursion computing all $V_{g,n}(b)$ (equivalently: intersection numbers of tautological classes; this reproves Witten–Kontsevich). The large-g asymptotics (Theorem 3.4(ii)) come from analysing the recursion's dominant terms; the sinh bound from term-by-term positivity. *See* [Mir07a, Mir07b, Mir13]; Wright's survey [Wri20] for the panorama; Mirzakhani–Zograf for the sharp V_g asymptotics.

Friedman, and Anantharaman–Monk. Friedman: expected counts of closed non-backtracking walks in random d-regular graphs are Ramanujan (Definition 3.11) — proved by a delicate classification of the ways a walk can traverse its "visited subgraph", with the tangles removed by a separate selective-trace argument. Anantharaman–Monk transplant the whole architecture: the average over $\mathcal{M}_g$ of geometric functions is expanded in powers of $1/g$ (each order a polynomial functional of the volume data); the Friedman–Ramanujan property of each order is proved via new length-coordinates adapted to a fixed closed geodesic, and surfaces with atypical short-geodesic patterns (the tangle analogues) are excised at small probability cost. *See* [Fri08] (or Bordenave's shorter proof [Bor20]); [AM23, AM25]; and above all the survey [Mon26], whose §2–3 are a guided tour of exactly this architecture.

Linear isoperimetry in $\mathbb{H}$ (Remark 2.3). The whole proof is the divergence theorem, once one has a vector field X with $|X| \le 1$ and $\operatorname{div} X \ge 1$: then $\operatorname{Area}(\Omega) \le \int_\Omega \operatorname{div} X = \oint_{\partial\Omega}\langle X,\nu\rangle \le \oint_{\partial\Omega}|X| = \ell(\partial\Omega)$. Two choices of X work. Radial: $X = \nabla r$ for $r = d(p,\cdot)$ has $|X| = 1$ and, by the radial Laplacian (12), $\operatorname{div} X = \coth r > 1$ — the same $\coth r > 1$ that powers the ground-state proof of McKean's bound. Horocyclic: $X = \nabla(-\log y)$ in the half-plane has $|X| = 1$ and $\operatorname{div} X = 1$ exactly, and is smooth everywhere (the radial field needs the singularity at p excised, which only helps the sign). The contrast with $\mathbb{R}^2$, where $\operatorname{div}\nabla r = 1/r \to 0$, is exactly the exponential-growth phenomenon (13): no

linear isoperimetric inequality holds in the plane. The same computation in $\mathbb{H}^n$ with $\operatorname{div}\nabla r = (n-1)\coth r$ gives Yau's $(n-1)\operatorname{vol}(\Omega) \le \operatorname{vol}(\partial\Omega)$. *See* [Yau75]; the sharp form $\ell^2 \ge 4\pi\,\mathrm{Area} + \mathrm{Area}^2$ needs symmetrisation instead.

C Annotated reading guide

A word on the references. This list is not complete, and favours what can be read straight after the course: the reference beside a theorem is the version most useful to read, not always the paper it first appeared in.

Textbooks (the two spines). Buser [Bus92] is *the* standard reference for Lectures 1–2 — **the geometric spine of the course**: Ch. 1–3 (hyperbolic building blocks, trigonometry, pants), Ch. 4 (collars), Ch. 5 (Bers' constant), Ch. 8 (small eigenvalues, Cheeger/Buser), Ch. 9 (trace formula). Levitin–Mangoubi–Polterovich [LMP23] is the modern first course on the variational toolkit (min–max, Weyl law, nodal geometry, Cheeger's inequality) — the right companion for students who have never seen spectral geometry.

Lecture 1. McKean [McK70] (three pages; the $\frac{1}{4}$). Sarnak [Sar03] for a panorama of the spectral theory of hyperbolic surfaces (arithmetic flavour).

Lecture 2. Most of this lecture is in Buser's book: [Bus92, Ch. 4] for the collar theorem with full proofs, [Bus92, Ch. 8] for small eigenvalues and the Cheeger–Buser circle, including his own $2g-2$ construction. Read those first; the papers below are what the book does *not* contain. Buser [Bus82] (the inequality itself), [Bus79] (the counterexample showing curvature is needed); Schoen–Wolpert–Yau [SWY80] (the source of Theorem 2.14, $\lambda_i \asymp_g L_i$ — the hard direction is the one the notes only sketch); Dodziuk–Randol [DR86] (a lower bound for λ_1 by a direct collar argument — the earlier and more transparent version of the method later pushed further in [DPRS87]); Otal–Rosas [OR09] (post-dates the book: the sharp $\lambda_{2g-2} > \frac{1}{4}$); Ballmann–Gromov–Schroeder [BGS85] (Margulis lemma and thick–thin in nonpositive curvature).

Lecture 3. Master reference, written as lecture notes at exactly this course's level: **Monk's 2026 survey [Mon26]** — read it immediately after the course. Monk–Naud [MN26] give the history of the whole subject, from Selberg to the 2025 proofs. For a shorter route through the same landscape, see **Petri's Séminaire Bourbaki exposé [Pet26]** (*Bass notes of random hyperbolic surfaces of large genus*), which takes the story up to Hide–Magee, Anantharaman–Monk and Hide–Macera–Thomas. Then: Mirzakhani [Mir07a] (integration formula; §§1–2 and 7), [Mir07b] (volumes and intersection theory), [Mir13] (random surfaces — read §1 and §4 with Exercise 3.1 in hand), [Mir10] (her own ICM exposition — the gentlest entry); Wright [Wri20] (tour of Mirzakhani's work, for the broader landscape); Monk [Mon22] and her thesis [Mon21] (Benjamini–Schramm, spectral distribution — the thesis's introduction is a survey in its own right); Anantharaman–Monk [AM25] (Friedman–Ramanujan functions; the proof of Theorem 3.14). For the

results quoted along the way: Wu–Xue [WX22] and Lipnowski–Wright [LW24] ($\frac{3}{16}$), Hide–Magee [HM23] and Hide–Macera–Thomas [HMT25], Friedman [Fri08] on the graph side, Strohmaier–Uski [SU13] for the Bolza numerics; Alon–Boppana is proved in Nilli [Nil91] ("A. Nilli" is a pseudonym of Noga Alon's); the Hoory–Linial–Wigderson survey [HLW06] is the place to read the graph side as a whole.

The graph side is moving too: Chen–Garza-Vargas–Tropp–van Handel [CGVTvH24] give a short proof of Friedman's theorem via strong convergence of random matrices — the method Hide–Macera–Thomas [HMT25] then transplanted to surfaces. Laura Monk (at Bristol) is leading a Lean formalisation of that proof, and it welcomes contributors (`https://lauramonk.github.io/FriedmanStrongConv/`).

D Notation glossary

Lecture 1: the hyperbolic plane and its spectrum

$\mathbb{H}$	hyperbolic plane $\{x+iy : y>0\}$, $ds^2=(dx^2+dy^2)/y^2$	(Def. 1.13)
$\mathrm{PSL}_2(\mathbb{R})$	Möbius isometry group of $\mathbb{H}$	(§1.5)
$X=\Gamma\backslash\mathbb{H}$	compact hyperbolic surface, genus $g\geq 2$	(Def. 1.17)
χ, $\operatorname{Area} X$	Euler characteristic $2-2g$; area $=-2\pi\chi=4\pi(g-1)$ (Gauss–Bonnet)	(§1.7)
Δ, λ_j	(positive) Laplacian; eigenvalues $0=\lambda_0<\lambda_1\leq\cdots$	(Thm. 1.3)
λ_1	the **spectral gap**	
$N(\lambda)$	eigenvalue counting function $\#\{j:\lambda_j\leq\lambda\}$	(Thm. 1.9)
$\mathcal{R}(f)$	Rayleigh quotient $\int\lvert\nabla f\rvert^2\,\mathrm{d}A/\int f^2\,\mathrm{d}A$	(Thm. 1.4)
$\lambda=s(1-s)$, $s=\frac{1}{2}+ir$	spectral parametrisation; $r\in\mathbb{R}$ tempered, $r\in i(0,\frac{1}{2}]$ exceptional	(Rem. 1.16)

Lecture 2: geometry controls the spectrum

$h(X)$	Cheeger constant	(Def. 2.1)
ℓ_γ, $\operatorname{sys}(X)$	length of the closed geodesic γ; shortest geodesic length	
$\ell_\gamma(X)$	the same length when X varies and γ is a fixed topological curve: a function on $\mathcal{M}_g$	(§3.2)
$\mathcal{C}(\gamma)$, $w(\ell)$	collar and its half-width $\operatorname{arcsinh}(1/\sinh(\ell/2))$	(Thm. 2.9)
$L_i(X)$	minimal length of a geodesic system cutting X into $i+1$ pieces	(Thm. 2.14)
$\operatorname{inj}(p)$	injectivity radius: largest r with $B(p,r)$ an embedded $\mathbb{H}$-ball	

Lecture 3: random surfaces

$\mathcal{T}_g$, $\mathcal{M}_g$, MCG_g	Teichmüller space, moduli space, mapping class group	(§3.1)
(ℓ_i,τ_i)	Fenchel–Nielsen coordinates (lengths, twists)	
ω_{WP}, V_g, $\mathbb{P}_g^{\mathrm{WP}}$	Weil–Petersson form, total volume, probability measure	
$V_{g,n}(b_1,\dots,b_n)$	WP volume of moduli of genus-g surfaces, n boundaries of lengths b_i	
whp	with $\mathbb{P}_g^{\mathrm{WP}}$-probability $\to 1$ as $g\to\infty$	
$\mathrm{Orb}(\gamma)$	the MCG_g-orbit of a simple closed curve: its *topological type*	(§3.2)
f, F	weight applied to a single length; the resulting *geometric function* $F(X)=\sum_{\gamma'\in\mathrm{Orb}(\gamma)} f(\ell_{\gamma'})$	(§3.2)

Lecture 3: graphs and the trace method

T_d, A, μ_+	d-regular tree; adjacency matrix; $\max(\mu_2,\lvert\mu_n\rvert)$	(§3.5)
C_k	Catalan number $\frac{1}{k+1}\binom{2k}{k}$, counting Dyck paths	(proof of Thm. 3.10)
h, $\widehat{h}$	test function in the Selberg trace formula and its Fourier transform	(§3.6)
β	defect of an exceptional eigenvalue: $\lambda_1=\frac{1}{4}-\beta^2$	(§3.6)

E Exercise sheets

Three sheets: two one-hour group sessions and a self-study sheet on the Lecture-3 material. Starred parts are optional challenges. These problems *extend* the lectures: you may use every result proved in the notes.

E.1 Exercise sheet 1

Exercise 1.1 (Flat tori: general lattices, and hearing the shape). Let $\Lambda \subset \mathbb{R}^2$ be a lattice ($\Lambda = \mathbb{Z}v_1 \oplus \mathbb{Z}v_2$ with v_1, v_2 linearly independent) and $T_\Lambda = \mathbb{R}^2/\Lambda$ with the Euclidean metric. Recall the *dual lattice* $\Lambda^* = \{w \in \mathbb{R}^2 : \ \langle w, v\rangle \in \mathbb{Z} \ \ \forall v \in \Lambda\}$.

(a) Show that the functions $e^{2\pi i\langle w,x\rangle}$, $w \in \Lambda^*$, are well defined on T_Λ, form a complete system of eigenfunctions, and conclude

$$\operatorname{spec}(T_\Lambda) = \Big\{4\pi^2|w|^2 : \ w \in \Lambda^*\Big\}, \qquad \lambda_1(T_\Lambda) = 4\pi^2 \min_{w\in\Lambda^*\setminus\{0\}} |w|^2.$$

(The lecture treated only rectangular lattices; the point here is the appearance of the *dual*.)

(b) (*You can hear a rectangular torus.*) Suppose $T_{a,b}$ and $T_{a',b'}$ (rectangular, $a \geq b$, $a' \geq b'$) have the same spectrum, with multiplicities. Show $a = a'$ and $b = b'$.

(c) (*Diameter and gap.*) Show $\operatorname{diam}(T_{a,b}) \geq a/2$. For $ab = 1$ and $a \to \infty$, check that the collapse of the gap is consistent with the Cheng-type bound $\lambda_1 \leq C/\operatorname{diam}^2$, and identify the bottleneck of the long thin torus.

(d) Compute the Neumann spectrum of the flat cylinder $C_{a,b} = [0, a] \times \mathbb{R}/b\mathbb{Z}$ and its gap. For $a = b = 1$, compare with the square torus: which has the smaller spectral gap, and by what factor? Explain geometrically what cutting the torus open did to the fundamental mode.

(d*) Can you hear a *general* flat torus of dimension 2? (True — reduce to a statement about positive-definite binary quadratic forms: show the spectrum determines the two successive minima of Λ^* and its covolume.)

Exercise 1.2 (Upper half-plane workout). Recall $\mathbb{H} = \{z = x + iy : y > 0\}$, $ds^2 = (dx^2 + dy^2)/y^2$, and the Möbius action of $\mathrm{PSL}_2(\mathbb{R})$.

(a) Verify the two identities used (but not checked) in the lecture: for $ad - bc = 1$,

$$\operatorname{Im} T(z) = \frac{\operatorname{Im} z}{|cz + d|^2}, \qquad |T'(z)| = \frac{1}{|cz + d|^2}.$$

(b) (*Transitivity, explicitly.*) For a target point $x_0 + iy_0 \in \mathbb{H}$, write down a matrix in $\mathrm{SL}_2(\mathbb{R})$ whose Möbius map sends $i \mapsto x_0 + iy_0$. Then show that the maps fixing i are exactly the $\left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix}\right)$, and identify their effect on directions at i.

(c) (*The distance formula.*) Prove that for $z, w \in \mathbb{H}$,

$$\cosh d(z,w) \;=\; 1 + \frac{|z-w|^2}{2\,\operatorname{Im} z\,\operatorname{Im} w}.$$

Hint: show the right-hand side is Möbius-invariant (use (a)); then reduce to two points on the imaginary axis, where $d = \log(y_2/y_1)$.

(d) Deduce that the hyperbolic circle of centre iy_0 and radius r is the *Euclidean* circle with centre $iy_0 \cosh r$ and radius $y_0 \sinh r$. Draw the picture: hyperbolic circles are Euclidean circles whose centres have drifted upward. What does this say about how $\mathbb{H}$ fools our Euclidean eyes?

Exercise 1.3 (Hardy's inequality with general exponents, and the rigidity of $\frac{1}{4}$). The lecture proved $\int_0^\infty u'^2\,\mathrm{d}y \geq \frac{1}{4}\int_0^\infty u^2 y^{-2}\,\mathrm{d}y$ and deduced $\operatorname{spec}(\mathbb{H}) \subset [\frac{1}{4}, \infty)$. Here you probe how rigid that $\frac{1}{4}$ is.

(a) (*General weights.*) For $\alpha < 1$ and $u \in C_c^\infty(0,\infty)$, prove

$$\int_0^\infty u'(y)^2\, y^\alpha\,\mathrm{d}y \;\geq\; \Big(\frac{1-\alpha}{2}\Big)^2 \int_0^\infty u(y)^2\, y^{\alpha-2}\,\mathrm{d}y,$$

by the same integration-by-parts scheme as in the lecture. Where does the argument break at $\alpha = 1$?

(b) (*A Dirichlet wall cannot lift the bottom.*) Let $U = \{y > 1\} \subset \mathbb{H}$, with the Dirichlet Laplacian (functions vanishing on $\{y = 1\}$). Show

$$\inf \operatorname{spec}\Big(\Delta, L^2(U)\Big) = \tfrac{1}{4}.$$

The lower bound is free (why?); for the upper bound, push test functions towards $y \to \infty$.

(c*) (*Sharpness on $\mathbb{H}$.*) Show the constant $\frac{1}{4}$ in McKean's theorem is attained in the limit: with $u_n(y) = y^{1/2}\varphi\big(\frac{\log y}{n}\big)$ and $f(x,y) = u_n(y)\,v(x/m)$, $\varphi, v \in C_c^\infty$ fixed, prove $\mathcal{R}(f) \to \frac{1}{4}$ when $n \to \infty$ and $m = m(n) \to \infty$ fast enough. Where have you seen the exponent $\frac{1}{2}$ before? (Answer: $\Delta y^s = s(1-s)y^s$.)

E.2 Exercise sheet 2

Part (c) of Exercise 2.3 is best revisited after Lecture 3.

Exercise 2.1 (Cheeger's inequality, transplanted to graphs). Let G be a connected d-regular graph on n vertices, A its adjacency matrix, $L = dI - A$ its Laplacian, with eigenvalues $0 = \nu_0 < \nu_1 \leq \cdots$. For $S \subset V$ let $e(S, S^c)$ be the number of edges between S and its complement, and define

$$h_G = \min_{0<|S|\leq n/2} \frac{e(S,S^c)}{|S|}.$$

The notes prove Cheeger's inequality for surfaces; here you transplant the *proof* to graphs — it is not written in the notes.

(a) Compute h_G for the cycle C_n and for the complete graph K_{d+1}.

(b) Prove the discrete Cheeger inequality

$$\nu_1 \;\geq\; \frac{h_G^2}{2d},$$

following the surface proof step by step: restrict a ν_1-eigenfunction u to its positivity set (arranged to have $\leq n/2$ vertices); prove the discrete coarea identity $\sum_{xy\in E} |v_x^2 - v_y^2| = \int_0^\infty e(S_t, S_t^c)\,\mathrm{d}t$ for $v = u^+$ and $S_t = \{v^2 > t\}$; then Cauchy–Schwarz. *You may use:* $\sum_{xy\in E}(v_x - v_y)^2 \leq \nu_1 \sum_x v_x^2$ *for* $v = u^+$ *(prove it if you have time: drop the helpful cross terms).*

(c) Sanity-check on C_n: compare $\nu_1 = 2 - 2\cos(2\pi/n)$ with $h_G^2/2d$. In the transplant, what replaced the factor 4 of the surface inequality by $2d$?

(d*) Prove the easy Buser-type converse $\nu_1 \leq 2h_G$ by testing with $u = \mathbf{1}_S - \frac{|S|}{n}$.

Exercise 2.2 (Inside the collar: what the notes left open).

Facts: the cylinder $C_\ell = \langle z \mapsto e^\ell z\rangle\backslash\mathbb{H}$; the distance of $z = x + iy$ to the imaginary axis satisfies $\sinh d = |x|/y$; the collar metric $ds^2 = d\rho^2 + \ell^2 \cosh^2\rho\, dt^2$ and the optimal crossing energy $4\ell/\pi$ are proved in the notes.

(a) (*Verify the collar metric.*) Writing $z = re^{i\varphi}$, set $t = \frac{\log r}{\ell}$ and define ρ by $\sinh\rho = \cot\varphi$ (signed). Show that in the coordinates (ρ, t) the hyperbolic metric becomes exactly $d\rho^2 + \ell^2\cosh^2\rho\, dt^2$ — the formula the notes quote without computation.

(b) (*The linear profile really is bad.*) Compute the energy of the naïve linear profile $f(\rho) = \rho/w$ across the collar and show

$$E_{\mathrm{lin}} = \frac{2\ell\sinh w}{w^2} \;\xrightarrow[\ell\to 0]{}\; \frac{4}{\big(\log(4/\ell)\big)^2}\,(1 + o(1)),$$

so $E_{\mathrm{lin}}/E_{\mathrm{opt}} \to \infty$: optimising the profile gains an order of magnitude. (The notes assert this; you are proving it.)

(c) (*Two waists, genus 3.*) Let X be a genus-3 surface containing two disjoint simple closed geodesics γ_1, γ_2, each of length ℓ, whose *union* separates X into two pieces of area 4π each. Build a test function that is ± 1 off both collars and crosses each with the optimal profile, and deduce

$$\lambda_1(X) \;\leq\; \frac{\ell}{\pi(\pi - 1)}\,(1 + o(1)) \qquad (\ell \to 0).$$

Why does neither γ_1 nor γ_2 *alone* give any bound at all?

(d) (*The cylinder's spectrum does not depend on the waist.*) Prove that $\inf\mathrm{spec}(\Delta, L^2(C_\ell)) = \frac{1}{4}$ for every $\ell > 0$.

- *Lower bound.* In the coordinates of (a), show that discarding the t-gradient gives
$$\int_{C_\ell} |\nabla f|^2 \, \mathrm{d}A \;\geq\; \ell \int_{\mathbb{R}} f_\rho^2 \cosh\rho \, \mathrm{d}\rho \cdot \int_0^1 \mathrm{d}t,$$
and reduce to the weighted 1-D Hardy inequality $\int_{\mathbb{R}} u'^2 \cosh\rho \, \mathrm{d}\rho \geq \frac{1}{4} \int_{\mathbb{R}} u^2 \cosh\rho \, \mathrm{d}\rho$. Prove this by the substitution $u = v(\cosh\rho)^{-1/2}$ and expanding. *This is the ground-state substitution of the notes' Remark "Another route to $\frac{1}{4}$" (§1.6), which also gives the second proof of McKean's bound in 1.3 — only the weight has changed, from the radial one to* $\cosh\rho$.
- *Upper bound.* Test with $u_n(\rho) = e^{-\rho/2}\varphi(\rho/n)$, where $\varphi \in C_c^\infty(\mathbb{R})$ with $\varphi \equiv 1$ near 0, and show $\mathcal{R}(u_n) \to \frac{1}{4}$ as $n \to \infty$.

Conclude: however short the waist, the bottom of the spectrum does not move. Contrast with part (c).

Exercise 2.3 (Non-backtracking walks: the graph avatar of geodesics). On a d-regular graph G, a walk is *non-backtracking* (NB) if it never immediately reverses an edge. Let $(A_\ell)_{xy}$ count NB walks of length ℓ from x to y ($A_0 = I$, $A_1 = A$), and write $\mathrm{girth}(G)$ for the length of a shortest cycle in G.

(a) Prove the recursions
$$A^2 = A_2 + dI, \qquad A\,A_\ell = A_{\ell+1} + (d-1)A_{\ell-1} \quad (\ell \geq 2).$$
Hint: $(A\,A_\ell)_{xy}$ counts pairs (edge xz, NB walk $z \to y$ of length ℓ); the concatenation is NB unless that walk returns to x first. Match the exceptions with NB walks $x \to y$ of length $\ell - 1$ and count the z producing each. Why is $\ell = 1$ different?

(b) Show that *any* tree has no closed NB walk of positive length. Now lift to the universal cover of G, which is the *infinite* d-regular tree, and deduce that closed NB walks in G are exactly the walks that genuinely wrap around cycles, and that the shortest closed NB walk has length $= \mathrm{girth}(G)$.

(c) Justify, in a sentence each, two new rows for the graphs↔surfaces dictionary of Lecture 3: girth ↔ systole; closed NB walks ↔ closed geodesics. *(Revisit after Lecture 3.)*

(d*) (*Where $2\sqrt{d-1}$ really comes from.*) By (a), $A_\ell = q_\ell(A)$ for a one-variable polynomial q_ℓ; write μ for its variable, which in the application is an eigenvalue of A. Setting $\mu = 2\sqrt{d-1}\cos\theta$, verify (initial cases + recursion) the closed formula
$$q_\ell(\mu) \;=\; (d-1)^{\ell/2}\Big(2\cos\ell\theta \;+\; \frac{d-2}{d-1} \cdot \frac{\sin\big((\ell-1)\theta\big)}{\sin\theta}\Big),$$
and deduce $|q_\ell(\mu)| \leq (\ell+1)(d-1)^{\ell/2}$ for all $|\mu| \leq 2\sqrt{d-1}$. *Hint: A is symmetric, so it is enough to bound the scalar $q_\ell(\mu)$ on the spectrum. Solve the recursion through its characteristic equation $x^2 - \mu x + (d-1) = 0$. Moral: $2\sqrt{d-1}$ is where the characteristic roots leave the circle $|x| = \sqrt{d-1}$, so beyond it NB counts stop being of size $(d-1)^{\ell/2}$. That this interval is also the spectrum of the infinite d-regular tree is a classical fact but not needed for the exercise; it is what makes $2\sqrt{d-1}$ the Ramanujan benchmark of Lecture 3.*

E.3 Exercise sheet 3

These are *not* worked in the notes: B3 derives a theorem of Mirzakhani stated there without proof, and B4 generalises the exponent bookkeeping. *Facts you may quote freely are printed in boxes.*

Exercise 3.1 (Separating geodesics are long: deriving Mirzakhani's bound).

Facts you may quote. Integration formula, separating type $(h, g-h)$, $1 \le h \le g/2$: $\mathbb{E}_g^{\mathrm{WP}}\Big[\sum_{\gamma \text{ sep},(h,g-h)} f(\ell_\gamma)\Big] = \frac{c_h}{V_g}\int_0^\infty t f(t) V_{h,1}(t) V_{g-h,1}(t)\,\mathrm{d}t$ with $c_h \le \frac{1}{2}$.
Volume facts: $V_{1,1}(t) = \frac{t^2+4\pi^2}{24}$; $V_{h,1}(t) \le V_{h,1}\frac{\sinh(t/2)}{t/2} \le V_{h,1}\, e^{t/2}$; and the ratio decay $\frac{V_{h,1}V_{g-h,1}}{V_g} \le \frac{C_h}{g^{2h-1}}$ with $\sum_h C_h < \infty$ (from the volume asymptotics of the notes).
For (d*) only — a *deterministic* bound, valid for every closed hyperbolic surface of genus g, not proved in the notes: $\lambda_1(X) \ge c\,\frac{L_1(X)}{g^2}$ for a universal $c > 0$, where $L_1(X)$ is as in the notes' Theorem on Schoen–Wolpert–Yau. *(Wu–Xue,* Optimal lower bounds for first eigenvalues of Riemann surfaces for large genus*, Amer. J. Math.* ***144*** *(2022), 1087–1114.)*

(a) Show that the expected number of separating geodesics of type $(1, g-1)$ and length $\le L$ is $O\Big(L^2 e^{L/2}/g\Big)$. (Note how the polynomial $V_{1,1}(t)$ — the genus-one side! — is what keeps the growth at $e^{L/2}$ rather than e^L.)

(b) Sum over the types $2 \le h \le g/2$ (both factors now exponential) and show their total contribution is $O\Big(e^L/g^3\Big)$. Conclude, via Markov:

$$\mathbb{P}_g^{\mathrm{WP}}\Big(\exists \text{ separating geodesic of length} \le a \log g\Big) \longrightarrow 0 \qquad \text{for every } a < 2,$$

— the theorem the notes quote from Mirzakhani's paper.

(c) Contrast with the nonseparating count of the notes ($\mathbb{P}_g^{\mathrm{WP}}(\mathrm{sys} \le \varepsilon) \asymp \varepsilon^2$ at fixed scale ε). What structural feature of the volume polynomials makes separating curves so much rarer? *(One sentence: products of moduli volumes are tiny compared with the connected volume.)*

(d*) Feed (b) into the Wu–Xue bound $\lambda_1 \ge c\, L_1(X)/g^2$ quoted in the box: what lower bound on λ_1 does length alone give whp? Compare with Mirzakhani's 0.002 and explain the moral. *(Note L_1 here is the shortest* separating *system, which is what (b) controls.)*

Exercise 3.2* (The error-exponent dial). Model the trace-method strategy of Lecture 3 as follows. Test functions h_L have $\widehat{h_L} \ge 0$ supported in $[-L, L]$ and $h_L \ge 0$ on $\mathbb{R} \cup i[-\frac{1}{2}, \frac{1}{2}]$; the geodesic terms of the trace formula carry weights $\approx e^{-\ell/2}$; and suppose the *averaged* geodesic counts up to length L are known with multiplicative error $e^{cL}\mathrm{poly}(L)$ for some exponent $c \in [\frac{1}{2}, 1]$, for all L.

(a) Show that after the $e^{-\ell/2}$ weights, the error contribution to the averaged geometric side is $O\Big(e^{(c-\frac{1}{2})L}\mathrm{poly}(L)\Big)$, while an exceptional eigenvalue $\lambda_1 = \frac{1}{4} - \beta^2$ contributes $\gtrsim e^{\beta L}$ to the spectral side. Conclude by Markov, letting $L/\log g \to \infty$ suitably:

$$\text{whp} \qquad \lambda_1 \;\ge\; \frac{1}{4} - \Big(c - \frac{1}{2}\Big)^2.$$

(b) Read the dial at its endpoints: $c = 1$ (no cancellation — trivial counting) and $c = \frac{1}{2}$ (the Friedman–Ramanujan property). Interpret both answers.

(c) Reconcile with the $\frac{3}{16}$ regime of the notes: there the expansion is only valid for $L \le 4 \log g$, with an additional error term of size $e^{L/2} \cdot \big(e^{L/2}/g\big)\mathrm{poly}$. Show detection then requires $e^{\beta L} \gg g\,\mathrm{poly}$, and recover whp $\lambda_1 \ge \frac{3}{16} - \varepsilon$ at $L = 4 \log g$.

(d*) Make a table of the bound $\frac{1}{4} - (c - \frac{1}{2})^2$ for $c = \frac{1}{2}, \frac{9}{16}, \frac{5}{8}, \frac{2}{3}, \frac{3}{4}, 1$. Which value of c gives the intermediate result $\frac{2}{9} - \varepsilon$ of Anantharaman–Monk (2023)? *(Answer: $c = \frac{2}{3}$.)* The moral of the table: along the *trace-method* line — Wu–Xue and Lipnowski–Wright, then Anantharaman–Monk twice — the successive theorems are not separate ideas but *one* inequality read at better and better cancellation exponents, and progress meant pushing c down to $\frac{1}{2}$. (The independent 2025 proof of Hide–Macera–Thomas reaches $\frac{1}{4}$ by a different route, strong convergence of random matrices, which this dial does not model.)

References

[AM85] Noga Alon and Vitali D. Milman. λ_1, isoperimetric inequalities for graphs, and superconcentrators. *J. Combin. Theory Ser. B*, 38(1):73–88, 1985.

[AM23] Nalini Anantharaman and Laura Monk. Friedman–Ramanujan functions in random hyperbolic geometry and application to spectral gaps I. Preprint, `arXiv:2304.02678`, 2023.

[AM25] Nalini Anantharaman and Laura Monk. Friedman–Ramanujan functions in random hyperbolic geometry and application to spectral gaps II. Preprint, `arXiv:2502.12268`, 2025.

[BGS85] Werner Ballmann, Mikhael Gromov, and Viktor Schroeder. *Manifolds of Nonpositive Curvature*, volume 61 of *Progress in Mathematics*. Birkhäuser, Boston, 1985.

[Bor20] Charles Bordenave. A new proof of Friedman's second eigenvalue theorem and its extension to random lifts. *Ann. Sci. Éc. Norm. Supér. (4)*, 53(6):1393–1439, 2020.

[Bus79] Peter Buser. Ueber den ersten eigenwert des Laplace-Operators auf kompakten Flächen. *Comment. Math. Helv.*, 54(1):477–493, 1979.

[Bus82] Peter Buser. A note on the isoperimetric constant. *Annales Scientifiques de l'École Normale Supérieure*, 15:213–230, 1982.

[Bus92] Peter Buser. *Geometry and Spectra of Compact Riemann Surfaces*, volume 106 of *Progress in Mathematics*. Birkhäuser, 1992. Reprinted in 2010.

[CGVTvH24] Chi-Fang Chen, Jorge Garza-Vargas, Joel A. Tropp, and Ramon van Handel. A new approach to strong convergence. Preprint, `arXiv:2405.16026`, 2024.

[Che70] Jeff Cheeger. A lower bound for the smallest eigenvalue of the laplacian. In *Problems in Analysis*, pages 195–199. Princeton University Press, 1970.

[Che75] Shiu-Yuen Cheng. Eigenvalue comparison theorems and its geometric applications. *Mathematische Zeitschrift*, 143:289–297, 1975.

[CLG09] Denis X. Charles, Kristin E. Lauter, and Eyal Z. Goren. Cryptographic hash functions from expander graphs. *Journal of Cryptology*, 22(1):93–113, 2009.

[Dod84] Józef Dodziuk. Difference equations, isoperimetric inequality and transience of certain random walks. *Trans. Amer. Math. Soc.*, 284(2):787–794, 1984.

[DPRS87] Jozef Dodziuk, Thomas Pignataro, Burton Randol, and Dennis Sullivan. Estimating small eigenvalues of riemann surfaces. In *The Mathematical Heritage of Hermann Weyl*, volume 64 of *Contemporary Mathematics*, pages 93–121. American Mathematical Society, 1987.

[DR86] Jozef Dodziuk and Burton Randol. Lower bounds for λ_1 on a finite-volume hyperbolic manifold. *Journal of Differential Geometry*, 24(1):133–139, 1986.

[Fri08] Joel Friedman. *A Proof of Alon's Second Eigenvalue Conjecture and Related Problems*, volume 195 of *Memoirs of the American Mathematical Society*. American Mathematical Society, 2008.

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

[HM23] Will Hide and Michael Magee. Near optimal spectral gaps for hyperbolic surfaces. *Annals of Mathematics*, 198:791–824, 2023.

[HMT25] Will Hide, Davide Macera, and Joe Thomas. Spectral gap with polynomial rate for Weil–Petersson random surfaces. Preprint, `arXiv:2508.14874`, 2025.

[Hub74] Heinz Huber. Über den ersten eigenwert des laplace-operators auf kompakten riemannschen flächen. *Commentarii Mathematici Helvetici*, 49:251–259, 1974.

[Led94] Michel Ledoux. A simple analytic proof of an inequality by P. Buser. *Proc. Amer. Math. Soc.*, 121(3):951–959, 1994.

[LMP23] Michael Levitin, Dan Mangoubi, and Iosif Polterovich. *Topics in Spectral Geometry*, volume 237 of *Graduate Studies in Mathematics*. American Mathematical Society, 2023.

[LP17] David A. Levin and Yuval Peres. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, second edition, 2017. With contributions by Elizabeth L. Wilmer, With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson.

[LPS88] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8:261–277, 1988.

[LW24] Michael Lipnowski and Alex Wright. Towards optimal spectral gaps of random surfaces. *Annals of Probability*, 52:545–575, 2024.

[McK70] Henry P. McKean. An upper bound to the spectrum of δ on a manifold of negative curvature. *Journal of Differential Geometry*, 4:359–366, 1970.

[Mir07a] Maryam Mirzakhani. Simple geodesics and Weil–Petersson volumes of moduli spaces of bordered riemann surfaces. *Inventiones Mathematicae*, 167:179–222, 2007.

[Mir07b] Maryam Mirzakhani. Weil–Petersson volumes and intersection theory on the moduli space of curves. *Journal of the American Mathematical Society*, 20:1–23, 2007.

[Mir10] Maryam Mirzakhani. On Weil–Petersson volumes and geometry of random hyperbolic surfaces. In *Proceedings of the International Congress of Mathematicians, Hyderabad*, pages 1126–1145, 2010.

[Mir13] Maryam Mirzakhani. Growth of Weil–Petersson volumes and random hyperbolic surfaces of large genus. *Journal of Differential Geometry*, 94:267–300, 2013.

[MN26] Laura Monk and Frédéric Naud. Spectral gaps on large hyperbolic surfaces. Preprint, `arXiv:2601.13988`, 2026.

[Mon21] Laura Monk. *Geometry and Spectrum of Typical Hyperbolic Surfaces*. PhD thesis, Université de Strasbourg, 2021.

[Mon22] Laura Monk. Benjamini–Schramm convergence and spectra of random hyperbolic surfaces of high genus. *Analysis & PDE*, 15:727–752, 2022.

[Mon26] Laura Monk. Typical hyperbolic surfaces have an optimal spectral gap. Preprint, `arXiv:2601.15157`, 2026. Current Developments in Mathematics 2025.

[MP19] Maryam Mirzakhani and Bram Petri. Lengths of closed geodesics on random surfaces of large genus. *Commentarii Mathematici Helvetici*, 94:869–889, 2019.

[Nil91] A. Nilli. On the second eigenvalue of a graph. *Discrete Mathematics*, 91(2):207–210, 1991.

[OR09] Jean-Pierre Otal and Eulalio Rosas. Pour toute surface hyperbolique de genre g, $\lambda_{2g-2} > \frac{1}{4}$. *Duke Mathematical Journal*, 150:101–115, 2009.

[Pet26] Bram Petri. Bass notes of random hyperbolic surfaces of large genus. Preprint, `arXiv:2607.06331`, 2026. Séminaire Bourbaki, exp. no. 1255.

[Pin73] Mark S. Pinsker. On the complexity of a concentrator. In *Proceedings of the 7th International Teletraffic Congress (ITC 7), Stockholm*, pages 318/1–318/4, 1973.

[Sar03] Peter Sarnak. Spectra of hyperbolic surfaces. *Bulletin of the American Mathematical Society*, 40:441–478, 2003. A Baltimore lecture-notes version was cited in the original bibliography.

[Sel56] Atle Selberg. Harmonic analysis and discontinuous groups in weakly symmetric riemannian spaces with applications to dirichlet series. *Journal of the Indian Mathematical Society*, 20:47–87, 1956.

[SU13] Alexander Strohmaier and Ville Uski. An algorithm for the computation of eigenvalues, spectral zeta functions and zeta-determinants on hyperbolic surfaces. *Communications in Mathematical Physics*, 317:827–869, 2013.

[SWY80] Richard Schoen, Scott Wolpert, and Shing-Tung Yau. Geometric bounds on the low eigenvalues of a compact surface. In *Geometry of the Laplace Operator*, volume 36 of *Proceedings of Symposia in Pure Mathematics*, pages 279–285. American Mathematical Society, 1980.

[Wol82] Scott A. Wolpert. The Fenchel–Nielsen deformation. *Annals of Mathematics*, 115:501–528, 1982.

[Wri20] Alex Wright. A tour through mirzakhani's work on moduli spaces of riemann surfaces. *Bulletin of the American Mathematical Society*, 57:359–408, 2020.

[WX22] Yunhui Wu and Yuhao Xue. Random hyperbolic surfaces of large genus have first eigenvalues greater than $\frac{3}{16} - \varepsilon$. *Geometric and Functional Analysis*, 32:340–410, 2022.

[Yau75] Shing-Tung Yau. Isoperimetric constants and the first eigenvalue of a compact Riemannian manifold. *Ann. Sci. École Norm. Sup. (4)*, 8(4):487–507, 1975.