Zeta Functions for Anosov flows

Mark Pollicott

Abstract. Dynamical zeta functions, by analogy with their more famous counterparts in number theory, are a useful tool to study certain types of dynamical systems. An important application is to the geodesic flow on a negatively curved surface. For surfaces of constant negative curvature the properties of the Selberg zeta function have been well understood for over half a century. However, understanding the properties of the corresponding zeta function for the more general setting of surfaces of variable negative curvature benefits from this more dynamical viewpoint.

Mathematics Subject Classification (2010). Primary 37C30; Secondary 11M36.

Keywords. Dynamical Zeta Functions, Thermodynamical Formalism.

1. Introduction

The best known setting for zeta functions is undoubtably that of analytic number theory, and so perhaps this is a good starting place to motivate the study of zeta functions for Anosov flows. We therefore begin with the best known zeta function, namely the Riemann zeta function, which is the complex function defined by

\[ \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad s \in \mathbb{C}, \]

which converges for \( \text{Re}(s) > 1 \). The Riemann zeta function was actually studied in 1737 by Euler. Indeed, it was Euler who proved the equivalent presentation

\[ \zeta(s) = \prod_p (1 - p^{-s})^{-1} \]

in terms of what is now called an Euler product over the prime numbers \( p \). However, when in 1859 Riemann was elected a member of the Berlin Academy of Sciences he reported “On the number of primes less than a given magnitude” [65] in a departure from his previous, and subsequent, research. In particular, he established the following basic properties of this zeta function.

**Theorem 1.1 (Riemann).** The zeta function \( \zeta(s) \) converges to a non-zero analytic function for \( \text{Re}(s) > 1 \). Moreover,

1. \( \zeta(s) \) has a single (simple) pole at \( s = 1 \); and
2. \(\zeta(s)\) extends to all complex numbers \(s \in \mathbb{C}\).

In particular, \(\zeta(s)\) has an analytic extension to the entire complex plane, except for a simple pole at \(s = 1\).

One of the main applications of the Riemann zeta function was to prove the prime number theorem (shown independently by Hadamard and de la Vallee Poussin in 1896). Let \(\pi(x)\) denote the number of primes \(p\) which are less than \(x\), for \(x > 0\).

**Theorem 1.2** (Prime Number Theorem). \(\pi(x) \sim \frac{x}{\log x} \text{ as } x \to +\infty\).

The asymptotic formula in the theorem means that \(\lim_{x \to +\infty} \frac{\pi(x)}{x/\log x} = 1\). The proof depends on the additional knowledge that \(\zeta(s)\) has no zeros on the line \(\text{Re}(s) = 1\).

It remains a major problem (famously posed in Hilbert’s 8th problem from his list of 23 problems from the 1900 International Congress of Mathematicians) to find the optimal asymptotic formulae. This can be formulated in terms of the zeros of the zeta function in the following slightly nonstandard form.

**Conjecture 1.3** (Riemann Hypothesis). 1 The Riemann zeta function \(\zeta(s)\) is analytic and non-zero on the half-plane \(\text{Re}(s) > \frac{1}{2}\), except for a simple pole at \(s = 1\).

The consequences of the validity of this conjecture for the behaviour of \(\pi(x)\) are well known. In particular, the Riemann hypothesis would improve the Prime Number Theorem (Theorem 1.2) by giving a very strong error term, i.e., we would know that

\[
\pi(x) = \text{Li}(x) + O \left( x^{1/2} \log x \right)
\]

where \(\text{Li}(x) = \int_2^x \frac{1}{\log u} du \) (with, of course, \(\text{Li}(x) \sim \frac{x}{\log x} \text{ as } x \to +\infty\)).

We will describe the analogous zeta functions in both geometric and dynamical settings (the Selberg and Ruelle zeta functions, respectively). In each case, we will be interested in understanding how far they can be extended analytically or meromorphically (the analogue of Theorem 1.1 and Conjecture 1.3, respectively).

## 2. The zeta function for geodesics

It is very striking that many of the features of the prime numbers and the Riemann zeta function \(\zeta(s)\) have counterparts in the geometry of compact surfaces of constant curvature. Let \(V\) denote a compact surface with constant negative Gaussian curvature \(\kappa = -1\). Instead of prime numbers we can consider closed geodesics \(\gamma\), of which there are a countable infinity on \(V\) since there is exactly one in every

---

1 According to Littlewood [41] this topped Hardy’s famous wish list from the 1920s: (1) Prove the Riemann Hypothesis; (2) Make 211 not out in the fourth innings of the last test match at the Oval; (3) Find an argument for the nonexistence of God which shall convince the general public; (4) Be the first man at the top of Mount Everest; (5) Be proclaimed the first president of the U.S.S.R., Great Britain, and Germany; and (6) Murder Mussolini.
conjugacy class of the fundamental group \( \pi_1(V) \). We will adopt the convention that closed geodesics are oriented (i.e., we count the two orientations of the same curve as two distinct closed geodesics), Let \( l(\gamma) \) denote the length of the closed geodesic \( \gamma \), say, then we recall the original definition of the Selberg zeta function.

**Definition 2.1.** The *Selberg zeta function* is defined by

\[
Z(s) = \prod_{\gamma} \prod_{n=0}^{\infty} \left( 1 - e^{-(s+n)l(\gamma)} \right), \quad s \in \mathbb{C}.
\]

This converges to an analytic function for \( \text{Re}(s) > 1 \). However, for later convenience we prefer to consider an alternative version of this zeta function (due to Ruelle) of the form

\[
\zeta_S(s) = \prod_{\gamma} \left( 1 - e^{-sl(\gamma)} \right)^{-1}, \quad s \in \mathbb{C},
\]

which again converges to a non-zero analytic function for \( \text{Re}(s) > 1 \). This has a reassuringly similar form to that of the Riemann zeta function, by formally replacing the prime numbers \( p \) by the weights \( e^{l(\gamma)} \) associated to each closed geodesic \( \gamma \). Clearly, we can write \( \zeta_S(s) = Z(s+1)/Z(s) \). The analogue of Theorem 1.1 is then the following [73], [30].

**Theorem 2.2.** The zeta function \( \zeta_S(s) \) converges to a non-zero analytic function for \( \text{Re}(s) > 1 \). Moreover,

1. \( \zeta_S(s) \) has a simple pole at \( s = 1 \); and
2. \( \zeta_S(s) \) extends to all complex numbers \( s \in \mathbb{C} \).

In this case the extension of \( \zeta_S(s) \) to \( \mathbb{C} \) is as a meromorphic function.

Pursuing the analogy between prime numbers and closed geodesics, the similar properties of the two zeta functions leads to a result on counting closed geodesics corresponding to that of the Prime Number Theorem (Theorem 1.2). Let \( \Pi(x) \) denote the number of prime closed geodesics \( \gamma \) with length \( l(\gamma) \) less than \( x \).

**Theorem 2.3** (Prime Geodesic Theorem). \( \Pi(x) \sim \frac{x}{\log x} \) as \( x \to +\infty \).

This geometric zeta function \( \zeta_S(s) \) has some advantages over that of the classical Riemann zeta function \( \zeta(s) \). In particular, the poles and zeros of \( \zeta_S(s) \) can be explicitly characterized, and an analogue of the Riemann hypothesis holds. We will return to this later in \( \S 6 \).

However, if one takes the broader view of Riemannian geometry it is natural to ask if these results generalise to geodesics on surfaces with variable negative curvature (or, more generally, higher dimensional manifolds with negative sectional curvatures). In fact, at about the same time that the above ideas were taking root, Anosov developed a completely different dynamical framework which would ultimately help address these questions [1].
3. Anosov flows

In order to formulate a dynamical analogue of the previous zeta functions we want to replace the prime numbers in $\zeta(s)$ (or closed geodesics in $\zeta_S(s)$) by closed orbits for appropriate flows which, in particular, we require to have a countable infinity of prime closed orbits. A particularly important class of such flows is that of Anosov flows.

Definition 3.1. We say that a $C^\infty$ flow $\phi_t : M \to M$ on a compact manifold $M$ is Anosov if the following hold.

1. There is a $D\phi_t$-invariant splitting $TM = E^0 \oplus E^s \oplus E^u$ such that
   
   (a) $E^0$ is one dimensional and tangent to the flow direction;
   
   (b) There exist $C, \lambda > 0$ such that $\|D\phi_t|E^s\| \leq Ce^{-\lambda t}$ and $\|D\phi_{-t}|E^u\| \leq Ce^{-\lambda t}$ for $t > 0$.

2. The flow is transitive (i.e., there exists a dense orbit).

A crucial feature of Anosov flows is that they have a countable number of closed orbits $\tau$ whose least periods $\lambda(\tau)$ tend to infinity. Moreover, for our purposes an important fact is that they can be used to study geodesics on surfaces of variable negative curvature via the associated geodesic flow.

3.1. Geodesic flows. Let $V$ be a compact surface with (variable) negative curvature $\kappa(x) < 0$, for $x \in V$, and let $M = SV := \{v \in TV : \|v\| = 1\}$ be the unit tangent bundle. Let $\phi_t$ be the associated geodesic flow, i.e., $\phi_t(v) = \dot{\gamma}_v(t)$ where $\gamma_v : \mathbb{R} \to V$ is the unit speed geodesic with $\dot{\gamma}_v(0) = v$. The closed orbits for the geodesic flow then correspond to closed geodesics on $V$.

Remark 3.2. Geodesic flows on surfaces of negative curvature, their dynamical properties and their analysis via symbolic coding were studied in a fundamental paper by Hadamard [29], only two years after his proof of The Prime Number Theorem (Theorem 1.2). This work was popularized in a 1906 book by the French
physicist Duhen, and subsequently translated into German by Adler. In 1909, Adler’s family shared a house with Einstein and his translation may (or may not) have influenced Einstein’s work on general relativity. However, Adler is better known for assassinating the prime minister of Austria, Count Karl von Stürghk on 21st October, 1916 [77].

![Figure 2. The geodesic flow on a negatively curved compact surface. This is an Anosov flow, with the bundles $E^s$, $E^u$ being associated to the horocycles.](image)

More generally, a geodesic flow on a compact surface with some positive curvature may be an Anosov flow providing there is sufficient negative curvature, in an appropriate sense. We can illustrate this by two particularly simple examples of Anosov geodesic flows.

![Figure 3. Two examples of Anosov geodesic flows: (i) The linkage example of Mackay and Hunt; (ii) A practice golf ball which resembles the surface of Donnay and Pugh.](image)

**Example.** Consider an idealized linkage, by which we mean a mechanical system consisting of a series of rigid rods where each rod has two either fixed pivots or movable joints connecting their ends. Furthermore, we can assume that all of the mass is concentrated on the joints (and friction, inertia, gravity, etc. can be neglected) and consider the time evolution of this idealised mechanical linkage...
in its phase space \[33\]. Its behaviour is described by a geodesic flow on its two
dimensional configuration space. In particular, Hunt and MacKay constructed
examples of triple linkages (based on the topological examples of Thurston and
Weeks) for which the flow is Anosov. Other examples were investigated in \[46\].

**Example.** Donnay and Pugh showed how to construct a surface which can be
embedded into three dimensional Euclidean space, and for which the geodesic flow
is Anosov \[16\]. The construction begins with two concentric spheres (with mild
positive curvature) which are then connected by a large number of judiciously placed
small tubes (with strong negative curvature).

One of the basic properties of Anosov flows is structural stability, by which any
sufficiently small perturbation of the flow still results in an Anosov flow and thus
gives a wealth of related examples. However, we next recall a second basic class of
examples which are fundamentally different to geodesic flows.

### 3.2. Suspensions of Anosov diffeomorphisms

Given any homeomorphism \( f : X \to X \) of a compact metric space \( X \) and a strictly positive continuous function \( r : X \to \mathbb{R}^+ \), we can associate a new space

\[
\tilde{X} = \{(x,u) : 0 \leq u \leq r(x) \}/(x,r(x)) \sim (f(x),0).
\]

We can then define a flow \( f_t : \tilde{X} \to \tilde{X} \) by \( f_t(x,u) = (x,u+t) \), subject to the equivalence relation.

We now specialise to the special case that the homeomorphism is a\( C^\infty \) Anosov
diffeomorphism \( f : X \to X \) on a compact manifold. The general definition can be
found in \[8\]. However, the simplest example to have in mind is that of an Arnol’d
CAT map (standing for Continuous Automorphism on a Torus), e.g., \( f : \mathbb{R}^2/\mathbb{Z}^2 \to \mathbb{R}^2/\mathbb{Z}^2 \) defined by \( f(x,y) = (2x + y, x+y) \). In the case that \( f : X \to X \) is an
Anosov diffeomorphism and \( r : X \to \mathbb{R}^+ \) is \( C^\infty \) the flow \( f_t : \tilde{X} \to \tilde{X} \) is a \( C^\infty \)
Anosov flow.

![Figure 4. A suspension flow \( f_t : \tilde{X} \to \tilde{X} \) over a transformation \( f : X \to X \).](image-url)
We say that an Anosov flow \( \phi_t : M \to M \) is topologically weak mixing if there is no non-trivial solution to \( F \circ \phi_t = e^{iat}F \) with \( a \in \mathbb{R} \) and \( F \in C^0(M, \mathbb{C}) \). Whereas Anosov geodesic flows are always topologically weak mixing, the suspension Anosov flows may not always be so (for example, when \( r \) is a constant function). However, topologically weak mixing is a generic assumption and we will henceforth assume it to simplify the exposition.

4. The Ruelle zeta function for Anosov flows

It should now be fairly clear what the natural generalisation of the zeta function \( \zeta_S(s) \) to Anosov flows \( \phi_t : M \to M \) in general, and geodesic flows in particular, should be. We will denote by \( \tau \) a primitive closed orbit and let \( \lambda(\tau) > 0 \) denote its period (i.e., for any \( x_\tau \in \tau \) the period is the smallest value \( t > 0 \) for which \( \phi_t x_\tau = x_\tau \)). However, to specify the domain of convergence of the zeta function we still need to introduce the notion of the topological entropy \( h(\phi) \) of the flow. The definition in the general case can be found in [55], for example, but there is a particularly simple equivalent formulation in the specific context of geodesic flows (on negatively curved surfaces) which we now recall.

**Example (Topological entropy of geodesic flows).** Let \( V \) be a surface with negative curvature. We can lift the Riemannian metric on \( V \) to the universal cover \( \tilde{V} \) and consider the rate of growth of volume of balls \( B_{\tilde{V}}(x, R) \) of radius \( R > 0 \) in \( \tilde{V} \). Then the topological entropy for the geodesic flow \( \phi_t : M \to M \) is given by

\[
\lim_{T \to +\infty} \frac{1}{T} \log \text{Vol}(B_{\tilde{V}}(x, R))
\]

for any \( x \in \tilde{V} \) [44]. In the special case of a surface of constant curvature \( \kappa = -1 \) we easily see that the geodesic flow has topological entropy \( h(\phi) = 1 \).

By analogy with the product form of the Riemann zeta function \( \zeta(s) \) (and the geometric zeta function \( \zeta_S(s) \)) we can now define the corresponding zeta function for an Anosov flow as follows.

**Definition 4.1.** The Ruelle zeta function for an Anosov flow is defined by

\[
\zeta_R(s) = \prod_{\tau} \left(1 - e^{-s\lambda(\tau)}\right)^{-1}, \quad s \in \mathbb{C},
\]

which converges to a non-zero analytic function for \( \text{Re}(s) > h(\phi) \).

There was also a version proposed by Smale [76] closer to original definition \( Z(s) \) of Selberg, but we will use the formulation of Ruelle.

**Example.** In the case of the geodesic flow on a compact negatively curved surface the closed orbits \( \tau \) of period \( \lambda(\tau) \) correspond to closed geodesics \( \gamma \) of length \( l(\gamma) = \lambda(\tau) \). In particular, we then have that \( \zeta_S(s) = \zeta_R(s) \).
Riemann | Selberg | Ruelle
--- | --- | ---
Number Theory | Geometry | Dynamical Systems
Primes $p$ | Closed geodesics $\gamma$ | Closed orbits $\tau$
$p$ | $e^{l(\gamma)}$ for lengths $l(\gamma)$ | $e^{\Lambda(\tau)}$ for periods $\Lambda(\tau)$
$\zeta(s) = \prod_p (1 - p^{-s})^{-1}$ | $\zeta_S(s) = \prod_{\gamma} (1 - e^{-sl(\gamma)})^{-1}$ | $\zeta_R(s) = \prod_{\tau} (1 - e^{-s\lambda(\tau)})^{-1}$

Table 1. Comparing the definitions of three zeta functions: Riemann zeta function, the geometric zeta functions and the Ruelle zeta function.

The following theorem shows that Theorem 2.2 for constant curvature surfaces generalises to Anosov flows. In particular, it shows that the theorem generalises to surfaces of variable negative curvature. Let $\phi_t : M \to M$ be a $C^\infty$ Anosov flow.

**Theorem 4.2.** The zeta function $\zeta_R(s)$ converges to a non-zero analytic function for $\text{Re}(s) > h(\phi)$. Moreover,

1. $\zeta_R(s)$ has a simple pole at $s = h(\phi)$; and
2. $\zeta_R(s)$ extends to all complex numbers $s \in \mathbb{C}$ as a meromorphic function.

The first part of this theorem was proved by Ruelle in [68]. The second part was proved in [26], with an alternative proof being given in [17]. If we have only have finite regularity (i.e., a $C^k$ Anosov flow $\phi_t : M \to M$ with $1 \leq k < +\infty$) then we still get an extension, albeit to a half plane. In particular, there exists $\lambda > 0$ (from the definition of the Anosov flow) such that $\zeta_R(s)$ has an extension to $\text{Re}(s) > h(\phi) - \lambda[\frac{k}{2}]$ [26].

In the general case of topologically weak mixing Anosov flows (which, we recall, includes geodesic flows on negatively curved surfaces) the zeta function $\zeta_R(s)$ has no zeros or poles on the line $\text{Re}(s) = h(\phi)$, other than $s = h(\phi)$, by analogy with the corresponding property for the Riemann zeta function $\zeta(s)$. Given a topologically weak mixing Anosov flow, we reuse the notation $\Pi(x) = \{\tau : \lambda(\tau) \leq x\}$, this time to denote the number of closed orbits $\tau$ with least period $\lambda(\tau)$ less than $x$. The following theorem was originally due to Margulis [47], [48] although the proof using zeta functions appears in [54], [55].

**Theorem 4.3** (Prime Orbit Theorem). For a topologically weak-mixing Anosov flow we have that $\Pi(x) \sim \frac{e^{h(\phi)}}{h(\phi)x}$ as $x \to +\infty$.

Restricting to the case of geodesic flows on compact negatively curved surfaces, we have the following generalisation of Theorem 2.3.

**Corollary 4.4.** For compact surfaces of variable negative curvature we have that $\Pi(x) \sim \frac{e^{h(\phi)}}{h(\phi)x}$ as $x \to +\infty$.

In particular, this gives the extension of Theorem 2.3 to the case of variable negative curvature.

---

2 An amusing reminiscence appears in Ruelle’s article [67].
Remark 4.5 (Zeta functions for Anosov diffeomorphisms). We have omitted a
detailed discussion of the case of zeta functions for Anosov diffeomorphisms $f : X \to X$, despite their mathematical and historical importance. In this context,
the natural definition of the zeta function is that of Artin and Mazur [2]. Let $N(n), n \geq 1$, denote the number of fixed points $f^n x = x$ for $f^n : X \to X$ and define
\[
\zeta_{AM}(z) = \exp \left( \sum_{n=1}^{\infty} \frac{z^n}{n} N(n) \right), \quad z \in \mathbb{C},
\]
which converges for $|z|$ sufficiently small. This zeta function can also be written in
Euler product form as $\zeta(z) = \prod_\tau (1 - z^{|\tau|} )^{-1}$ with the product over prime closed
orbits $\tau = \{x, f x, \cdots, f^{n-1} x\}$ of least period $|\tau| = n$. For Anosov diffeomorphisms,
or even more generally for Axiom A diffeomorphisms, this is a rational function,
(i.e., a quotient of two polynomials in $z$) [44], [28], [20]. This is perhaps reminiscent
of the results for the Lefschetz zeta function, the Weil zeta functions for finite fields,
and the Ihara zeta functions for finite graphs [80], [6]. There are also closely related
results for interval maps [31].

More generally, one might weight $\tau$ by taking the values along the points in the
orbit of a suitable function $F : X \to \mathbb{C}$ and then associate a more general version
of the Artin-Mazur zeta function
\[
\zeta_{R,F}(z) = \exp \left( \sum_{n=1}^{\infty} \frac{z^n}{n} \sum_{T^n x = x} \exp \left( \sum_{i=0}^{n-1} F(T^i x) \right) \right).
\]
Of course, this reduces to the Artin-Mazur zeta function above when $F$ is identically zero. On the other hand, this gives the zeta function for the suspension
Anosov flow when $F = -sr$ (where $s \in \mathbb{C}$ and $r : X \to \mathbb{R}^+ )$. Much is known
about the domain of $\zeta_{R,F}(z)$ through the work of many authors [66], [68], [69],
[55], [3], [43].

5. Techniques for $C^\infty$ Anosov flows

We briefly recall the two principle approaches to studying zeta functions for Anosov
flows. Broadly speaking, these share the common strategy of using suitable operators
on appropriate Banach spaces. In both approaches, the extension of the zeta
function is related to the spectral properties of the corresponding operator and the
poles and zeros are related to spectra of the operators. The earlier approach used
symbolic dynamics and a reduction to the study of transfer operators on classical
Banach spaces of Hölder functions for subshifts of finite type [55]. The more recent
approaches avoid this somewhat non-canonical reduction and involve similar, but
technically different, operators on specially tailored Banach spaces of distributions
[26].
5.1. **Symbolic dynamics.** The classical approach to studying the Ruelle zeta function $\zeta_R(s)$ was based on the use of symbolic dynamics, whereby the Anosov system was modelled by the suspension of a two sided subshift of finite type $\sigma : X \to X$ by a Hölder continuous function $r : X \to \mathbb{R}$. This is similar in spirit to the suspension construction described in §3.2, although in the general case the subshift arises from considering the Poincaré map on a finite number of judiciously chosen codimension one transverse sections. We refer to [64] and [9] for more details.

In this approach one replaces the two sided subshift of finite type (which is a homeomorphism) by the one sided subshift of finite type (which is a finite-to-one local homeomorphism). This corresponds to artificially suppressing the effect of the contracting direction for the flow. One can then considers the Banach space of Hölder continuous functions on the one sided shift space and an associated transfer operators (parameterised by complex numbers $s \in \mathbb{C}$) which averages over the preimages [68], [55].

**Remark 5.1.** This approach has the advantage that it applies in the even more general setting of Smale’s Axiom A flows [76] and which, as we will see later in §7.3.1, has applications to geodesic flows on infinite volume surfaces. However, in this more general context of Axiom A flows the zeta function $\zeta_R(s)$ may not extend meromorphically to $\mathbb{C}$, and there are examples of Axiom A flows for which the zeta function has an essential singularity [25], [68].

5.2. **Anisotropic spaces.** Despite its early success, the previous approach has the distinct disadvantage that one cannot make use of the smoothness of the flow. A more recent approach has been to work with simpler operators, but to consider more sophisticated Banach spaces of distributions. The origins for these ideas lie in the papers of Rugh [72] and Fried [21] for real analytic Anosov diffeomorphisms and flows, respectively, and by Kitaev [40] for Anosov diffeomorphisms of finite differentiability. Two different approaches to defining these Banach spaces were developed for diffeomorphisms by Baladi-Tsujii [3] and Goëzel-Liverani [27]. The approach in the latter paper was then extended to flows in [11] and [12]. However, in order to apply this method to the Ruelle zeta function $\zeta_R(s)$ it is necessary to generalise it to Banach spaces of forms [26], as was already anticipated in the work of Ruelle [66]. In particular, the extension of the zeta function is via closely related complex functions called determinants which, as the name suggests, are extended using spectral properties of the associated transfer operators. The determinants associated to forms can then be combined to recover the zeta function.

6. **Error terms for counting closed orbits**

In the introduction we recalled the importance of the Riemann Hypothesis (Conjecture 1.3) for the Riemann zeta function. We now consider its analogues in the context of the zeta functions for closed geodesics on constant and variable negatively curved manifolds, respectively. We begin with the classical results in the case of the geometric zeta function $\zeta_S(s)$ for a compact surface $V$ of constant
curvature $\kappa = -1$.

### 6.1. The surface case.

We begin with the case that $V$ is a compact surface of constant negative curvature $\kappa = -1$. Let $\Delta : L^2(V) \to L^2(V)$ be the usual self-adjoint Laplace-Beltrami operator with eigenvalues $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots$ satisfying $\Delta \psi_n + \lambda_n \psi_n = 0$. In particular, we have the following partial analogue of the Riemann Hypothesis (Conjecture 1.3) [30].

**Theorem 6.1.** For a compact surface $V$ of constant curvature $\kappa = -1$ the zeta function $\zeta_S(s)$ has a non-zero analytic extension to a half plane, $\text{Re}(s) > 1 - \epsilon$ where

$$\epsilon = \begin{cases} \frac{1}{2} - \sqrt{\frac{1}{4} - \lambda_1} & \text{if } 0 < \lambda_1 < \frac{1}{4} \\ \frac{1}{2} & \text{if } \lambda_1 \geq \frac{1}{4} \end{cases}$$

except for a simple pole at $s = 1$.

Combining this with additional bounds on the modulus of the zeta function in this region gives an error term estimate for number $\Pi(x)$ of geodesics $\gamma$ with length $l(\gamma)$ less than $x$ of the following form.

**Corollary 6.2.** For $0 < \epsilon' < \epsilon$, $\Pi(x) = \text{Li}(e^x) + O\left(e^{(1-\epsilon')x}\right)$.

The error term has a simple geometric interpretation in terms of the geometry of the surface $V$. More precisely, the value $\lambda_1$ is proportional to the length of the shortest closed geodesic dividing the surface into two parts [74].

**Remark 6.3.** Hilbert and Polya proposed the idea of tying to understand the location of the zeros of the Riemann zeta function in terms of eigenvalues of some (as of yet) undiscovered self-adjoint operator whose necessarily real eigenvalues are related to the zeros. This idea has yet to reach fruition for the Riemann zeta function, but the approach works particularly well for the Selberg Zeta function where the associated operator is the Laplacian. Interestingly, it was Selberg who presented an alternative elementary proof (i.e., not using $\zeta(s)$, but significantly harder) of the Prime Number Theorem (Theorem 1.2) [73].

The corresponding results for surfaces of variable negative curvature and the Ruelle zeta function $\zeta_R(s)$ require a more dynamical proof which, unfortunately, gives a less quantifiable estimate on the size of the extension (and consequently the error term) [14], [62].

**Theorem 6.4.** For a compact surface of (variable) negative curvature $\kappa < 0$ there exists $\epsilon > 0$ such that $\zeta_R(s)$ has an analytic zero-free extension to $\text{Re}(s) > h(\phi) - \epsilon$, except for a simple pole at $s = h(\phi)$.

The same argument as in the case of the Selberg zeta function then leads to an exponential error term when counting closed geodesics on a surface of variable curvature [62]. Let $\Pi(x)$ again denote the number of closed geodesics $\gamma$ with length $l(\gamma) \leq x$. 
**Corollary 6.5.** For $0 < \epsilon' < \epsilon$, $\Pi(x) = \text{Li}(e^{h(\phi)x}) + O(e^{(h(\phi)-\epsilon')x})$.

We do not know, for example, if such error terms can be achieved for closed orbits of general weak mixing Anosov flows. In the wider context of Axiom A flows the zeta function may have poles arbitrarily close to the line $Re(s) = h(\phi)$, and consequently no such exponential error term could be expected [70], [60]. However, weaker error terms can sometimes be obtained under quite modest assumptions [63], [15].

### 6.2. The higher dimensional case.

There is a partial generalisation of Theorem 6.4 to higher dimensional manifolds [26]. We say that the variable negative sectional curvatures of a compact manifold $V$ are $a$-pinched if they all lie in the range $[-1, -a]$.

**Theorem 6.6.** Let $V$ be a compact manifold with variable negative sectional curvatures that are $\frac{1}{a}$-pinched. There exists $\epsilon > 0$ such that $\zeta_R(s)$ has an analytic zero-free extension to $Re(s) > h(\phi) - \epsilon$, except for a simple pole at $s = h(\phi)$.

Again, by complete analogy with the derivations of Corollaries 6.2 and 6.5, this has the following corollary under the same hypotheses as the theorem.

**Corollary 6.7.** For $0 < \epsilon' < \epsilon$, $\Pi(x) = \text{Li}(e^{h(\phi)x}) + O(e^{(h(\phi)-\epsilon')x})$

These can be viewed as a generalisation of previous results. Firstly, this generalizes Corollary 6.2 in the particular case of constant negative curvature [32], [30]. Secondly, this partly generalizes Corollary 4.4 for the case of geodesic flow on variable negative curvature manifolds by adding an exponential error estimate. However, we do not know whether the conclusions of Theorem 6.6 and Corollary 6.7 remain true without any pinching condition on the sectional curvatures.

### 7. Applications

The theory of dynamical zeta functions has a surprisingly wide range of applications, from which we present a small selection.

#### 7.1. Decay of correlations and resonances.

There is a complementary problem to counting closed orbits for Anosov flows in which one considers mixing (or decay of correlations) with respect to Gibbs measures. The problem of estimating the error terms on counting functions for closed orbits naturally corresponds to estimates on error terms in decay of correlations.

Let $\phi_t : M \to M$ be an Anosov flow and let $F : M \to \mathbb{R}$ and $G : M \to \mathbb{R}$ be two smooth functions.

**Definition 7.1.** We define the correlation function for a $\phi$-invariant probability measure $\mu$ by

$$\rho(t) := \int F\phi_tFd\mu - \int Fd\mu \int Gd\mu.$$
We say that the flow \( \phi_t : M \to M \) is (strong) mixing relative to the measure \( \mu \) if \( \rho(t) \to 0 \) for any \( F, G \in C^\infty(M) \).

A basic question to ask is about the speed at which \( \rho(t) \to 0 \) as \( t \to +\infty \). For this problem, a natural class of measures \( \mu \) to study are Gibbs measures for H"{o}lder continuous function \( A : M \to \mathbb{R} \) which includes, for example, the Sinai-Ruelle-Bowen measure (which is precisely the normalised Liouville measure in the case of geodesic flows) and the Bowen-Margulis measure of maximal entropy. In particular, when \( A = 0 \) the Gibbs measure is the Bowen-Margulis measure, and when \( A \) is the infinitesimal expansion along the unstable manifolds then the Gibbs measure is the Sinai-Ruelle-Bowen measure. In either case, we can conveniently characterize the associated Gibbs measure \( \mu_A \) in terms of weighted closed orbits:

\[
\int B d\mu_A = \lim_{T \to +\infty} \frac{\sum_{\lambda(\tau) \leq T} \lambda_B(\tau) e^{-\lambda_A(\tau)}}{\sum_{\lambda(\tau) \leq T} \lambda(\tau) e^{-\lambda_A(\tau)}}
\]

for any \( B \in C^0(M) \), where \( \lambda_A(\tau) = \int_0^{\lambda(\tau)} A(\phi_t x_\tau) dt \) and \( \lambda_B(\tau) = \int_0^{\lambda(\tau)} B(\phi_t x_\tau) dt \), for any \( x_\tau \in \tau \) [52], [9].

**7.1.1. Geodesic flows on surfaces.** If it has been known since the work of Fomin and Gelfand [19] that the geodesic flow on compact surfaces with constant curvature \( \kappa = -1 \) has exponential decay of correlations with respect to the normalized Liouville measure. Their proof used representation theory and the associated decay of matrix coefficients. However, these methods do not extend to the geometric setting of manifolds with variable negative curvature and a different approach is required [14].

**Theorem 7.2.** Let \( \phi_t : M \to M \) be the geodesic flow for a compact surface with variable negative curvature and let \( \mu \) be a Gibbs measure for a H"{o}lder continuous function. Then the correlation function \( \rho(t) \) tends to zero exponentially fast, i.e., there exist constants \( C, \epsilon > 0 \) such that \( |\rho(t)| \leq Ce^{-\epsilon t} \) for all \( t > 0 \).

In particular, this result applies to the important examples of the Bowen-Margulis and normalized Liouville measures described above.

**7.1.2. Geodesic flows in higher dimensions.** It was also shown in [19] that the geodesic flow on a three dimensional manifold with constant curvature \( \kappa = -1 \) has exponential decay of correlations with respect to the Liouville measure, and the basic method generalises to arbitrary dimensions [50]. Moreover, it also applies to frame flows for three dimensional manifolds, which has been useful, for example, in the recent work of Kahn and Markovic [38]. However, for manifolds with variable negative sectional curvatures a dynamical viewpoint is again necessary [14].

**Theorem 7.3.** Let \( \phi_t : M \to M \) be the geodesic flow for a compact manifold with variable negative sectional curvatures that are \( \frac{1}{3} \)-pinched and let \( \mu \) be a Gibbs measure for a H"{o}lder continuous function. Then the correlation function \( \rho(t) \) tends to zero exponentially fast.
The proof of Dolgopyat of the above theorem is stated for normalized Liouville measure, but there is additional property required for more general Gibbs measures, which can apparently be deduced for geodesic flows using a Shadowing Lemma of Mohsan [57] (see also [78], [79]). In particular, the result applies to both the Bowen-Margulis and normalized Liouville measures, although in these particular cases the following stronger results are known [42], [26].

**Theorem 7.4.** Let \( \phi_t : M \to M \) be the geodesic flow for a compact manifold with variable negative sectional curvatures.

1. If \( \mu \) is the normalized Liouville measure, then the correlation function \( \rho(t) \) tends to zero exponentially fast.

2. If the sectional curvatures are \( \frac{1}{9} \)-pinched and \( \mu \) is the Bowen-Margulis measure then the correlation function \( \rho(t) \) tends to zero exponentially fast.

By contrast, in the more general setting of Axiom A flows it is possible to give examples there the flow mixes arbitrarily slowly [70], [60].

**7.1.3. Fourier transforms and resonances.** A standard approach to understanding the asymptotic behaviour of such functions \( \rho(t) \) is by considering the Fourier transform. More precisely, we write

\[
\hat{\rho}(z) = \int_{-\infty}^{\infty} e^{izt} \rho(t) dt, \quad z \in \mathbb{C},
\]

where it is defined. The following well known classical result gives explicit connection between the domain of \( \hat{\rho}(z) \) and the asymptotic behaviour of \( \rho(t) \).

**Theorem 7.5** (Paley-Wiener). For a function \( \rho(t) \) the following are equivalent.

1. The Fourier transform \( \hat{\rho}(z) \) has an analytic extension to some strip \( |\text{Im}(z)| < \epsilon \), and is integrable along lines parallel to the real axis.

2. \( \rho(t) \) tends to zero exponentially fast.

An early result on the domain of the Fourier transform \( \hat{\rho}(z) \) of the correlation function was the following [60], [71].

**Theorem 7.6.** Let \( \phi_t : M \to M \) be a \( C^\infty \) Anosov flow and let \( \mu \) be a Gibbs measure for a Hölder continuous function. Then there exists \( \epsilon > 0 \) so that the function \( \hat{\rho}(z) \) has a meromorphic extension to \( |\text{Im}(z)| < \epsilon \).

The poles in the meromorphic extension given in Theorem 7.6 are sometimes called resonances. The original proof of Theorem 7.6 used the method of symbolic dynamics and thus even applies in the more general context of \( C^1 \) Axiom A flows. In the particular context of the normalized Liouville measure [11], [12] and the Bowen-Margulis measure [26] there are stronger results.

Perhaps somewhat surprisingly, the proof of Theorem 7.6 shows in the case that the Gibbs measure is the Bowen-Margulis measure \( \mu_0 \) that the poles of \( \hat{\rho}(z) \) are intimately related to the poles of the Ruelle zeta function \( \zeta_R(s) \).
Figure 5. (a) The poles for \( \zeta_R(s) \); and (b) The poles for \( \tilde{\rho}(z) \). We can informally think of this as translating by \(-h(\phi)\), turning the picture through 90 degrees, and adding its reflection.

**Theorem 7.7.** The poles \( z = \pm b \pm ia \) for \( \tilde{\rho}(z) \) (with \( a, b \in \mathbb{R}^+ \)) give rise to poles \( s = h(\phi) + a \pm ib \) for \( \zeta_R(s) \).

### 7.2. Computation of numerical values.
Dynamical zeta functions can sometimes be used to give alternative expressions for certain numerical dynamical characteristics, such as the Hausdorff dimension of invariant sets or the Lyapunov exponents, and thus provide an alternative method for their computation which is often quite efficient. This basic method, based on what are now known as cycle expansions, was pioneered by Cvitanović and his coauthors [13]. We briefly describe two applications of this approach.

#### 7.2.1. Hausdorff Dimension.
We can consider a rational map \( T : \hat{\mathbb{C}} \to \hat{\mathbb{C}} \) with a hyperbolic Julia set \( J \) (i.e., \( J \) is the closure of the union of the periodic points \( T^n z = z \) for \( |(T^n)'(z)| > 1 \), and we require that \( \sup_{z \in J} |(T^n)'(z)| > 1 \)). The Hausdorff Dimension \( \dim_H(J) \) of the Julia set can then be approximated using the values \( |(T^n)'(z)| \) of the derivatives at period points \( T^n z = z \) with periods less than \( N \). McMullen showed how these values could be used to compute approximations \( d_N \) to the Hausdorff dimension satisfying \( \dim_H(J) = d_N + O(\theta^N) \), for \( N \geq 1 \) and a fixed value \( 0 < \theta < 1 \), using approximations based on eigenvalues of matrices [49]. Using an approach based on the Ruelle zeta function shows that precisely the same values for periodic points, but used in different combinations, leads to a faster approximation \( D_N \) with \( \dim_H(J) = D_N + O(\theta^{N^{3/2}}) \), for \( N \geq 1 \) and a fixed value \( 0 < \theta < 1 \). Analogous results hold for the limit sets of Schottky groups and certain related Kleinian groups [34], as well as for continued fractions with deleted digits [35].

#### 7.2.2. Lyapunov Exponents.
Given a finite set of \( d \times d \) matrices \( A_1, \ldots, A_k \), with \( d \geq 2 \) and \( k \geq 2 \), and a probability vector \( (p_1, \ldots, p_k) \) one can associate the
Figure 6. Two hyperbolic Julia sets: (i) a Douady Rabbit; and (ii) a quasi-circle.

(largest) Lyapunov exponent [24] defined by

$$\lambda = \lim_{n \to +\infty} \frac{1}{n} \sum_{i_1, \ldots, i_n} p_{i_1} \cdots p_{i_n} \log \|A_{i_1} \cdots A_{i_n}\|.$$ 

In the particular case that the matrices are strictly positive then it is possible to approximate $\lambda$ using the maximal eigenvalues $\lambda_{i_1, \ldots, i_n}$ of the finite products $A_{i_1} \cdots A_{i_n}$ with $1 \leq n \leq N$ [58]. Using the Ruelle zeta function these values can be used to get approximations $\lambda_N$ to the Lyapunov exponent satisfying $\lambda = \lambda_N + O(\theta^{N^{1+\frac{1}{d}}})$, for $N \geq 1$ and a fixed value $0 < \theta < 1$ [61]. This has applications, for example, to computing entropy rates for binary symmetric processes. There are similar types of estimates for the Lyapunov exponents for $C^\omega$ Markov expanding maps [36].

### 7.3. Variations on the theme of the geometric zeta function.

The dynamical viewpoint sometimes provides a useful tool for extending the zeta function $\zeta_S(s)$ for geodesic flows to related settings. We briefly illustrate this viewpoint with the examples of infinite volume surfaces and the semi-classical zeta function.

#### 7.3.1. Zeta functions for infinite volume surfaces.

The setting of geodesic flows on infinite area surfaces associated to convex cocompact Fuchsian groups is one in which the dynamical viewpoint proves particularly useful. Although the associated geodesic flow is not Anosov the restriction to the recurrent part of the flow, which contains all of the closed orbits, is essentially a real analytic Axiom A flow and so the associated zeta function $\zeta_S(s)$ can be studied by adapting the dynamical approach of [66] and, in particular, it can be shown to have a meromorphic extension to $\mathbb{C}$.

However, in this setting the zeros and poles of the zeta function $\zeta_S(s)$ are now more difficult to describe than in the case of compact surfaces, associated
to cocompact Fuchsian groups. As a consequence of the dynamical approach and estimates on the zeta function some results are known on the distribution of zeros and poles [56], [4]. Moreover, it is known that a weak analogue of Theorem 6.1 still holds, in as much as there is a non-zero analytic extension to a half-plane \(\text{Re}(s) > h(\phi) - \epsilon\) in the spirit of Theorem 6.4 [51]. But the empirical behaviour of the resonances appears to be very different from that of the zeta function for compact surfaces [5].

7.3.2. Zeros for semi-classical zeta functions. Recently, Faure and Tsujii proved new results on the semi-classical zeta function

\[
\zeta_{SC}(s) = \exp \left( - \sum_{\tau} \sum_{m=1}^{\infty} \frac{e^{-s m \lambda(\tau)}}{m \det(I - D(\tau)^m)^{1/2}} \right)
\]

where \(D(\tau)\) is the Jacobian for the Poincaré map for \(\tau\) intersecting a small transverse section to the flow [18]. They have shown that some of the results on the locations of zeros for the original Selberg zeta function \(\zeta_S(s)\) have analogs for \(\zeta_{SC}(s)\). In particular, they show that the zeros, with only finitely many exceptions, lie in vertical strips.

7.4. Variations on the theme of the Ruelle zeta function. There are also interesting variations on the Ruelle zeta function \(\zeta_R(s)\) for Anosov flows. We illustrate this with a more general weighted Ruelle zeta function and \(L\)-functions.

7.4.1. The generalized Ruelle zeta function. A more general version of the Ruelle zeta function takes the form:

\[
\zeta_A^R(s) = \prod_{\tau} \left( 1 - e^{\lambda_A(\tau) - s\lambda(\tau)} \right)^{-1}, \quad s \in \mathbb{C},
\]

where \(A : M \to \mathbb{C}\) is a \(C^\infty\) function and as before \(\lambda_A(\tau) = \int_0^{\lambda(\tau)} A(\phi_t x) dt\), for any \(x \tau \in \tau\) This converges to a non-zero analytic function for \(\text{Re}(s) > P(A)\), where \(P(A)\) is the pressure of the function \(A\). (In the special case that \(A = 0\) is identically zero then \(P(0) = h(\phi)\) and the zeta function reduces to the original Ruelle zeta function, i.e., \(\zeta_A^R(s) = \zeta_R(s)\).) This was studied in [69] and was shown to have a simple pole at \(P(s)\) and a meromorphic extension to a larger domain. This leads to corresponding asymptotic and equidistribution results for weighted closed orbits [52]. Finally, a closely related zeta function is the differential zeta function introduced in [53].

7.4.2. \(L\)-functions. In prime number theory, the Riemann zeta function \(\zeta(s)\) has a useful generalization to \(L\)-functions which are used in the study of the distribution of primes in congruence classes.
The analogue of these complex functions for Anosov flows are given by

\[ L_{R_X}(s) = \prod_{\tau} \det \left( 1 - e^{-s\lambda(\tau)} R_X([\tau]) \right)^{-1}, \quad s \in \mathbb{C}, \]

where \( R_X : \pi_1(M) \to U(n) \) is an irreducible unitary representation, and which again converges for \( \text{Re}(s) > h(\phi) \) [55]. In the particular case that \( n = 1 \) and \( R_X = I \) is trivial then the \( L \)-function reduces to the original Ruelle zeta function, i.e., \( L_I(s) = \zeta_R(s) \).

A particularly elegant variant of this approach is where one considers the special case of the geodesic flow \( \phi_t : M \to M \) on the unit tangent bundle \( M = SV \) of a compact negatively curved surface \( V \) and \( \chi : H_1(V,\mathbb{Z}) \to \mathbb{C} \) is a character. This associates to each closed geodesic \( \gamma \) with homology class \([\gamma]\) the weight \( \chi([\gamma]) \). Using properties of these \( L \)-functions one can show that for any \( \alpha \in H_1(V,\mathbb{Z}) \) the number \( \Pi(x,\alpha) \) of closed geodesics with length \( l(\gamma) \leq T \) and \([\gamma] = \alpha\) satisfies

\[ \Pi(x,\alpha) \sim C e^{\frac{h(\phi)}{2} x} \left( \frac{T}{b+1} \right)^{b+1}, \quad \text{as} \quad x \to +\infty, \]

where \( b \) is the first Betti number of \( M \) [59], [39].

Remark 7.8. Among the many topics we have not discussed, are the Patterson conjecture [10], the Lefschetz theorem for flows [23], [37], and results on the closely related Poincaré series. For many other topics related to counting problems, we refer the reader to [75].

References


[33] Hunt, T. J., MacKay, R. S., Anosov parameter values for the triple linkage and a physical system with a uniformly chaotic attractor, Nonlinearity 16 (2003), 1499–1510.


Zeta functions


[75] Sharp, R., Periodic orbits of hyperbolic flows, 73–138, in [48].


Department of Mathematics, Zeeman Building, Warwick University, Coventry, CV4 7AL, England
E-mail: masdbl@warwick.ac.uk