From Dynamical Systems, Theory and Applications, (Ed. by Jurgen Moser), Springer Lecture Notes in Physics 38, 1975.

SPECTRAL INVARIANTS AND SMOOTH ERGODIC THEORY

D. A. Lind

Department of Mathematics, University of California, Berkeley 94720

1. Introduction

My purpose here is to briefly discuss the influence of spectral theory on the study of measure preserving transformations, and then to indicate how spectral invariants may play a role in deciding whether a given measure preserving transformation is isomorphic to a smooth one. As part of this discussion, I give a simple construction of a diffeomorphism of the two dimensional torus which has an everywhere discontinuous eigenfunction. I also hope to stimulate interest in the recent work of Anosov and Katok on constructing diffeomorphisms with prescribed ergodic behavior.

2. Spectral Theory of Measure Preserving Transformations

In the following X will denote a measure space with a σ-algebra of measurable subsets on which a probability measure μ is defined. To avoid unpleasant and needless pathology, we require X to be a so-called Lebesgue space, namely that X be measure isomorphic to the unit interval equipped with Lebesgue measure. For an axiomatic treatment of Lebesgue spaces, see Rohlin [17]. All "naturally occurring" measure spaces are Lebesgue.

A one-to-one and onto transformation T of one measure space to another is said to be measurable if $T^{-1}(E)$ is measurable whenever E is. T is called measure preserving if $T^{-1}(E)$ and E have the same measure. The objects of study in ergodic theory are measure preserving transformations of a space onto itself. Two such transformations, T acting on X and T' acting on X', are essentially the same from this point of view if there is a measure preserving transformation $\Phi: X \to X'$ such that $\Phi(Tx) = T'(\Phi x)$ for almost every $x \in X$. The basic problem of ergodic theory is to find ways of detecting whether two transformations, specified in perhaps very different ways, are isomorphic. Historically, spectral invariants have played a key role in this study.

The idea of using spectral theory to study measure preserving transformations goes back to Koopman [13]. He observed that a (measure preserving) transformation T of X induces a unitary operator U_T on $L^2(X)$ by the formula $(U_T f)(x) = f(Tx)$. Isomorphic transformations induce unitarily equivalent operators. Thus properties of the induced operators which are unitary invariants are the same for isomorphic transformations. These so-called spectral invariants allow us to prove that some transformations are not isomorphic.

The importance of these invariants becomes clear when one realizes that for over two decades, until Kolmogorov introduced entropy in 1958, they were the only useful ones known.

A transformation T is ergodic if whenever $T(E) = E$, then $\mu(E) = 0$ or 1. This is easily equivalent to the property that the only functions in $L^2(X)$ invariant under U_T are constant. This means that T is ergodic if and only if 1 is a simple eigenvalue of U_T, so that ergodicity is a spectral invariant.

In the same spirit, the point spectrum of U_T is a spectral invariant. Recall that a complex number λ is in the point spectrum P_T of U_T if there is a nonzero eigenfunction $f \in L^2(X)$ such that $U_T f(x) = \lambda f(x)$. P_T is a subgroup of the multiplicative unit circle group $\mathbb{T}$.

A good example to keep in mind is translation on the n-dimensional torus group $\mathbb{T}^n = \mathbb{T} \times \cdots \times \mathbb{T}$ with Haar measure. If $\lambda_i \in \mathbb{T}$, then translation on $\mathbb{T}^n$ by $(\lambda_1, \ldots, \lambda_n)$ is a measure preserving transformation whose point spectrum is the subgroup of $\mathbb{T}$ generated by $\lambda_1, \ldots, \lambda_n$. This translation is ergodic if whenever integers m_j are such that $\lambda_1^{m_1} \ldots \lambda_n^{m_n} = 1$, then all the $m_j = 0$. In this case the λ_i are said to be independent over the rationals. The eigenfunctions for this translation are just the multidimensional exponentials (the characters of $\mathbb{T}^n$), and since these span $L^2(\mathbb{T}^n)$, the point spectrum comprises the entire spectrum.

Koopman's idea is completely successful in this type of situation. Say that an ergodic transformation T has discrete spectrum if the eigenfunctions of U_T span $L^2(X)$.

<u>Discrete Spectrum Theorem</u>. If S and T are transformations with discrete spectrum, then S is measure isomorphic to T if and only if U_S is unitarily equivalent to U_T. Furthermore, this occurs if and only if $P_S = P_T$.

It can also be shown that any countable subgroup of $\mathbb{T}$ is the point spectrum of a transformation with discrete spectrum, and that every discrete spectrum transformation is isomorphic to a translation on a compact abelian group. For further details, see Halmos [9].

These facts give a very satisfying classification of transformations with discrete spectrum. However, as soon as the transformation has any stronger mixing properties such as weak mixing (see [6] for a definition), the point spectrum reduces to $\{1\}$ and so is useless for classification. Indeed, all Bernoulli shifts have unitarily equivalent induced operators and so are indistinguishable from the spectral point of view. Since there are nonisomorphic Bernoulli shifts, spectral invariants are useful but far from complete.

subgroup of $\mathbb{T}$ generated
f whenever integers m_j
e $m_j = 0$. In this case
rationals. The eigen-
multidimensional exponen-
e span $L^2(\mathbb{T}^n)$, the

in this type of situa-
has discrete spectrum

are transformations
isomorphic to T if
U_T. Furthermore, this

subgroup of $\mathbb{T}$ is the
ete spectrum, and that
morphic to a transla-
details, see Halmos

ification of transforma-
on as the transforma-
as weak mixing (see [6]
to {1} and so is use-
li shifts have
are indistinguishable
are nonisomorphic
ful but far from

3. Smooth Transformations

Let M be a compact, connected, smooth (i.e., C^∞) manifold, and suppose that T is a smooth diffeomorphism of M which preserves a smooth measure μ, so that μ is given by a smooth positive density function on M. The manifold M now becomes the measure space we called X in Section 2, and the σ-algebra is the μ-completion of the Borel σ-algebra on M generated by the compact sets. A measure preserving transformation is defined to be smooth if it is measure isomorphic to such a diffeomorphism. Certain smooth systems, arising from the Hamiltonian flow on the energy surface of a mechanical system, were the original motivating examples of ergodic theory. An immediate question is: Are all ergodic transformations smooth? Or has ergodic theory, by throwing away the differentiable structure of the original examples, greatly enlarged the class of objects it studies?

The only definite result in this direction is that smooth transformations have finite entropy. A good source for the definition of entropy and related results is Billingsley [6]. Kouchnirenko [14] first proved this fact by using the isoperimetric inequality (see Arnold and Avez [4]). Bowen [7] has given a neat proof that the entropy of T is bounded above by $\max\{0, \dim M \cdot \sup_{x\in M} \| dT\, T_x M \|\}$, where the norm on the derivative is taken with respect to a suitable Riemannian metric on M.

The example in the previous section of translation T on $\mathbb{T}^n$ by the element $(\lambda_1,\ldots,\lambda_n)$ is certainly smooth.

The rank of P_T, written $\operatorname{rk} P_T$, is defined to be the largest number of elements in P_T which are independent over the rationals. Since P_T for the torus translation is generated by $\{\lambda_1,\ldots,\lambda_n\}$, it is easy to check that $\operatorname{rk} P_T \leq n$, where n is the dimension of the manifold $\mathbb{T}^n$.

Up until about five years ago, in all smooth cases which could be checked, it turned out that

$$\text{(1)} \qquad \operatorname{rk} P_T \leq \dim M .$$

Indeed, Kolmogorov observed this in his address to the 1954 International Congress of Mathematicians [12], and it was conjectured to be generally true by Arnold and Avez [4; Appendix 16].

If the eigenfunctions of the diffeomorphism T are assumed to have certain smoothness properties, then finite bounds for $\operatorname{rk} P_T$ are known. We give below a simple proof that (1) holds if the eigenfunctions are once differentiable. Avez [5] has shown that it holds under the assumption of mere continuity of the eigenfunctions, and under the same hypothesis Arnold and Avez [4, Appendix 16] have shown that $\operatorname{rk} P_T \leq \dim H_1(M;\mathbb{R})$, the bound being the first Betti number of M.

Unfortunately, as was already known to Kolmogorov in 1953 [11], there are smooth systems with everywhere discontinuous eigenfunctions. Since examples seem to be available only in Russian, we give below a construction of this phenomenon on $\mathbb{T}^2$. Such badly behaved eigenfunctions obstructed further progress on the problem. Then about five years ago Anosov and Katok [3] published some constructions of diffeomorphisms which demolished the conjecture (1), but they left many questions open. We will describe their main result, and point out some areas which remain murky.

Let us first prove (1) for the point spectrum of well behaved eigenfunctions. Let D_T be the subgroup of P_T consisting of eigenvalues corresponding to once differentiable eigenfunctions.

<u>Theorem</u>. If T is an ergodic diffeomorphism of a manifold M, then $\operatorname{rk} D_T \leq \dim M$.

<u>Proof</u>: Let $\{\lambda_1,\ldots,\lambda_r\} \subset D_T$ be independent over the rationals.

Choose differentiable eigenfunctions f_j for λ_j. Since T is ergodic, $|f_j|$ is constant almost everywhere, hence everywhere by continuity. Normalize the eigenfunctions so that $|f_j| \equiv 1$ on M. Define a map $F: M \to \mathbb{T}^r$ by $F(x) = (f_1(x),\ldots,f_r(x))$. Notice that $F(Tx) = (\lambda_1,\ldots,\lambda_r)\cdot F(x)$. Hence invariance of M under T implies that the image $F(M)$ is invariant under translation by $(\lambda_1,\ldots,\lambda_r)$. Rational independence of the λ_i implies that the powers of $(\lambda_1,\ldots,\lambda_r)$ are dense in $\mathbb{T}^r$. Since $F(M)$ is compact and invariant under a dense set of translations, it must be all of $\mathbb{T}^r$. Now F is a differentiable onto mapping, so that the dimension of the range is bounded by the dimension of the domain. This gives $r \leq \dim M$, which proves the result.

4. Discontinuous Eigenfunctions

Our construction uses measure preserving flows, so we begin with some preliminary material. A measure preserving flow $\{T_t: t \in \mathbb{R}\}$ on a measure space X is a collection of measure preserving transformations satisfying the group property $T_sT_t = T_{s+t}$ together with the measurability assumption that $(x,t) \mapsto T_tx$ is a jointly measurable transformation from $X \times \mathbb{R}$ to X. A function $\psi: X \to \mathbb{R}$ is an eigenfunction for $\{T_t\}$ with eigenvalue $\lambda \in \mathbb{R}$ if $\psi(T_tx) = \exp(i\lambda t)\psi(x)$. This function is therefore an eigenfunction for each T_{t_0} with eigenvalue $\exp(i\lambda t_0)$.

We can build such flows from transformations by using the following suspension technique first introduced by von Neumann [15]. Suppose B is a measure space and $T: B \to B$ is measure preserving. Let $f: B \to (0,\infty)$ be measurable. There is a natural measure on the the region $X = \{(x,y): x \in B,\ 0 \leq y \leq f(x)\}$ under the graph of f obtained by restricting to X the product of the measure on B with Lebesgue measure on $\mathbb{R}$. The flow $\{S_t\}$ on X defined by moving a point (x,y) vertically at unit speed until $y = f(x)$, identifying $(x,f(x))$ with $(Tx,0)$, and continuing to flow vertically at $(Tx,0)$, is measurable and measure preserving. $\{S_t\}$ is said to be the flow built under the function f with base transformation T. Ambrose and Kakutani [2] have shown that all measure preserving flows can be put into this form.

In this construction, if B is a manifold, T a diffeomorphism of B, and f a smooth positive function on B, then X becomes a manifold M by identifying $(x,f(x))$ with $(Tx,0)$, i.e., by gluing the top and bottom edges of X together using f. Also, $\{S_t\}$ is a smooth flow on M.

A complex valued function on a manifold is everywhere discontinuous if on every open subset it does not agree almost everywhere with a continuous function. The function has an essential discontinuity at a point if it is not equal almost everywhere to a function continuous at the point.

The construction of an everywhere discontinuous eigenfunction for a smooth ergodic transformation proceeds as follows. Using a function constructed via harmonic analysis, we produce a measurable but not continuous conjugacy Φ between an irrational flow $\{S_t\}$ on $\mathbb{T}^2$ and a smooth flow $\{T_t\}$ on a manifold M diffeomorphic to $\mathbb{T}^2$. Then an exponential eigenfunction for $\{S_t\}$ carries over under Φ to everywhere discontinuous eigenfunction for $\{T_t\}$. A proper choice of t_0 then gives the ergodic diffeomorphism $T = T_{t_0}$ desired.

We now switch to additive notation on $\mathbb{T}$, it being considered as $[0,1)$ with addition modulo 1. For the sake of clarity, suppose first that we can find a function $\phi\colon \mathbb{T} \to [0,1)$ and $\omega \in \mathbb{T}$ such that $\phi(x+\omega) - \phi(x)$ is smooth, while $\exp(2\pi i\, \phi(x))$ has an essential discontinuity at 0. We will later modify the argument to handle unbounded ϕ, and then construct such a function. Put $g(\xi) = \phi(\xi+\omega)-$

$-\ \phi(\xi) + 1$, a smooth positive function on $\mathbb{T}$. Let $\{S_t\}$ acting on X be the flow built under the constant function 1 with translation on $\mathbb{T}$ by ω as base transformation, and $\{T_t\}$ acting on M be the smooth flow built under g with the same base transformation as $\{S_t\}$. A typical point in X is denoted by (x,y), and one in M by (ξ,η). There is a natural map $x = \beta(\xi)$ from the base of $\{T_t\}$ to that of $\{S_t\}$.

We will define a conjugacy Φ: $M \to X$ using ϕ. Basically, Φ restricted to the base of $\{T_t\}$ is $\phi \circ \beta$, and then Φ extends to all of M by using the flow. More precisely,

$$\Phi(\xi,\eta) = \begin{cases} (\beta(\xi),\eta+\phi(\beta(\xi))) , & \text{if } 0 \le \eta \le 1-\phi(\beta(\xi)) , \\ (\beta(\xi)+\omega,\eta+\phi(\beta(\xi))-1) , & \text{if } 1-\phi(\beta(\xi)) \le \eta \le g(\xi) . \end{cases}$$

It is easy to check that Φ maps orbits to orbits and preserves the direction and speed of the flow. An easy argument using Fubini's theorem shows that Φ is measurable and measure preserving. Hence Φ is a conjugacy between $\{T_t\}$ and $\{S_t\}$.

Let χ be the eigenfunction for $\{S_t\}$ defined on X by $\chi(x,y) = \exp(2\pi iy)$. Then $\psi = \chi\circ\Phi$ is an eigenfunction for $\{T_t\}$ with eigenvalue 1. Now $\psi(\xi,0) = \exp(2\pi i\ \phi(\beta(\xi)))$ has an essential discontinuity at $\beta = 0$. But since $\psi \circ T_t = [\exp(2\pi it)]\psi$, this discontinuity can be translated by $\{T_t\}$ along the orbit of $(0,0)$, which is dense in M. Thus in every open subset of M the eigenfunction has an essential discontinuity, so that it is everywhere discontinuous.

Unhappily, it is not easy to find a bounded discontinuous function ϕ on $\mathbb{T}$ such that $\phi(x+\omega) - \phi(x)$ is smooth. Below we will produce a necessarily unbounded ϕ and an irrational $\omega \in T$ for which $\phi(x+\omega) - \phi(x)$ is smooth, bounded in absolute value by 1/4, and such that $\exp(2\pi i\ \phi(x))$ has an essential discontinuity at 0.

Let us indicate how to modify the definition of Φ if ϕ is unbounded. Basically, we use more of the orbit of $\{S_t\}$ to wrap the graph of ϕ.

For any real number a, let $[a]$ denote the greatest integer in a, and $\{a\}$ the fractional part of a. We then define

$$\Phi(\xi,\eta) = \begin{cases} (\xi+[\phi(\beta(\xi))]\omega,\ \eta+\{\phi(\beta(\xi))\}), & \text{if } 0 \leq \eta \leq 1 - \{\phi(\beta(\xi))\}, \\ (\xi+[\phi(\beta(\xi))]\omega+\omega, \eta+\{\phi(\beta(\xi))\}-1) & \text{if } 1-\{\phi(\beta(\xi))\} \leq \eta \leq g(\xi). \end{cases}$$

The same arguments as before show that Φ is a conjugacy between $\{T_t\}$ and $\{S_t\}$. Φ reduces to the previous conjugacy if $0 \leq \phi \leq 1$. We also have as before that $\psi = \chi \circ \Phi$ is an everywhere discontinuous eigenfunction for $\{T_t\}$.

To obtain an ergodic diffeomorphism, choose t_0 such that $\exp(2\pi i t_0)$ and $\exp(2\pi i t_0 \omega)$ are independent over the rationals. Then S_{t_0} is isomorphic to translation on $\mathbb{T}^2$ by $(\exp(2\pi i t_0)$, $\exp(2\pi i t_0 \omega))$, and so is ergodic. Since T_{t_0} is isomorphic to S_{t_0}, $T = T_{t_0}$ is the required diffeomorphism.

The problem of finding a suitable ϕ and ω was considered by Furstenberg [8] in a somewhat different context. Our solution is parallel to his. Write the Fourier series of ϕ as

$$\phi(x) \sim \sum_{-\infty}^{\infty} c_n e^{2\pi i n x} .$$

If $h(x) = \phi(x+\omega) - \phi(x)$, then

$$h(x) \sim \sum_{-\infty}^{\infty} c_n \left(e^{2\pi i n \omega} - 1\right) e^{2\pi i n x} .$$

If ω is abnormally well approximable by rationals, the factor $e^{2\pi i n \omega} - 1$ becomes small often enough to force smoothness of h while at the same time allowing the c_n's to be large enough to force an essential discontinuity of f at 0.

The Fourier series we consider will be lacunary. For any sequence of integers $n_k \to \infty$, with $n_{-k} = n_k$, the function of $\{S_t\}$ to wrap the

greatest integer in
define

$\leq \eta \leq 1 - \{\phi(\beta(\xi))\}$,
$-\{\phi(\beta(\xi))\} \leq \eta \leq g(\xi)$.

conjugacy between
jacy if $0 \leq \phi \leq 1$.
where discontinuous

t_0 such that
over the rationals.
$(\exp(2\pi i t_0)$,
s isomorphic to

was considered by
. Our solution is
as

inx .

nals, the factor
smoothness of h
large enough to force

nary. For any
ne function

$$\phi(x) = \sum_{k\neq 0} \frac{1}{|k|} \exp(2\pi i n_k x)$$

is in $L^2(\mathbb{T})$. However, ϕ has an essential discontinuity at 0 since the Cesaro sums of the Fourier series of ϕ diverge at 0. By replacing ϕ with an appropriate multiple, we can guarantee that $\exp(2\pi i\, \phi(x))$ also has an essential discontinuity at 0. We will find n_k, ω such that

$$(2) \qquad n_k^s |\exp(2\pi i n_k \omega) - 1| \to 0 \quad \text{as} \quad k \to \infty$$

for all $s \geq 0$. This forces h to have derivatives of all orders. Let $n_1 = 5$, $n_{k+1} = n_k^{k+1}$, and put $\omega = \sum_1^\infty 1/n_k$. Then if $k > s+1$,

$$n_k^s\{n_k \omega\} = n_k^{s+1} \sum_{j=k+1}^{\infty} \frac{1}{n_j} \leq n_k^{s+1} \left(\frac{2}{n_k^{k+1}}\right) < \frac{2}{n_k},$$

which proves (2) and also that $|h| < 1/4$. This yields the desired ϕ and ω.

We remark that such a ϕ cannot be bounded. The proof uses lacunarity of the Fourier series of ϕ and the method of Riesz products. Let

$$P_N(x) = \prod_{k=1}^{N} (1 + \cos 2\pi n_k x) .$$

Since the sequence $\{n_k\}$ is lacunary, the Fourier coefficient of P_N at 0 is 1, and at each $\pm n_k$ is $\frac{1}{2}$. Also, $P_N \geq 0$ since each factor is nonnegative. Suppose there were a C with $|\phi| \leq C$. Then we would have

$$\int_0^1 \phi(x)\, P_N(x)\, dx = \sum_{k=1}^{N} \frac{1}{k} \to \infty \quad \text{as} \quad N \to \infty ,$$

while

$$\left| \int_0^1 \phi(x)\, P_N(x)\, dx \right| \leq C \int_0^1 |P_N(x)|\, dx = C \int_0^1 P_N(x)\, dx = C .$$

This contradiction shows that ϕ must be unbounded.

We finish by briefly discussing the recent work of Anosov and Katok, and pointing out a few of the many still uncharted portions of the field.

What Anosov and Katok have done in [3] is to construct, for any given $n = 0,1,2,\ldots,\infty$, a diffeomorphism whose point spectrum has rank n on any manifold which admits a measure preserving action of the circle group (such as the two dimensional disk). However, the specific eigenvalues for these diffeomorphism are abnormally well approximable by rationals, and have quite a specific form. Thus while these examples destroy the original conjecture (1), arbitrary rank is a long way from arbitrary spectrum.

Generally, very little is known about available ergodic behavior for diffeomorphisms of a given manifold. For instance, do the various counterexamples constructed by Ornstein have smooth versions? There are some known Bernoulli diffeomorphisms, such as toral automorphisms (Katznelson [10]), and smooth Bernoulli flows, such as geodesic flow on the unit tangent bundle of a manifold of negative curvature (Ornstein and Weiss [16]). However, can any manifold support such Bernoulli diffeomorphisms or flows? A negative answer would mean that there are topological restrictions on the energy surface of a Bernoulli mechanical system, and this could have physical consequences.

[1] Abraham
(1967).
[2] Ambrose
measura
[3] Anosov,
ergodic
Soc. 23
[4] Arnold,
mechani
[5] Avez, A.
C. R. A
[6] Billing
New Yor
[7] Bowen,
spaces,
[8] Furstenb
torus,
[9] Halmos,
Japan, n
[10] Katznel
Israel
[11] Kolmogor
invarian
763-766
[12] Kolmogor
classica
North Ho
English

References

[1] Abraham, R., Foundations of mechanics, Benjamin, New York (1967).

[2] Ambrose, W., and Kakutani, S., Structure and continuity of measurable flows, Duke Math. J. 9 (1942) 25-42.

[3] Anosov, D. V., and Katok, A. B., New examples in smooth ergodic theory. Ergodic diffeomorphisms, Trans. Moscow Math. Soc. 23 (1970) 1-35.

[4] Arnold, V. I., and Avez, A., Ergodic problems of classical mechanics, Benjamin, New York, 1958.

[5] Avez, A., Spectre discrete des systèmes ergodiques classiques, C. R. Acad. Sci. Paris 264 (1967) 49-52.

[6] Billingsley, P., Ergodic theory and information, Wiley, New York, 1965.

[7] Bowen, R., Entropy for group automorphisms and homogeneous spaces, Trans. Amer. Math. Soc. 153 (1971) 401-414.

[8] Furstenberg, H., Strict ergodicity and transformations of the torus, Amer. Jour. Math. 83 (1961) 573-601.

[9] Halmos, P. R., Lectures on ergodic theory, Publ. Math. Soc. Japan, no. 3, Math. Soc. Japan, Tokyo (1959).

[10] Katznelson, Y., Ergodic automorphism of T^n are Bernoulli, Israel Math. Jour. 10 (1971), 186-195.

[11] Kolmogorov, A. N., On dynamical systems with an integral invariant on the torus, Dokl. Akad. Nauk SSSR (N.S.) 93 (1953) 763-766 (in Russian).

[12] Kolmogorov, A. N., General theory of dynamical systems and classical mechanics, Proc. 1954 International Congress Math., North Holland: Amsterdam, 1957, pp. 315-333. (Russian; an English translation appears in Appendix D of [1]).

[13] Koopman, B. O., Hamiltonian systems and transformations in Hilbert spaces, Proc. Nat. Acad. Sci. 17 (1931) 315-318.

[14] Kouchnirenko, A. G., An estimate from above for the entropy of a classical system, Dokl. Akad. Nauk SSSR 161 (1965) 37-38.

[15] von Neumann, J., Zur Operatorenmethode in der klassichen Mechanik, Ann. of Math. 33 (1932) 587-642.

[16] Ornstein, D., and Weiss, B., Geodesic flows are Bernoullian, Israel J. Math. 14 (1973) 184-198.

[17] Rohlin, V. A., On the fundamental ideas of measure theory, Amer. Math. Soc. Transl., Ser. 1, Vol. 10, 1-54.