LINEAR AND FRACTIONAL RESPONSE FOR THE SRB MEASURE OF SMOOTH HYPERBOLIC ATTRACTORS AND DISCONTINUOUS OBSERVABLES

VIVIANE BALADI, TOBIAS KUNA, AND VALERIO LUCARINI

ABSTRACT. We consider a smooth one-parameter family $t \mapsto (f_t : M \to M)$ of diffeomorphisms with compact transitive Axiom A attractors Λ_t, denoting by $d\rho_t$ the SRB measure of $f_t|_{\Lambda_t}$. Our first result is that for any function θ in the Sobolev space $H_p^r(M)$, with $1 < p < \infty$ and $0 < r < 1/p$, the map $t \mapsto \int \theta \, d\rho_t$ is α-Hölder continuous for all $\alpha < r$. This applies to $\theta(x) = h(x)\Theta(g(x) - a)$ (for all $\alpha < 1$) for h and g smooth and Θ the Heaviside function, if a is not a critical value of g. Our second result says that for any such function $\theta(x) = h(x)\Theta(g(x) - a)$ so that in addition the intersection of $\{x \mid g(x) = a\}$ with the support of h is foliated by "admissible stable leaves" of f_t, the map $t \mapsto \int \theta \, d\rho_t$ is differentiable. (We provide distributional linear response and fluctuation-dissipation formulas for the derivative.) Obtaining linear response or fractional response for such observables θ is motivated by extreme-value theory.

1. INTRODUCTION

Response theory has long been an essential ingredient of statistical physics, because it provides reliable and intuitive tools to describe the changes in the physical invariant measure of a system resulting, e.g., from the introduction of a forcing or of a parametric modulation, in terms of some statistical properties of the unperturbed system. The theory — heuristic and not mathematically rigorous — was initially proposed by Kubo [20] in a seminal paper on statistical mechanical systems weakly driven outside of the thermodynamical equilibrium described by the equilibrium measure in the canonical ensemble. In this case, it is possible to use response theory to derive the so-called fluctuation-dissipation theorem which provides a fundamental connection between the free and forced (linear) fluctuations of a system [20] and has an enormous importance in a multitude of fields of physical sciences [26]. Response formulas describe the change in the expectation value of a given observable θ of the system.

Response theory has been extremely successful. However the limits involved in constructing the response formulas were not always proved to be well-defined. Ruelle [29, 28, 30] (see also [18]) established a rigorous response theory for transitive Axiom A [1] attractors. By definition,

---

[1] According to the *chaotic hypothesis* of Gallavotti and Cohen [13], Axiom A systems can be taken as effective models for dynamical systems with many degrees of freedom, and they hence have physical relevance.

Axiom A attractors [2] are uniformly hyperbolic, and it is well known that they possess an invariant SRB (physical) measure. The SRB measure is in general singular with respect to Lebesgue measure (so, from a physical point of view, the system can be viewed as out of equilibrium), but Ruelle showed that, *when used to average a smooth enough observable θ,* this measure is nevertheless differentiable with respect to the parameters controlling small modifications of the system, and he provided explicit expressions for the derivatives. Ruelle's theory has been successfully applied to predict the response to perturbations of out of equilibrium systems [25, 27, 12, 10], including the very high-dimensional case of a climate model [21]. Applying Ruelle's formulas to the perturbation of a statistical mechanical system whose statistics is described by the canonical ensemble (in contact with a reservoir at constant temperature), the classical results of Kubo can be straightforwardly recovered.

Modern spectral methods (see e.g. [16]), based on a transfer (Ruelle–Perron–Frobenius) operator acting on a suitable Banach space of anisotropic distributions, have provided new proofs and extensions of Ruelle's results. However, up to now, all existing rigorous works (in particular [29, 16]) establishing linear response for hyperbolic systems require smooth (at least C^1) observables θ. (The situation is very different for locally expanding dynamics, where bounded observables are more than enough, see e.g. [1] and references therein. This is due to the fact that the transfer operator there acts on Banach spaces of functions and not distributions.) The goal of the present short note[3] is to investigate response theory of Axiom A systems for non differentiable observables, allowing in particular discontinuities in the observable. Our results below (Theorems 2.1 and 3.3) apply to observables involving the Heaviside function $\Theta(x)$, for example $\theta(x) = \Theta(g(x))h(x)$, with g and h smooth, assuming that zero is not a critical value of g. Note that the expectation[4] value of $\Theta(g(x))$ gives the fraction of the total measure where g has positive value, while the expectation value of $h(x)\Theta(g(x))$ divided by the expectation value of $\Theta(g(x))$ gives the average of h conditional on the fact that g is positive. Therefore, such discontinuous observables have a clear probabilistic and physical interpretation. Here is an example regarding the analysis of extreme events [22] in chaotic dynamical systems: In [24] it is suggested that one can construct estimates of the parameters describing the statistics of extremes of a general observable g for Axiom A systems through suitable combinations of the expectation value of quantities of the form $(g(x) - a)^n \Theta(g(x) - a)$, for $n \in \mathbb{N}$, where a is a (noncritical) threshold close enough to the global maximum of g on the attractor of the system if g is smooth, while a tends to infinity if g is unbounded. Therefore, some regularity in the dependence of such discontinuous observables on perturbations of the dynamics will give control on the response of extreme events — an obviously relevant practical problem. (We refer in particular to [24, Section III], see Remark 2.3 below. The discussion around (2.2) below mentions related open questions.)

In the present work, we consider one-parameter families of transitive compact hyperbolic attractors $f_t : \Lambda_t \to \Lambda_t$, with Λ_t contained in a Riemann manifold M, denoting the SRB measure by ρ_t, and we obtain two main results:

---

[2]The terms Axiom A attractor and (uniformly) hyperbolic attractor are indeed used interchangeably in the literature and in the present paper.

[3]In this paper, we only consider discrete-time dynamical systems, i.e., iterations of a map $f : M \to M$, or of one-parameter families of maps $f_t : M \to M$.

[4]Replacing Θ by a Dirac mass δ_D also gives interesting observables, see Remarks 2.2 and 3.4.

In Section 2, we prove Theorem 2.1, which says that for any $1 < p < \infty$, any $0 < r < 1/p$, and any observable θ in the generalised (isotropic, ordinary) Sobolev space $H_p^r(M)$, the map

$$(1.1) \qquad\qquad t \mapsto \int \theta \, d\rho_t$$

is α-Hölder continuous for all exponents $\alpha < r$. (Remark 2.3 shows that this result applies to observables of the type $h(x)\Theta(g(x) - a)$, and in that case r can be taken arbitrarily close to 1.) We expect that it is possible to find examples of hyperbolic attractors f_t and observables $\theta \in H_p^r(M)$ for which the map (1.1) is not differentiable and maybe to show that $\alpha = r$ is the optimal regularity limitation. This is not the first time that a *fractional* (i.e., weak) formulation of the response has been obtained: It is known that, in nonuniformly hyperbolic situations where strong bifurcations are present, differentiability of the response can be violated (even for C^∞ observables). In fact, 1/2-Hölder upper *and lower* bounds have been obtained in the setting of the logistic family (i.e., unimodal interval maps, see [1, 4] and references therein). We emphasize that it is often however extremely difficult to distinguish numerically between differentiability (strong, linear, response) and Hölder continuity (weak, fractional, response), as is clearly discussed in the illuminating preprint [15] of Gottwald, Wormell, and Wouters, which came to our attention while we were finishing the present work.

In Section 3, our second main result, Theorem 3.3, gives ordinary (i.e., strong) linear response for observables $h(x)\Theta(g(x) - a)$, whose singularity set is "compatible" with the stable cone of the transitive attractor. The precise compatibility condition involves "admissible stable leaves," (see Definition 3.2), and requires that the set

$$\mathcal{W}_a = \{x \mid g(x) = a\} \cap \mathrm{supp}(h)$$

be foliated by admissible stable leaves. This condition, which implies transversality of the normal vector to $\mathcal{W}_a$ with the "wave-front set" of the SRB measure, will not surprise experts, but to our knowledge it is the first time that it is written down. By ordinary linear response, we mean here differentiability of the map (1.1), together with formulas for the derivative, both of linear response-type (3.3) and fluctuation-dissipation-type (3.2) (both formulas being in the sense of distributions!). We explain next how to interpret (distributional) fluctuation-dissipation expressions in hyperbolic settings, if the SRB measure $d\rho_0$ of f_0 is singular with respect to Lebesgue. The starting point is that $d\rho_0$ can be approached (in anisotropic norm) by a sequence $\rho_0^{(n)} dm$ of absolutely continuous measures, with smooth densities, either by mollification, or by iterating Lebesgue measure under the dynamics f_0^n. Then, setting $X_0 = \partial_t f_t|_{t=0} \circ f_0^{-1}$, our results on the distributional fluctuation-dissipation formula (3.2) easily imply that the terms of the sum

$$-\sum_{k=0}^{\infty} \int (\theta \circ f_0^k)[\mathrm{div}(X_0)\rho_0^{(n)} + \langle X_0, \mathrm{grad}\rho_0^{(n)} \rangle] \, dm$$

decay exponentially (uniformly in n), and the sum converges to $\partial_t \int \theta \, d\rho_t|_{t=0}$ as $n \to \infty$, if the observable θ lies in the dual of a suitable anisotropic space. (Admittedly, $\mathrm{grad}\rho_0^{(n)}$ is not very nice if $\rho_0^{(n)}$ is constructed by iterating Lebesgue measure, it is slightly easier to control if $\rho_0^{(n)}$ is obtained by convolution with an approximation of the Dirac mass.)

We would like to emphasize that, in hyperbolic situations (as opposed to locally expanding maps), anisotropic norms are still required to get exponential convergence to equilibrium or linear response *even if the SRB measure is absolutely continuous with respect to Lebesgue.* In particular, fluctuation-dissipation expressions such as (3.2) require the observable to have some modulus of continuity in the stable cones. However, in the case of an absolutely continuous SRB measure, the gradient of the SRB measure will be contained in a space of arbitrarily mild anisotropic distributions (in the notation below, instead of $W_p^{u-1,s-1}$ or $\mathcal{B}^{u-1,|s-1|}$ for $s < 0$, one can use $W_p^{u-1,\sigma}$ or $\mathcal{B}^{u-1,|\sigma|}$ for $\sigma < 0$ arbitrarily close to 0), and one can then hope to weaken slightly the condition on the observable. The toy-model case of smooth enough locally expanding maps is much easier to tackle, since then linear response holds for observables in spaces of distributions, so that one could obviously replace the Heaviside function by a Dirac mass δ_D, or derivatives of a Dirac mass. (To summarise, there is a hierarchy of difficulties, with locally expanding easiest, singular SRB measure hardest, and absolutely continuous SRB measure in the middle.)

At a technical level, we work with anisotropic Banach norms. For Theorem 2.1, it is more convenient to use a norm introduced in [8]. One of the lessons of the present work is that, for the fractional response of Theorem 2.1, only a weak, dynamics-independent, condition on the observable is needed, and a "cheap mollification trick" does the job. To prove Theorem 3.3, the geometric anisotropic norms from [16] are more convenient. The compatibility condition on $\mathcal{W}_a$ given above implies the transversality needed to check that the observable lies in the dual of the anisotropic space containing the gradient of the SRB measure, which is just what the proof requires.

2. Fractional response for hyperbolic attractors and observables in $H_p^r(M)$

Let M be a smooth d-dimensional Riemann manifold. Recall that a nonempty compact set $\Lambda_t \subset M$ is a (uniformly) hyperbolic attractor for a diffeomorphism $f_t : \Lambda_t \to \Lambda_t$, if there exists an open neighbourhood V_t of Λ_t so that $f_t(\overline{V_t}) \subset V_t$ and $\cap_{n \geq 0} f_t^n(V_t) = \Lambda_t$, and if there exist $C > 1$ and $\nu < 1$ so that tangent space $TM|_{\Lambda_t}$ over Λ_t decomposes into $E^u \oplus E^s$, with E^u and E^s two Df_t-invariant bundles, of respective dimensions $d_u \geq 1$ and $d_s \geq 1$, so that

$$\|D_x f_t^n|_{E_x^s \to E_{f_t^n(x)}^s}\| \leq C\nu^n, \quad \|D_x f_t^{-n}|_{E_x^u \to E_{f_t^{-n}(x)}^u}\| \leq C\nu^n, \quad \forall x \in \Lambda_t, \, \forall n \geq 1.$$

The attractor is transitive if there exists a point with a dense orbit. It is well known (see e.g. [19]) that if $f_t : \Lambda_t \to \Lambda_t$ is a transitive hyperbolic attractor as just defined, then f_t has a unique SRB (invariant probability) measure on Λ_t. In what follows, there will always be a compact subset Λ of M containing all the hyperbolic attractors Λ_t involved, so for all practical purposes we can assume that M is compact, denoting by dm normalised Lebesgue measure on M, and we can work with finite systems of smooth charts and partitions of unity.

For real numbers r and $1 < p < \infty$, let $H_p^r(M)$ be the Sobolev space of functions $\varphi \in L_p(M)$ so that, denoting by Δ the Laplacian in $\mathbb{R}^d$ and by $\widetilde{\varphi}_j$ the function φ in the jth chart, the L_p norm of each $(\mathrm{id} + \Delta)^{r/2}(\widetilde{\varphi}_j)$ is finite. If $r > 0$, then $H_p^r(M)$ is the closure of $C^u(M)$ for any $u > r$, for the same norm, see e.g. [2]. If $r = 0$ then $H_p^r(M) = L_p(M)$. Also, $H_p^{-r}(M)$ is the dual space to $H_{p'}^r(M)$ if $1/p + 1/p' = 1$, where duality is given by the scalar product

$\int_M \varphi_1 \bar{\varphi}_2 \, dm$. Recall that an element of $H_p^r(M)$ is continuous if $r > d/p$, but beware that we shall only be interested in the case $r < 1/p$.

The following theorem is our first technical result on compact transitive hyperbolic attractors. Its main interest lies in the fact that the assumption on the observable θ does not involve the dynamical system. Also, the proof (which uses a standard mollification trick) is remarkably simple.

Theorem 2.1 (Fractional response). *Fix $\beta \in (0,1)$. Let $t \mapsto f_t$, for $t \in [-\epsilon_0, \epsilon_0]$, be a $C^{2+\beta}$ family of C^3 diffeomorphisms f_t on a smooth Riemann manifold M, so that f_t has a transitive compact hyperbolic attractor $\Lambda_t \subset M$. Let ρ_t be the (unique) SRB measure of f_t on Λ_t. Let $\theta : M \to \mathbb{C}$ be so that $\theta \in H_p^r(M)$ for some $1 < p < \infty$ and $0 < r < 1/p$. Then there exists $\epsilon_1 \in (0, \epsilon_0]$ so that for any $\alpha < r$ the map*

$$t \mapsto \int_M \theta \, d\rho_t$$

is α-Hölder continuous on $[-\epsilon_1, \epsilon_1]$.

An important ingredient of the proof of Theorem 2.1 is the transfer operator $\mathcal{L}_t$ associated to f_t by setting

$$(2.1) \qquad \mathcal{L}_t(\varphi)(x) = \frac{\varphi(f_t^{-1}(x))}{|\mathrm{Det} D f_t(f_t^{-1}(x))|} \, , \quad \varphi \in C^2 \, ,$$

where $|\mathrm{Det} D f_t|$ is the Jacobian with respect to Lebesgue measure (so that the dual of the transfer operator leaves Lebesgue measure on Λ invariant). We shall let this operator act on anisotropic Banach spaces of distributions on which it is known to have a spectral gap.

Remark 2.2. Previous results on linear response ([29]–[18], [16, §8]) restrict to observables θ which are C^r for some $r \geq 1$. As pointed out by the referee, Demers and Liverani [11] obtained *fractional* (Hölder) response for two-dimensional piecewise hyperbolic (or smooth hyperbolic) systems and piecewise Hölder test functions. An inescapable requirement is that $\int \theta d\rho_t$ be well defined. Practically, θ must lie in the dual space to an anisotropic Banach space containing ρ_t. In particular, we cannot take θ to be a Dirac mass on M, since ρ_t cannot integrate such distributions in general (see the proofs of Theorems 2.1 and 3.3 for more information about ρ_t). Note however that one can probably show that ρ_t integrates distributions involving the Dirac distribution under a condition in the spirit of Theorem 3.3 below (see [16, Prop. 4.4], although this proposition cannot be applied directly since what we need is a statement on the *dual* space). This inescapable requirement does not suffice for full-fledged linear response, since what is relevant there is integration against gradρ_t and not against ρ_t itself, see Section 3 and Remark 3.4.

Remark 2.3 (Application to extreme value theory). Let $\Theta : \mathbb{R} \to \{0, 1\}$ be the Heaviside step function

$$\Theta(v) = 0 \, , \text{ for } v \leq 0 \, , \quad \Theta(v) = 1 \, , \text{ for } v > 0 \, .$$

An application of Theorem 2.1 to extreme value theory can be obtained by taking

$$\theta(x) = \theta_a(x) = \Theta(g(x) - a)$$

for a C^2 function $g : M \to \mathbb{R}$ (we assume for simplicity here that M is compact) and any threshold $a \in \mathbb{R}$ which is not a critical value of g. The theorem then gives α-Hölder continuity of the response for all $\alpha < 1$. (Indeed, $\Theta(g(x) - a)$ is just the indicator function of the set $\mathcal{D}_a = \{x \in M \mid g(x) > a\}$. We claim that this implies that $\theta(x) \in H_p^r(M)$ for all $1 < p < \infty$ and all $0 \le r < 1/p$, so that we can take any $\alpha < 1$ in Theorem 2.1 by letting p tend to 1 and r tend to $1/p$. If a is larger than the global maximum of g or smaller than the global minimum of g, then the claim is obvious since $\mathcal{D}_a$ coincides with the empty set or with M, respectively. Otherwise, recall (see e.g. [31]) that the indicator function of a compact interval belongs to $H_p^r(\mathbb{R})$ for all $0 \le r < 1/p$ if $1 < p < \infty$. If a is not a critical value of g then the implicit function theorem gives that the level set $\mathcal{W}_a = \{x \in M \mid g(x) = a\}$ is a finite union of smooth codimension-one submanifolds and it is the boundary of $\mathcal{D}_a$. Therefore, by Fubini[5], the indicator function of $\mathcal{D}_a$ belongs to $H_p^r(M)$, for all $1 < p < \infty$ and all $0 \le r < 1/p$, as claimed, since M is compact.)

As a last item of this remark, assume (without loss of generality) that the global maximum of g is equal to 0. For $\theta_b(x) = \Theta(g(x) - b)$, it would be interesting to study $\int \theta_b \, d\rho_t$ as a function of b, in particular in view of considering the regularity of the limit of the map[6]

$$(2.2) \qquad\qquad t \mapsto \frac{\int \theta_{as} \, d\rho_t}{\int \theta_a \, d\rho_t} \, ,$$

for fixed $s \in (0, 1)$, as $a \to 0$ (so that numerator and denominator both tend to zero, since 0 is the global maximum). Indeed, for fixed t, proving the existence of the limiting quotient (2.2) as $a \to 0$ and finding an expression for this limit are key steps to construct explicit extreme value laws for the observable $g(x)$. (We refer e.g. to [22, (3.1.7)] for i.i.d. random variables and to [22, (4.2.6)] for dynamical systems.) For occurrences of similar quotients in the physics literature, see [23, (12)] and [24, (5)], where two different kinds of observables are considered and the limiting expression is derived under the assumption that the limit (2.2) exists in a suitable sense ([24, Section III] includes a discussion about varying t). We also refer to the overview in [22, Chapter 8, (8.2.4), (8.2.13)]. (The notation in [22, 23, 24] is $T = a < 0$, and $Z + T = as$, so that $Z = T(s - 1) > 0$: This additive formulation does not give a useful limit as T tends to zero for fixed Z, but it is convenient for experiments or simulations.) The study of the limiting quotient map (2.2) is beyond the scope of this paper, and we postpone it to further works. The first test case would be that of locally expanding maps.

Proof of Theorem 2.1. Let V be a common open isolating (in fact, attracting) neighbourhood for all the f_t (up to reducing the value of ϵ_0): $\Lambda_t = \cap_{n \in \mathbb{Z}} f_t^n(V) = \cap_{n \in \mathbb{Z}_+} f_t^n(V)$, with $f_t(\overline{V})$ a strict subset of V. We can assume that the closure of V is compact. The support of each ρ_t is contained in V by construction.

Fix a finite system of C^∞ charts $\psi_i : U_i \to M$, $i = 1, \ldots, N$, with open bounded domains $U_i \subset \mathbb{R}^d$, so that $\overline{V} \subset \cup_{i=1}^M \psi_i(U_i)$, and fix a C^∞ partition of unity on V, i.e., functions ϕ_j

---

[5]See e.g. the proof of [6, Lemma 23] and the references therein to [31]. We emphasize however that we only need here θ to belong to a suitable (dual) space and not to be a multiplier.

[6]The quotient (2.2) is also interesting for a function $\theta_a = h(x)\Theta(g(x) - a)$ as in the statement of Theorem 3.3 below, and the fact that ρ_t is equivalent to Lebesgue measure along unstable leaves renders plausible that the limit as $a \to 0$ exists, in view of the proof of Theorem 3.3.

with $\sum_{j=1}^{N} \phi_j(y) = 1$ for all $y \in V$, with the support of ϕ_i compactly contained in $\psi_i(U_i)$. Let $\eta : \mathbb{R}^d \to \mathbb{R}_+$ be a C^∞ function supported in $[-1, 1]^d$, with $\int \eta(x)\, dx = 1$. For any small enough $\varepsilon > 0$, put $\eta_\varepsilon(v) = \varepsilon^{-d}\eta(v/\varepsilon)$, and set for $x \in M$

$$(2.3) \qquad \theta_\varepsilon(x) = \mathbb{M}_\varepsilon(\theta) := \sum_{i=1}^{N} \phi_i(x)\Big(\int_{\mathbb{R}^d} \eta_\varepsilon(v)\theta(\psi_i(\psi_i^{-1}(x) - v))\, dv\Big).$$

In the rest of the proof of the theorem, we ignore the charts, slightly abusing notation. Since the ψ_i and ϕ_i are C^∞, and there are finitely many of them, this does not cause any problems (we refer to [8], [7, §5], or [2] for details).

In [8] it is proved that the transfer operator $\mathcal{L}_t$ defined by (2.1) acting on[7] a Banach space $W_{p'}^{u,s}(f_t, V) = W_{p',\dagger}^{u,s}(f_t, V)$ of anisotropic distributions has essential spectral radius strictly smaller than 1, for all real numbers $1 < p' < \infty$ and $u - 2 < s < 0 < u < 2$. In fact, up to further reducing ϵ_0, we can use the same space $W_{p'}^{u,s}(f_0, V)$ (noted $W_{p'}^{u,s}$ from now on, for simplicity) for all $\mathcal{L}_t$ with $|t| < \epsilon_0$. Recall that $\int \mathcal{L}_t(\varphi)\, dm = \int \varphi\, dm$ so that the dual of $\mathcal{L}_t$ fixes Lebesgue measure. By standard arguments (see [17] or [2]), the spectral radius of $\mathcal{L}_t$ on $W_{p'}^{u,s}$ is equal to 1, and 1 is a simple eigenvalue of $\mathcal{L}_t$, for all t. (The spectral properties listed in this paragraph are often referred to as Perron–Frobenius properties.)

It follows from [8, App. A] (see also [2]) that we have the following bounded inclusions for any $r' > u$ and $\widetilde{r} > |s|$:

$$H_{p'}^{r'}(M) \subset W_{p'}^{u,s} \subset H_{p'}^{-\widetilde{r}}(M) \text{ and } H_p^{\widetilde{r}}(M) \subset (W_{p'}^{u,s})^* \subset H_p^{-r'}(M),$$

where $1/p + 1/p' = 1$. (We shall only use the inclusions involving $\widetilde{r}$.) Also, the partial derivatives of an element of $W_{p'}^{u,s}$ belong to $W_{p'}^{u-1,s-1}$.

Denoting by $\widetilde{\rho}_t$ the fixed point of $\mathcal{L}_t$ in $W_{p'}^{u,s}$, it is well-known that the distribution $\varphi \mapsto \widetilde{\rho}_t(\varphi)$ defined for $\varphi \in H_p^{\widetilde{r}}(M)$ extends to a nonnegative Radon measure, which is f_t-invariant. Normalising gives a probability measure, which is in fact the unique SRB measure $d\rho_t$ of f_t. Slightly abusing notation, we shall not distinguish between $\widetilde{\rho}_t(\varphi)$ and $\int \varphi d\rho_t$, in particular we drop the tilde from now on. Note also the fixed point property implies that $\widetilde{\rho}_t$ also belongs to the (smaller) anisotropic space $\widetilde{W}_{p'}^{u,s}$ obtained from $W_{p'}^{u,s}$ by taking a narrower stable cone and a wider unstable cone.

Now, there exists C so that each component of $\operatorname{grad}\rho_t$ has $\widetilde{W}_{p'}^{u-1,s-1}$-norm bounded by $C\|\rho_t\|_{\widetilde{W}_{p'}^{u,s}}$. Since we can take $u \in (1, 1 + \beta)$ and $u - 2 < s < 0$, the transfer operator $\mathcal{L}_t$ acting on $W_{p'}^{u-1,s-1}$ also has essential spectral radius strictly smaller than 1 and enjoys the Perron-Frobenius spectral properties described above. For any $C^{1+\beta}$ function v and any j, since $u - 1 < \beta$ and $\|\partial_j v\|_{C^\beta} \le \|v\|_{C^{1+\beta}}$, we have
(2.4)
$$\|(\partial_j v)\rho_t\|_{W_{p'}^{u-1,s-1}} \le \|v\|_{C^{1+\beta}}\|\rho_t\|_{\widetilde{W}_{p'}^{u-1,s-1}}, \quad \|v(\partial_j\rho_t)\|_{W_{p'}^{u-1,s-1}} \le \|v\|_{C^\beta}\|\partial_j\rho_t\|_{\widetilde{W}_{p'}^{u-1,s-1}},$$

using the Leibniz-type bound obtained from applying the Lasota–Yorke estimates in [8] to the identity map. For any smooth vector field Y, integrating by parts and using that the

---

[7] We recall the construction of the space in Appendix A.

manifold M is boundaryless, we find

$$(2.5) \qquad \int_M \operatorname{div}(Y)\rho_t + \langle \operatorname{grad}\rho_t, Y \rangle = 0 \,,$$

(where $\langle Z, Y \rangle$ denotes the scalar product of vector fields) in the sense of distributions. Hence (recalling that Lebesgue measure is a simple fixed point of each dual operator $\mathcal{L}_t^*$), the resolvent $(1 - \mathcal{L}_t)^{-1}$ is well-defined when acting on $\operatorname{div}(Y)\rho_t + \langle \operatorname{grad}\rho_t, Y \rangle \in W_{p'}^{u-1,s-1}$. Recalling (2.4), the image $(1-\mathcal{L}_t)^{-1}[\operatorname{div}(Y)\rho_t + \langle \operatorname{grad}\rho_t, Y \rangle]$ is a distribution whose $W_{p'}^{u-1,s-1}$-norm is bounded by a constant (depending only on $\|Y\|_{C^{1+\beta}}$) times the $\widetilde{W}_{p'}^{u,s}$-norm of ρ_t.

We may apply the abstract theorem of [16, §8] to the present spaces $W_p^{u,s}$ (although the spaces $W_p^{u,s}$, taken from [8], are not isomorphic to the spaces used in [16]): Mutatis mutandis, we just follow the arguments in [2, §5.3, App. A.3], where [16, §8] was applied to a slightly different anisotropic space introduced in [9]. Translating the statement from [16, Theorem 2.7] gives that $t \mapsto \rho_t$ is differentiable when the values are viewed in $W_{p'}^{u-2,s-2}$. (This would only allow to show Hölder regularity in Theorem 2.1 for $\alpha < 1/2$.) However, we claim that differentiability holds in the stronger norm $W_{p'}^{u'-1,s'-1}$, for $u - 1 < u' < u$ and $s - 1 < s' < s$, arbitrarily close to u and s, respectively. To check this, we follow the argument of [2, Theorem 2.38], bootstrapping from the fact that for $\delta \in (0, 1)$ there exists C so that for all φ

$$(2.6) \qquad \|\mathcal{L}_t\varphi - \mathcal{L}_0\varphi\|_{W_{p'}^{u,s}} \leq C|t|^\delta \|\varphi\|_{W_{p'}^{u+\delta,s+\delta}} \,.$$

The bound (2.6) can be proved by using the mollifiers defined by (2.3): First note that

$$\|\mathcal{L}_t\varphi - \mathcal{L}_0\varphi\|_{W_{p'}^{u,s}} \leq \|\mathcal{L}_t\mathbb{M}_\varepsilon(\varphi) - \mathcal{L}_0\mathbb{M}_\varepsilon(\varphi)\|_{W_{p'}^{u,s}}$$
$$+ \|\mathcal{L}_t(\mathbb{M}_\varepsilon(\varphi) - \varphi)\|_{W_{p'}^{u,s}} + \|\mathcal{L}_0(\mathbb{M}_\varepsilon(\varphi) - \varphi)\|_{W_{p'}^{u,s}}$$
$$\leq C|t|\|\mathbb{M}_\varepsilon(\varphi)\|_{\widehat{W}_{p'}^{u+1,s+1}} + 2C\|\mathbb{M}_\varepsilon(\varphi) - \varphi\|_{\widehat{W}_{p'}^{u,s}} \,,$$

where the (larger) anisotropic space $\widehat{W}_{p'}^{u,s}$ are obtained from $W_{p'}^{u,s}$ by taking a suitable wider stable cone and a suitable narrower unstable cone. Second observe that[8]

$$\|\mathbb{M}_\varepsilon(\varphi)\|_{\widehat{W}_{p'}^{u+1,s+1}} \leq C\varepsilon^{\delta-1}\|\varphi\|_{W_{p'}^{u+\delta,s+\delta}} \text{ and } \|\mathbb{M}_\varepsilon(\varphi) - \varphi\|_{\widehat{W}_{p'}^{u,s}} \leq C\varepsilon^\delta \|\varphi\|_{W_{p'}^{u+\delta,s+\delta}} \,.$$

Finally, taking $\varepsilon = |t|$ proves (2.6).

Next, using that

$$H_p^{1+r}(M) \subset (W_{p'}^{u'-1,s'-1})^* \text{ if } 1 + r > |s' - 1| \,,$$

---

[8]To prove both claims, use $W_{p',\dagger\dagger}^{u,s}(T, V)$ from [8, App. A] and the results therein, since $\|(1+\Delta)^{\delta/2}\varphi\|_{W_{p',\dagger\dagger}^{u,s}}$ is equivalent to $\|\varphi\|_{W_{p',\dagger\dagger}^{u+\delta,s+\delta}}$. The change of cones is needed both to handle the multiplication by ϕ_i in the definition of $\mathbb{M}_\varepsilon$ and the change of charts $\psi_i \circ \psi_j^{-1}$.

the differentiability of $t \mapsto \rho_t \in W_{p'}^{u'-1,s'-1}$ gives that for any $r > 0$ (taking $s' < s < 0$ close enough to zero), any $1 < p' < \infty$, and any function $\varphi \in H_p^{1+r}(M)$, the map

$$t \mapsto \int \varphi \, d\rho_t$$

is differentiable, and its derivative at any $|t_0| < \epsilon_0$ is given by the following fluctuation-dissipation formula

$$(2.7) \qquad \partial_t \left(\int_M \varphi \, d\rho_t \right)|_{t=t_0} = -\left((1 - \mathcal{L}_{t_0})^{-1} [\operatorname{div}(X_{t_0}) \rho_{t_0} + \langle \operatorname{grad} \rho_{t_0}, X_{t_0} \rangle] \right)(\varphi) \,,$$

where the $C^{1+\beta}$ vector field X_{t_0} is defined by

$$(2.8) \qquad \qquad X_{t_0} = (\partial_t f_t)|_{t=t_0} \circ f_{t_0}^{-1} \,.$$

In particular, the derivative $\partial_t (\int_M \varphi \, d\rho_t)|_{t=t_0}$ depends continuously on t_0.

Let $1 < p < \infty$ and $0 < r < 1/p$ be so that $\theta \in H_p^r(M)$. Fix $|s| < r$ small and $\widetilde{r} \in (|s|, r)$ close to $|s|$. Let p' be so that $1/p + 1/p' = 1$.

There exists a constant C (see, e.g., the proof of [7, Lemma 5.3]), so that for any $\varepsilon > 0$

$$(2.9) \qquad \qquad \|\theta_\varepsilon\|_{H_p^{1+\widetilde{r}}} \leq C \|\theta\|_{H_p^r} \varepsilon^{r-1-\widetilde{r}} \,.$$

To simplify notation, assume that $t_0 = 0$. For fixed small $\varepsilon > 0$, decomposing $\theta = \theta_\varepsilon + (\theta - \theta_\varepsilon)$, it suffices to estimate the terms $|\int \theta_\varepsilon (d\rho_t - d\rho_0)|$ and $|\int (\theta_\varepsilon - \theta) d\rho_t| + |\int (\theta_\varepsilon - \theta) d\rho_0|$.

On the one hand, it follows from the linear response formula (2.7), the mean value theorem, and (2.9) that

$$\left| \int \theta_\varepsilon (d\rho_t - d\rho_0) \right| \leq |t| \sup_\tau |((1 - \mathcal{L}_\tau)^{-1} [\operatorname{div}(X_\tau) \rho_\tau + \langle \operatorname{grad} \rho_\tau, X_\tau \rangle])(\theta_\varepsilon)|$$

$$\leq \widetilde{C}_X C' |t| \sup_\tau \|\rho_\tau\|_{W_{p'}^{u,s}} \|\theta_\varepsilon\|_{H_p^{1+\widetilde{r}}} \leq C \cdot C' \cdot \widetilde{C}_X \cdot |t| \sup_\tau \|\rho_\tau\|_{W_{p'}^{u,s}} \|\theta\|_{H_p^r} \varepsilon^{r-1-\widetilde{r}} \,.$$

(We used $\|\theta_\varepsilon\|_{(W_{p'}^{u-1,s-1})^*} \leq C' \|\theta_\varepsilon\|_{H_p^{1+\widetilde{r}}}$ for any $\widetilde{r} > |s|$.)

On the other hand,

$$(2.10) \qquad \left| \int (\theta_\varepsilon - \theta) d\rho_t \right| + \left| \int (\theta_\varepsilon - \theta) d\rho_0 \right| \leq 2 \|\theta_\varepsilon - \theta\|_{H_p^{\widetilde{r}}} \max_{\tau=0,t} \|\rho_\tau\|_{H_{p'}^{-\widetilde{r}}} \,.$$

Since $\|\rho_\tau\|_{H_{p'}^{-\widetilde{r}}} \leq \|\rho_\tau\|_{W_{p'}^{u,s}}$ if $\widetilde{r} > |s|$, we have $\max_{\tau=0,t} \|\rho_\tau\|_{H_{p'}^{-\widetilde{r}}} < \infty$. Since $\widetilde{r} < r < 1/p$, we have

$$(2.11) \qquad \qquad \|\theta_\varepsilon - \theta\|_{H_p^{\widetilde{r}}} \leq C \varepsilon^{r-\widetilde{r}} \|\theta\|_{H_p^r} \,,$$

e.g., by [7, Lemma 5.4]. Choose $\varepsilon = |t|$. For each $0 < \alpha < r$, we can get

$$|t|^{1+r-1-\widetilde{r}} = |t|^{r-\widetilde{r}} \leq |t|^\alpha \,,$$

by taking small enough $\widetilde{r} > |s| > 0$. This proves that $t \mapsto \int \theta \rho_t$ is α-Hölder. $\qquad \square$

3. Linear response for hyperbolic attractors and observables with transversal singularities

As pointed out in Remark 2.2, a necessary condition for our study is that $\int \theta \, d\rho_t$ be well defined for the observable θ. In order to get linear response, the formula (2.7) in the proof of Theorem 2.1 shows that we need the stronger condition that θ belongs to the dual of a space containing the distribution $\langle \mathrm{grad}\rho_t, Y \rangle$ for a smooth vector field Y, in particular containing the components (in the sense of distributions) of $\mathrm{grad}\rho_t$. We may thus hope that observables θ with singularities satisfy this stronger condition if their "wave front set" (WFS) is transversal to the WFS of $\mathrm{grad}\rho_t$. (We only use the notion of a WFS in this informal discussion, and the reader will not need to know its precise definition, which is not exactly the classical one since our objects are C^r for some finite $r > 1$, possibly small with respect to the dimension. It suffices to mention that, in the present setting, the WFS of ρ_t and $\mathrm{grad}\rho_t$ are contained in stable cones.)

Let us be more concrete, considering the anisotropic spaces $W_p^{u,s}$ (containing ρ_t) used in the proof of Theorem 2.1. To fix ideas, we work in dimension two and focus on the toy-model situation of charts (x, y) where the horizontal direction $y = 0$ is expanding and the vertical direction $x = 0$ is contracting. We consider $p = p' = 2$ to simplify matters, and in this informal discussion we shall pretend that the dual of $W_2^{u,s}$ is $W_2^{-u,-s}$. Consider the toy function $\theta(x, y) = \Theta(x)h(y)$, where Θ is the Heaviside function, and h is a C^∞ function with rapid decay (we ignore compact support issues in the x variable in the present heuristic outline). This is clearly the simplest transversal toy function. In Fourier space, with parameters (ξ, η), we find $\hat{\Theta}(\xi)\hat{h}(\eta)$, where the decay of $\hat{\Theta}$ at infinity is $\sim |\xi|^{-1}$ and for any $N \geq 1$ the decay of $\hat{h}(\eta)$ at infinity is $\sim C_N|\eta|^{-N}$. It is thus reasonable to hope that a function of the type of $\theta(x, y)$ belongs not only to the dual $W_2^{-u,-s}$ of a space $W_2^{u,s}$ for some $s < 0 < u$, but also to the dual $W_2^{-(u-1),-(s-1)}$ of a space $W_2^{u-1,s-1}$. Indeed , in any fixed "stable" cone $|\xi| \leq \tilde{c}|\eta|$, we can get the required decay $C(1 + |\xi|^2 + |\eta|^2)^{(s-1)/2}$ (with $s < 0$ so that $s - 1 < -1$) by choosing N large enough depending on $\tilde{c}$. The condition in the "unstable" cone is satisfied if $(u - 1 + 1)2 > 2$, that is, if $u > 1$. (We would like to emphasize that, although we have just argued that functions of the type $\theta(x, y)$ may lie in anisotropic spaces $W_p^{u,s}$, they are in general not bounded multipliers in these spaces, see [3].)

Our second main theorem below (Theorem 3.3) will justify the optimism induced by the above toy-model computation.

To formulate this second result, we first recall the notion of admissible stable leaves from Gouëzel and Liverani [16, §3]. Let $t \mapsto f_t$, for $t \in [-\epsilon_0, \epsilon_0]$, be a C^3 family of C^4 diffeomorphisms f_t on a smooth d-dimensional Riemann manifold M, with a transitive compact hyperbolic attractor $\Lambda_t \subset M$ (as defined in the beginning of Section 2), and let V be a common attracting neighbourhood for all the f_t (up to reducing ϵ_0). We may and do assume that $\overline{V}$ is compact. Using an adapted Mather metric and further reducing ϵ_0 if necessary, we can assume that there exists $0 < \nu < 1$ so that for all small enough $|t|$, the expansion of $D_x f_t$ along E_x^u is stronger than $\nu^{-1} > 1$, while its contraction along E_x^s is stronger than $\nu < 1$, and the angle between E_x^s and E_x^u is everywhere arbitrarily close to $\pi/2$. For small enough $\kappa > 0$, we define the stable cone at $x \in V$ by

$$\mathcal{C}^s(x) = \{w + v \in T_x M \mid w \in E^s(x), v \perp E^s(x), \|v\| \leq \kappa\|w\|\} \, .$$

If $\kappa > 0$ is small enough then $D_x f_0^{-1}(\mathcal{C}^s(x) \backslash \{0\})$ belongs to the interior of $\mathcal{C}^s(f_0^{-1}(x))$, and $D_x f_0^{-1}$ expands the vectors in $\mathcal{C}^s(x)$ by ν^{-1}.

Definition 3.1 (Admissible charts). *There exist real numbers $r_i \in (0,1)$ and C^4 coordinate charts $\psi_1, \ldots, \psi_N$, with ψ_i defined on $(-r_i, r_i)^d \subset \mathbb{R}^d$ (with its Euclidean norm), such that the attracting neighbourhood V is covered by the open sets*

$$\left(\psi_i((-r_i/2, r_i/2)^d) \right)_{i=1\ldots N},$$

and the following conditions hold: $D\psi_i(0)$ is an isometry, $D\psi_i(0) \cdot \left(\mathbb{R}^{d_s} \times \{0\} \right) = E^s(\psi_i(0))$, and the C^4-norms of ψ_i and its inverse are bounded by $1 + \kappa$ (with κ as introduced above).

Next, we may pick $c_i \in (\kappa, 2\kappa)$ such that the corresponding stable cone in charts

$$\mathcal{C}_i^s = \{ w + v \in \mathbb{R}^d \mid w \in \mathbb{R}^{d_s} \times \{0\}, v \in \{0\} \times \mathbb{R}^{d_u}, \|v\| \leq c_i \|w\| \}$$

satisfies $D_x \psi_i(\mathcal{C}_i^s) \supset \mathcal{C}^s(\psi_i(x))$ and $D_{\psi_i(x)} f_t^{-1}(D_x \psi_i(x) \mathcal{C}_i^s) \subset \mathcal{C}^s(f_t^{-1}(\psi_i(x)))$ for any $x \in (-r_i, r_i)^d$ and all $|t| < \epsilon_0$ (reducing ϵ_0 if necessary).

Let $G_i(K)$ be the set of graphs of C^4 maps $\chi : U_\chi \to (-r_i, r_i)^{d_u}$ defined on a subset U_χ of $(-r_i, r_i)^{d_s}$, with

$$|D\chi| < c_i \text{ and } |\chi|_{C^4} \leq K,$$

where a function defined on a subset U of $\mathbb{R}^{d_s}$ is C^4 if it admits a C^4 extension to an open neighbourhood of U, the norm being the infimum of the norms of such extensions. In particular, the tangent space to the graph of χ belongs to the interior of the cone $\mathcal{C}_i^s$.

Uniform hyperbolicity of f_0 implies (see [16, Lemma 3.1]) that if K is large enough, then there exists $K' < K$ such that, for any $W \in G_i(K)$ and any $1 \leq j \leq N$, the set $\psi_j^{-1}(f_t^{-1}(\psi_i(W)))$ belongs to $G_j(K')$ for all $|t| < \epsilon_0$ (reducing again ϵ_0 if needed).

Definition 3.2 (Admissible graphs and admissible stable leaves). *An admissible graph is a map χ defined on a ball*

$$\bar{B}(w, K_1 \delta) \subset (-2r_i/3, 2r_i/3)^{d_s}$$

for some small enough $\delta > 0$ and large[9] enough K_1, taking its values in $(-2r_i/3, 2r_i/3)^{d_u}$ with $\mathrm{range}(\mathrm{Id}, \chi) \in G_i(K)$. The set of admissible stable leaves is

$$\left\{ W := \psi_i \circ (\mathrm{Id}, \chi)(\bar{B}(w, \delta)) \mid \chi : \bar{B}(w, K_1 \delta) \to \mathbb{R}^{d_u} \right.$$

$$\left. \text{is an admissible graph on } B_i := (-2r_i/3, 2r_i/3)^{d_s} \right\}.$$

We may now state our second theorem, which exhibits a simple transversality condition which is sufficient to ensure linear response for hyperbolic attractors:

Theorem 3.3 (Linear response for transversal Heaviside singularities). *Let $t \mapsto f_t$, for $t \in [-\epsilon_0, \epsilon_0]$, be a C^3 family of C^4 diffeomorphisms f_t on a smooth Riemann manifold M, with a transitive compact hyperbolic attractor $\Lambda_t \subset M$. Let ρ_t be the (unique) SRB measure of f_t on Λ_t. Let*

$$\theta(x) = h(x)\Theta(g(x) - a),$$

---

[9]If $\kappa > 0$ chosen above is small enough, then $\nu(1+\kappa)^2 \sqrt{1+4\kappa^2} < 1$. Let $K_1 > 1$ be such that $K_1 > 1 + K_1 \nu_1 (1+\kappa)^2 \sqrt{1+4\kappa^2}$, and take $\delta > 0$ so that $6K_1 \delta < \min_i(r_i)$, see [16].

where $\Theta : \mathbb{R} \to \mathbb{R}$ *is the Heaviside function,* $h : M \to \mathbb{C}$ *is* C^3 *and* $g : M \to \mathbb{R}$ *is* C^4, *with* $a \in \mathbb{R}$ *not a critical value of* g, *and so that*[10] *the intersection*

$$\mathcal{W}_a := \{x \mid g(x) = a\} \cap \mathrm{supp}(h)$$

admits a C^4 *foliation by admissible stable leaves. Then there exists* $\epsilon_1 \in (0, \epsilon_0]$ *so that for any* $|t_0| < \epsilon_1$ *the map* $t \mapsto \int_M \theta \, d\rho_t$ *is differentiable at* t_0, *with (recalling (2.1) and (2.8))*

$$(3.1) \qquad \partial_t(\int_M \theta \, d\rho_t)|_{t=t_0} = -\big((1 - \mathcal{L}_{t_0})^{-1}[\mathrm{div}(X_{t_0})\rho_{t_0} + \langle \mathrm{grad}\rho_{t_0}, X_{t_0}\rangle]\big)(\theta) \,,$$

where the gradient is in the sense of distributions.

In addition, if (f_{t_0}, ρ_{t_0}) *is*[11] *mixing, we have the exponentially convergent sums*

$$(3.2) \qquad \partial_t(\int_M \theta \, d\rho_t)|_{t=t_0} = -\sum_{k=0}^{\infty} \big(\mathcal{L}_{t_0}^k[\mathrm{div}(X_{t_0})\rho_{t_0} + \langle \mathrm{grad}\rho_{t_0}, X_{t_0}\rangle]\big)(\theta)$$

$$(3.3) \qquad\qquad = \sum_{k=0}^{\infty} \rho_{t_0}\big(\langle \mathrm{grad}(\theta \circ f_{t_0}^k), X_{t_0}\rangle\big) \,.$$

If the stable dimension d_s is equal to $d - 1$ (for example, if $d = 2$ with $d_s = d_u = 1$) then the assumption reduces to requiring that $\mathcal{W}_a = \{x \mid g(x) = a\} \cap \mathrm{supp}(h)$ is an admissible stable leaf.

More general transversal test functions can be considered: Our proof of Theorem 3.3 actually shows that any function θ lying in the dual to a Gouëzel–Liverani space $\mathcal{B}^{0,|s-2|}$ (the corresponding norm is recalled in (3.4)) with $s < 0$ satisfies the conclusion of the theorem. Since our goal here is to highlight a sufficient condition which works for a natural class of test functions, we refrain from a more general statement in order to keep the proof as free as possible from technicalities. (In any case, the regularity conditions in the blueprint theorem above are probably not optimal, in particular we expect that lying in the dual to a space $\mathcal{U}_p^{0,|s-1|}$ with $s < 0$ from [3] should suffice.)

Remark 3.4 (Potential generalisations involving the Dirac mass). We expect that there is an analogue of Theorem 3.3 when $h(x)\Theta(g(x)-a)$ is replaced by the distribution $h(x)\delta_D(g(x)-a)$ if a is not critical and $\mathcal{W}_a$ is foliated by admissible stable leaves, up to increasing the smoothness requirement. (See also [16, Prop 4.4] for an occurrence of the Dirac mass in the anisotropic setting.)

Proof of Theorem 3.3. To prove Theorem 3.3, it is convenient to use the anisotropic Banach spaces from [16, 17], based on taking suprema over admissible stable leaves (instead of working with cones in Fourier space as we did with the spaces from [8] used in the proof of Theorem 2.1).

Since each f_t is C^4, the transfer operator $\mathcal{L}_t$ has a spectral gap on the anisotropic spaces $\mathcal{B}^{u,|s|}$ of [16] for integer $u \geq 1$ and real $s < 0$ with $u + |s| < 3 = 4 - 1$. Using admissible stable leaves and the notations introduced from Definition 3.2, we recall the definition of the norm of $\mathcal{B}^{u,|s|}$ in coordinates (see [16, Lemma 3.2]). Let $u \in \{1, 2\}$, let $-3 + u < s < 0$ be

---

[10]This assumption implies the announced transversality condition: The normal to $\mathcal{W}_a$ at each point lies in the unstable cone.

[11]As usual, we can reduce to the mixing case by spectral decomposition (see e.g. [19, Theorem 18.3.1]).

a real number, and set (the test function ω below is compactly supported in $\bar{B}(w,\delta)$, the measure dm is Lebesgue measure on $\mathbb{R}^{d_s}$, and χ ranges over admissible graphs on B_i)

$$(3.4) \qquad \|\varphi\|_{u,s} = \max_{\substack{0 \le u' \le u \\ u' \in \mathbb{Z}}} \max_{\substack{|\alpha|=u' \\ 1 \le i \le N}} \sup_{\chi} \sup_{|\omega|_{C^{|s|+u'}} \le 1} \int_{B(w,\delta)} [\partial^\alpha(\varphi \circ \psi_i)] \circ (\mathrm{Id}, \chi) \cdot \omega \, dm \,.$$

The space $\mathcal{B}^{u,|s|}$ is then the closure of C^3 functions for the above norm.

We need further preparations: We may and shall focus on a neighbourhood of $\mathcal{W}_a$, since the function θ is C^2 elsewhere. Since a is not a critical value, $\mathcal{W}_a$ is a codimension-one manifold by the implicit function theorem. Our assumption guarantees that this manifold has a C^4 foliation by d_s-dimensional leaves which are stable admissible leaves. The idea is to foliate a chart neighbourhood by stable leaves, e.g. like in the proof of [5, Lemma 2.11]. More precisely, we consider the C^4 foliation of $\mathcal{W}_a$ by stable leaves (which exists by assumption), noting $W_a(z)$ the leaf through $z \in \mathcal{W}_a$. Fixing small balls around finitely many points $x_j \in \mathcal{W}_a$, we may use C^4 charts $\widetilde{\psi}_j$ satisfying the assumptions in Definition 3.1 and, in addition, mapping $\mathbb{R}^{d_s} \times \{\vec{0}\}$ to $W_a(x_j)$, while mapping the codimension-one hyperspace $\mathbb{R}^{d-1} \times \{0\}$ to $\mathcal{W}_a$. (Recall also that the d-th axis $\{0\} \times \mathbb{R}$ is transversal to the stable cone $\mathcal{C}_j^s$.) It will be convenient to require the charts to satisfy the following stronger assumption: First notice that since a is not a critical value of g, there is a small neighbourhood I_a of a so that any $b \in I_a$ is not a critical value of g. In addition, $\mathcal{W}_b$ is also foliated by stable leaves. In fact, we may ensure that there exist C^3 local diffeomorphisms $\tau_j : \mathbb{R} \to \mathbb{R}$ with $\tau_j(0) = 0$, $\tau_j'(0) = 1$, and $\tau_j''(0) = 0$, so that for any $b \in I_a$, the chart $\widetilde{\psi}_j$ maps $\mathbb{R}^{d_s} \times \{\tau_j(a-b)\}$ to $W_b(x_j)$ and maps the codimension-one hyperspace $\mathbb{R}^{d-1} \times \{\tau_j(a-b)\}$ to $\mathcal{W}_b$. We let $\widetilde{\phi}_j$ be an adapted partition of unity.

We now move on to the proof. Assume without loss of generality that $t_0 = 0$. Set

$$(3.5) \qquad \gamma := (1 - \mathcal{L}_0)^{-1}[\mathrm{div}(X_0)\rho_0 + \langle \mathrm{grad}\rho_0, X_0 \rangle] \,.$$

Since the results of [16] imply that $\rho_t \in \mathcal{B}^{u,|s|}$ for $u = 2$ and $-1 < s < 0$, and since X_0 is a C^2 vector field, we have $\gamma \in \mathcal{B}^{u-1,|s-1|} \subset \mathcal{B}^{u-2,|s-2|}$. (The Leibniz inequality is not difficult here since each u' in (3.4) is an integer, in particular, we do not need different sets of "cones" as in the proof of Theorem 2.1.)

We can apply [16, Theorem 2.7, §8] (as we are using the same spaces here): The result there says that the map $t \mapsto \rho_t \in \mathcal{B}^{u-2,|s-2|}$ is differentiable and, for any φ in the dual space to $\mathcal{B}^{u-2,|s-2|}$ we have (recalling the notation (3.5))

$$(3.6) \qquad \partial_t \int \varphi \, d\rho_t|_{t=0} = -\gamma(\varphi) = -\big((1 - \mathcal{L}_0)^{-1}[\mathrm{div}(X_0)\rho_0 + \langle \mathrm{grad}\rho_0, X_0 \rangle]\big)(\varphi) \,.$$

(Contrary to the situation in the proof of Theorem 2.1, it is not clear that we can take φ in the dual space to $\mathcal{B}^{u-2,|s'-1|}$ for $s - 1 < s' < s < 0$ in (3.6).)

To prove the first claim of the theorem, it thus suffices to show that the function $\theta(x) = h(x)\Theta(g(x) - a)$ belongs to the dual space to $\mathcal{B}^{u-2,|s-2|}$. (Recall that $u = 2$ and $-1 < s < 0$.) The special properties of the charts $\widetilde{\psi}_j$ derived from our transversality assumption will be essential here. Let $\widetilde{\gamma}$ be a C^3 function on M supported in a small ball around x_j and put

$$h_j = |\det D\widetilde{\psi}_j|(\widetilde{\phi}_j \cdot h) \circ \widetilde{\psi}_j \,.$$

Then the contribution of $\int \widetilde{\gamma}\theta \, dm = \int \widetilde{\gamma}(y)h(y)\Theta(g(y) - a) \, dm$ to the jth chart is

$$\int_{\mathbb{R}^d} |\det D\widetilde{\psi}_j|(\widetilde{\phi}_j\theta\widetilde{\gamma}) \circ \widetilde{\psi}_j \, dx$$

$$= \int_{[-1,1]^{d_u}} \Theta(x_d)dx_{d_s+1}\ldots dx_d \int_{B_j} dx_1 \ldots dx_{d_s} h_j(x)\widetilde{\gamma}(\widetilde{\psi}_j(x))$$

$$= \int_{[0,1]} dx_d \int_{[-1,1]^{d_u-1}} dx_{d_s+1}\ldots dx_{d-1} \int_{B_j} dx_1 \ldots dx_{d_s} h_j(x)\widetilde{\gamma}(\widetilde{\psi}_j(x)).$$

Obviously,

$$|\int_{[0,1]} dx_d \int_{[-1,1]^{d_u-1}} dx_{d_s+1}\ldots dx_{d-1} \int_{B_j} dx_1 \ldots dx_{d_s} h_j(x)\widetilde{\gamma}(\widetilde{\psi}_j(x))|$$

$$\leq 2^{d_u-1} \sup_{x_{d_s+1},\ldots,x_d} |\int_{B_j} dx_1 \ldots dx_{d_s} h_j(x)\widetilde{\gamma}(\widetilde{\psi}_j(x))|.$$

Since h_j is C^3, while $s - 2 > -3$ and $u = 2$, taking $u' = 0$ in the definition (3.4) of the norm, the integral over $B_j \subset [-1,1]^{d_s}$ in the right-hand side is bounded by $\|h_j\|_{C^2}\|\widetilde{\gamma}\|_{u-2,s-2}$. In particular,[12] we can let $\widetilde{\gamma}$ tend to $\widetilde{\phi}_j \cdot \gamma \in \mathcal{B}^{u-2,|s-2|}$. We have proved that θ lies in the dual space to $\mathcal{B}^{u-2,|s-2|}$.

It remains to show (3.2)–(3.3) in the mixing case. For this, first notice that (see [17]) mixing implies that 1 is the only eigenvalue of $\mathcal{L}_0$ on the unit circle so that, using (2.5), the distribution $(1 - \mathcal{L}_0)^{-1}(\gamma)$ may be replaced by the exponentially convergent — in the Banach norm — sum $\sum_{k=0}^{\infty} \mathcal{L}_0^k(\gamma)$. Second, note that $(\mathcal{L}_0^k(\gamma))(\theta) = \gamma(\theta \circ f_0^k)$. Finally, integrate by parts, in the sense of distributions, noting that the derivative of θ is not only a distribution of order zero but also an element of the dual of $\mathcal{B}^{u,|s|}$. This can be checked by using the ideas developed above, and is left to the reader. □

APPENDIX A. DEFINITION OF THE ANISOTROPIC SPACES $W_p^{u,s}$ FROM [8].

For the reader's convenience, we recall the definition of the space $W_p^{u,s}(T, V) = W_{p,\dagger}^{u,s}(T, V)$ from [8, App. A] associated to a C^κ diffeomorphism T with a transitive compact hyperbolic attractor Λ with an attracting neighbourhood V, for $\kappa > 1$. (Formally the paper [8] only considers the Anosov case when $\Lambda = M$, but, since E^u and E^s can be extended to $\overline{V}$, in the sense of [9], this does not cause any problems.)

A cone in $\mathbb{R}^d$ is a subset which is invariant under scalar multiplication. We write $\mathbf{C} \Subset \mathbf{C}'$ for two cones $\mathbf{C}$ and $\mathbf{C}'$ in $\mathbb{R}^d$ if $\overline{\mathbf{C}} \subset$ interior $(\mathbf{C}') \cup \{0\}$. A cone $\mathbf{C}$ is called d'-dimensional if $d' \geq 1$ is the maximal dimension of a linear subset of $\mathbf{C}$.

Definition A.1 (Cone pairs and cone systems). *A cone pair is* $\mathbf{C}_\pm = (\mathbf{C}_+, \mathbf{C}_-)$, *where* $\mathbf{C}_+$ *and* $\mathbf{C}_-$ *are closed cones in* $\mathbb{R}^d$, *with nonempty interiors, of respective dimensions* d_u *and* d_s *and so that* $\mathbf{C}_+ \cap \mathbf{C}_- = \{0\}$. *A cone system is*

$$\Theta = (\mathbf{C}_\pm, \varphi_+, \varphi_-),$$

---

[12]In practice, we can take the sequence corresponding to $(1 - \mathcal{L}_0)^{-1}[\text{div}X_0\mathcal{L}_0^n(1_V) + \langle X_0, \text{grad}(\mathcal{L}_0^n(1_V))\rangle]$.

with $\mathbf{C}_\pm = (\mathbf{C}_+, \mathbf{C}_-)$ a cone pair and $\varphi_\pm : \mathbf{S}^{d-1} \to [0,1]$ two C^∞ functions on the unit sphere $\mathbf{S}^{d-1}$ in $\mathbb{R}^d$ satisfying

$$\varphi_+(\xi) = 1 \ \text{if} \ \xi \in \mathbf{S}^{d-1} \cap \mathbf{C}_+, \quad \varphi_+(\xi) = 0 \ \text{if} \ \xi \in \mathbf{S}^{d-1} \cap \mathbf{C}_-, \quad \varphi_- = 1 - \varphi_+ \,.$$

Next, introduce for real numbers u and s the functions

$$\Psi_{u,\Theta_+}(\xi) = (1 + \|\xi\|^2)^{u/2} \varphi_+ \left(\frac{\xi}{\|\xi\|} \right) \quad \text{and} \quad \Psi_{s,\Theta_-}(\xi) = (1 + \|\xi\|^2)^{s/2} \varphi_- \left(\frac{\xi}{\|\xi\|} \right) .$$

For a cone system Θ and $\varphi \in C_0^\infty(\mathbb{R}^d)$, we define norms for $1 < p < \infty$ and $s \le 0 \le u$ by

$$(\text{A.1}) \qquad \|\varphi\|_{W_p^{\Theta,u,s}} \;=\; \|\Psi_{u,\Theta_+}^{Op}(\varphi)\|_{L_p} + \|\Psi_{s,\Theta_-}^{Op}(\varphi)\|_{L_p} \,,$$

where for a function $\Psi : \mathbb{R}^d \to \mathbb{R}$ the operator Ψ^{Op} is defined by the convolution

$$\Psi^{Op}(\varphi) = (\mathbf{F}^{-1}\Psi) * \varphi \,,$$

where $\mathbf{F}^{-1}$ denotes the inverse Fourier transform

$$\mathbf{F}^{-1}(\Psi)(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ix\xi} \Psi(\xi) \, d\xi, \quad x \in \mathbb{R}^d \,.$$

Definition A.2 (Admissible charts and partition of unity). *Admissible charts and partition of unity for T are: A finite system of C^∞ local charts $\{(V_i, \psi_i)\}_{1 \le N}$, with open subsets $V_i \subset M$, and diffeomorphisms $\psi_i : U_i \to V_i$ such that $\overline{V} \subset \cup_i V_i$, and U_i is a bounded open subset of $\mathbb{R}^d$ for each $1 \le i \le N$, together with a finite C^∞ partition of unity $\{\phi_i\}$ for M, subordinate to the cover $\mathcal{V} = \{V_i\}$ of $\overline{V}$.*

Definition A.3 (Admissible cone pairs). *Since Λ is a hyperbolic attractor for T we may choose C^∞ admissible local charts $\{(V_i, \psi_i)\}_{1 \le i \le N}$ and cone pairs $\{\mathbf{C}_{i,\pm} = (\mathbf{C}_{i,+}, \mathbf{C}_{i,-})\}_{1 \le i \le N}$, so that the following conditions hold[13]:*

- *If $x \in V_i$, the cone $(D\psi_i^{-1})_x^*(\mathbf{C}_{i,+})$ contains the (d_u-dimensional) normal subspace of $E^s(x)$, and the cone $(D\psi_i^{-1})_x^*(\mathbf{C}_{i,-})$ contains the (d_s-dimensional) normal subspace of $E^u(x)$.*
- *If $T(V_i) \cap V_{i'} \ne \emptyset$, then the C^κ map*

$$F = \psi_i^{-1} \circ T^{-1} \circ \psi_{i'} : \psi_{i'}^{-1}(T(V_i) \cap V_{i'}) \to U_i \,,$$

  extends to a bilipschitz C^1 diffeomorphism of $\mathbb{R}^d$ so that, using A^{tr} to denote the transposition of a matrix A,

$$DF_x^{tr}(\mathbb{R}^d \setminus \mathbf{C}_{i,+}) \Subset \mathbf{C}_{i',-} \,, \qquad \forall x \in \mathbb{R}^d \,.$$

Definition A.4 (The spaces $W_p^{u,s}(T,V)$). *Fix a finite set of admissible charts, a partition of unity, and admissible cone pairs, and choose a cone system for each cone pair. Then for any $1 < p < \infty$ and $u - (\kappa - 1) < s \le 0 \le u$, the Banach space $W_p^{u,s}(T,V)$ is the completion of $C^\infty(\overline{V})$ for the norm*

$$(\text{A.2}) \qquad \qquad \|\varphi\|_{W_p^{u,s}} := \max_{1 \le i \le N} \|(\phi_i \cdot \varphi) \circ \psi_i^{-1}\|_{W_p^{\Theta_i, u, s}} \,,$$

---

[13]$\mathbf{C}_{i,\pm}$ are locally constant cone fields in the cotangent bundle $T^*\mathbb{R}^d$, so that the conditions are expressed with respect to normal subspaces.

recalling (A.1) *(We do not claim that choosing different charts, partitions of unity, cone pairs, or cone systems would lead to equivalent norms.)*

References

[1] V. Baladi, *Linear response, or else*, in *ICM Seoul 2014, Proceedings, Volume III* (2014) 525–545.

[2] V. Baladi, *Dynamical Zeta Functions and Dynamical Determinants for Hyperbolic Maps,* Monograph (2016). `https://webusers.imj-prg.fr/~viviane.baladi/baladi-zeta2016.pdf`

[3] V. Baladi, The quest for the ultimate anisotropic Banach space, `arXiv:1607.00654` To appear J. Stat. Phys.

[4] V. Baladi, M. Benedicks, and D. Schnellmann, *Whitney–Hölder continuity of the SRB measure for transversal families of smooth unimodal maps,* Invent. Math. **201** (2015) 773–844.

[5] V. Baladi, M.F. Demers, and C. Liverani, *Exponential decay of correlations for finite horizon Sinai billiard flows,* `arXiv:1506.02836`.

[6] V. Baladi and S. Gouëzel, *Good Banach spaces for piecewise hyperbolic maps via interpolation,* Ann. Inst. Henri Poincaré, Anal. non. lin. **26** (2009) 1453–1481.

[7] V. Baladi and C. Liverani, *Exponential decay of correlations for piecewise contact hyperbolic flows,* Comm. Math. Phys. **314** (2012) 689–773.

[8] V. Baladi and M. Tsujii, *Anisotropic Hölder and Sobolev spaces for hyperbolic diffeomorphisms,* Ann. Institut Fourier **57** (2007) 127–154.

[9] V. Baladi and M. Tsujii, *Dynamical determinants and spectrum for hyperbolic diffeomorphisms,* in K. Burns, D. Dolgopyat, Ya. Pesin (eds) Probabilistic and Geometric Structures in Dynamics, pp. 29–68, Contemp. Math., **469**, Amer. Math. Soc., Providence, RI (2008).

[10] B. Cessac and J.A. Sepulchre, *Linear Response in a class of simple systems far from equilibrium,* Physica D **224** (2007) 13–28.

[11] M. Demers and C. Liverani, *Stability of statistical properties in two-dimensional piecewise hyperbolic maps,* Trans. Amer. Math. Soc. **360** (2008), 4777–4814.

[12] G. L. Eyink, Y.W.N. Haine, and D.J. Lee, *Ruelle's linear response formula, ensemble adjoint schemes and Lévy flights,* Nonlinearity **16** (2004) 1867–1889.

[13] G. Gallavotti and E.G.D. Cohen, *Dynamical ensembles in nonequilibrium statistical mechanics,* Phys. Rev. Lett. **74** (1995) 2694–2697.

[14] P. Giulietti, C. Liverani, and M. Pollicott, *Anosov flows and dynamical zeta functions,* Ann. of Math. **178** (2013) 687–773.

[15] G. A. Gottwald, J. P. Wormell, and J. Wouters, *On spurious detection of linear response and misuse of the fluctuation-dissipation theorem in finite time series*, Phys. D **331** (2016) 89–101.

[16] S. Gouëzel and C. Liverani, *Banach spaces adapted to Anosov systems,* Ergod. Th. Dynam. Sys. **26** (2006) 189–218.

[17] S. Gouëzel and C. Liverani, *Compact locally maximal hyperbolic sets for smooth maps: fine statistical properties,* J. Diff. Geom. **79** (2008) 433–477.

[18] M. Jiang, *Differentiating potential functions of SRB measures on hyperbolic attractors,* Ergod. Th. Dynam. Sys. **32** (2012) 1350–1369.

[19] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems,* Cambridge University Press, Cambridge (1995).

[20] R. Kubo, *The fluctuation-dissipation theorem,* Rep. Prog. Phys. **29** (1966) 255–284.

[21] V. Lucarini, F. Lunkeit, and F. Ragone, *Predicting climate change using response theory: global averages and spatial patterns,* J. Stat. Phys. doi: 10.1007/s10955-016-1506-z (2016).

[22] V. Lucarini, D. Faranda, A.C.M. Freitas, J.M. Freitas, M. Holland, T. Kuna, M. Nicol. M. Todd, and S. Vaienti, *Extremes and Recurrence in Dynamical Systems*, Wiley, New York (2016).

[23] V. Lucarini, D. Faranda, and J. Wouters, *Universal behaviour of extreme value statistics for selected observables of dynamical systems,* J. Stat. Phys. **152** (2012) 63–73.

[24] V. Lucarini, D. Faranda, J. Wouters, and T. Kuna, *Towards a general theory of extremes for observables of chaotic dynamical systems,* J. Stat. Phys. **154** (2014) 723–750.

[25] V. Lucarini and S. Sarno, *A statistical mechanical approach for the computation of the climatic response to general forcings,* Nonlin. Proc. Geophys. **18** (2011) 7–28.

[26] U. Marini Betttolo Marconi, A. Puglisi, A. Rondoni, and A. Vulpiani, *Fluctuation-dissipation: Response theory in statistical physics,* Phys. Rep. **461** (2008) 111–195.

[27] C.H. Reick, *Linear response of the Lorenz system,* Phys. Rev. E **66** (2002) 036103.

[28] D. Ruelle, *General linear response formula in statistical mechanics, and the fluctuation-dissipation theorem far from equilibrium*, Phys. Letters A **245** (1998) 220–224.

[29] D. Ruelle, *Differentiation of SRB states,* Comm. Math. Phys. **187** (1997) 227–241. *Correction and complements: "Differentiation of SRB states"* Comm. Math. Phys. **234** (2003) 185–190.

[30] D. Ruelle, *A review of linear response theory for general differentiable dynamical systems*, Nonlinearity **22** (2009) 855–870.

[31] R. Strichartz, *Multipliers on fractional Sobolev spaces,* J. Math. Mech. **16** (1967) 1031–1060.

Sorbonne Universités, UPMC Univ. Paris 6, CNRS, Institut de Mathématiques de Jussieu, (IMJ-PRG), 4, Place Jussieu, 75005 Paris, France
E-mail address: `viviane.baladi@imj-prg.fr`

Department of Mathematics and Statistics, University of Reading, RG66AX, Reading, UK
E-mail address: `t.kuna@reading.ac.uk`

Department of Mathematics and Statistics, University of Reading, RG66AX, Reading, UK, CEN, University of Hamburg, Hamburg 20144, Germany
E-mail address: `v.lucarini@reading.ac.uk`