Rigorous Diffusion Properties for the Sawtooth Map

J. Bellissard* and S. Vaienti**

Centre de Physique Theorique, C.N.R.S. Luminy, Case 907, F-13288 Marseille Cedex 9, France

Abstract. We get a rigorous bound for the diffusion constant of the hamiltonian dynamical system generated by a sawtooth map on a cylinder. The momentum variable properly renormalized then behaves almost like a brownian motion in the limit of infinite coupling constants. The strategy of the proof is a rigorous reformulation of the Random Phase Approximation.

0. Introduction

In this paper we consider the area-preserving sawtooth map:

$$A_{n+1} = A_n + Kg(\theta_n),$$
$$\theta_{n+1} = \theta_n + A_{n+1} \,(\text{mod } 2\pi),$$

(0.1)

where $g(\theta)$ is a 2π-periodic piecewise continuous function of zero average. This map is a simple model for a certain number of physical situations: charged particles in magnetic fields, plasma confinement in nuclear fusion, stochastic ionization, etc. . . . In fact it describes a "kicked" rotor subject to a sequence of periodic impulses $g(\theta)$, or alternatively, since it can be written as a second order difference equation:

$$\theta_{n+1} - 2\theta_n + \theta_{n-1} = Kg(\theta_n)(\text{mod } 2\pi)$$

(0.2)

it describes the motion of a particle receiving an impulse determined by a periodic one-dimensional potential.

   A lot of analytical and numerical work has been devoted to the study of (0.1) for smooth g's, the most important example being $g(\theta) = \sin\theta$, namely the Chirikov–Taylor "standard map" [1, 2]. While at small values of K the dynamics is regular (stability of KAM tori [3–8]), it has been observed since a long time ago that for K large, the solutions of (0.1) admit a diffusive behavior in the phase space $(A, \theta) \in \mathbb{R} \times \mathbb{T}$. The standard argument in this respect is the so-called *random phase*

approximation (RPA) [1, 2]. If we define ΔA_n by $\Delta A_n(A, \theta) = A_n(A, \theta) - A$ we get a 2π-periodic function of A, that will be considered as a stochastic process if (A, θ) is considered as a random variable on the space $\mathbb{T}^2$ with probability given by the normalized Haar measure μ_L. Neglecting the *two-point correlation function* $\langle g(\theta)g(\theta_n(A, \theta)) \rangle$, $n > 0$, where $\langle \cdot \rangle$ is the expectation on $\mathbb{T}^2$, the variance of ΔA_n is:

$$\langle (\Delta A_n)^2 \rangle = K^2 n \langle g^2 \rangle. \tag{0.3}$$

The intuitive idea is that, for $n > 0$, the phases $\theta_n = \theta_n(A, \theta)$ oscillate so fast at large K's, that they become uncorrelated from θ. Therefore, the evolution of ΔA is diffusive and the diffusion coefficient D, defined by:

$$D = \lim_{n \to \infty} \frac{\langle (\Delta A_n)^2 \rangle}{\langle g^2 \rangle n}, \tag{0.4}$$

(provided the limit exists), is approximately equal to K^2, for K large. This last limit is called the *quasi-linear diffusion coefficient* and is denoted by D_L. Further calculations [1, 2, 23–27] using the formal infinite expansion for D in force–force correlations, give an higher-order correction of the type:

$$D \sim D_L[\tfrac{1}{2} - J_2(K) + J_2^2(K)], \tag{0.5}$$

where J_2 is the Bessel function of order 2. Amazingly enough, very little is known mathematically on the existence of D and the validity of (0.4) and (0.5). The main difficulty is the non-ergodicity of the motion due to the possible existence of islands of stability, so that the usual methods are useless to prove the convergence of the expansion involved in (0.5).

   Our aim in this paper is to take seriously the RPA method and to derive rigorous estimates on the diffusion coefficient. The argument leading to (0.3) is actually much too simple. A certain number of additional steps must be investigated in order to estimate how fast the correlation function converges to zero at large K's. To illustrate these ideas, we will consider the sawtooth map. This is because this latter model avoids complicated and probably inessential combinatorics. We must note at this point that if K is a positive integer, this map becomes a linear automorphism of the torus and the diffusion coefficient can easily be computed using analytic methods ([10–11, 16–18]). However for $K \notin \mathbb{Z}$ the situation is quite different. While it is possible to prove the ergodicity of the map (and the proof is far from being trivial [15]), the exact computation of D is still unsolved; approximate estimates can be found in [9] and recently a new technique based on a suitable partition of the phase space into "resonances" has been developed, but not all the statistical assumptions of this method are justified yet [19–22].

   Moreover the technical difficulties of establishing exponential decay of correlations from the construction of a Markov partition has stimulated a search for alternative, more direct techniques of estimating the rate of decay of correlations. The current paper is in this spirit.

   We remember that Markov partitions for dynamical systems with singularities (in particular billiards) have been constructed by Bunimovich and Sinai [28–30]; see also Levi [31] and a recent paper by Krüger and Troubetzkoy [32].

   Our way to estimate the diffusion coefficient is the following: we firstly recall the equality (0.3) that holds in the RPA and neglecting the higher-order corrections

to D for K large; then we introduce a continuous effective time $t \in [0, 1]$ and set $n = [Kt]$, where $[\cdot]$ denotes the integer part. The variance of $\Delta A_{[Kt]}$ scales now like $K^3 t \langle g^2 \rangle$, so that if we define the stochastic process:

$$\xi_K(t, \omega) = \frac{\Delta A_{[Kt]}(\omega)}{\sqrt{\langle g^2 \rangle K^3}}, \quad \omega = (A, \theta) \in \mathbb{T}^2, \tag{0.6}$$

on the probability space of the initial values, we get the following limit (see Theorem 1.1 below):

$$\langle (\xi_K(t, \omega))^2 \rangle \xrightarrow[K \to \infty]{} t. \tag{0.7}$$

Equation (0.6) is a necessary (but not sufficient) condition for the process $\xi_K(t, \omega)$ to converge in probability to the normalized Wiener process. Let us however remark that ΔA_n measures the diffusion of the action variable in the cylinder $\mathbb{R} \times \mathbb{T}$. After a rescaling of the discrete time $n = [Kt]$ and a suitable normalization, (0.6) suggests that $\Delta A_{[Kt]}$ is approximately distributed for K large, as the path up to time $t = 1$ of a brownian particle (see Billingsley, p. 5 in [12]). What we prove in the present paper is the existence and the validity of the limit (0.7). We also get a rigorous finite-time estimate for the diffusion coefficient $D_n = \frac{\langle (\Delta A_n)^2 \rangle}{\langle g^2 \rangle n}$ after n steps. More precisely, we will prove that the difference $\left| \frac{D_n}{D_L} - 1 \right|$ is of order K^{-1} for times n up to $[K]$.

The proofs of these results are based upon a detailed analysis of the two-point correlation function $\langle g(\theta) g(\theta_n(A, \theta)) \rangle$, and don't require ergodicity for the map. We believe that this method can be extended to a larger class of transformations, including the smooth ones. The core of the proof is a detailed control of the points of discontinuity of $\theta_n(A, \theta)$, $\theta \in [-\pi, \pi)$, for A fixed.

For smooth g's, these points should be replaced by a small neighbourhood of the critical points of θ_n, where the derivative of the map becomes small. These latter points are responsible for islands of stability in the phase space, and for breaking the global ergodicity.

We actually prove for the sawtooth map (and we guess it should also be true for the standard map), that those singular regions give a contribution of order $K^{-\alpha}$ (for some $\alpha \geq 2$) to the correlations, which is sufficient in our probabilistic approach. We incidentally note that for integer K's our method recovers the exponential decay of the two-point correlation function as expected, showing in this case that the diffusion coefficient D coincides with the quasi-linear one in the limit $K \to \infty$.

1. Notations and Statements of the Results

The sawtooth map (SM) T is defined on the cylinder $\mathbb{R} \times \mathbb{T}$ with coordinates (A, θ) by:

$$T(A, \theta) = (A + Kg(\theta), A + \theta + Kg(\theta) \bmod 2\pi), \tag{1.1}$$

where K is a real positive number and $g(\theta)$, $\theta \in \mathbb{R}$ is a 2π-periodic piecewise linear function given by $g(\theta) = \theta$, for $\theta \in [-\pi, \pi)$. We set $T^n(A, \theta) = (A_n(A, \theta), \theta_n(A, \theta))$ for

$n \in \mathbb{Z}$. We will show in Sect. 2 that:

$$A_n(A, \theta) = A + K \sum_{s=0}^{n-1} g(\theta_s(A, \theta)). \tag{1.2}$$

In particular, the action increment $\Delta A_n(A, \theta) = A_n(A, \theta) - A$ is a 2π-periodic function of both variables (A, θ). Remarking that T commutes with the translation of A by 2π, we denote by $\tilde{T}$ the map induced by T on the 2-torus $\mathbb{T}^2$ just making A an angle variable in (1.1)

$$\tilde{T}(A, \theta) = (A + K g(\theta) \bmod 2\pi, A + \theta + K g(\theta) \bmod 2\pi). \tag{1.3}$$

$\tilde{T}$ preserves the Lebesgue measure $\mu_L = \dfrac{dA \, d\theta}{4\pi^2}$ on $\mathbb{T}^2$.

Then we introduce the probability space $\mathscr{P} = \{\mathbb{T}^2, \eta, \mu_L\}$, where η is the Borel σ-algebra and we denote by $\mathbb{E}, \mathbb{V}$ respectively the corresponding *expectation* and *variance* on $\mathscr{P}$. One easily checks that ΔA_n is actually defined on $\mathscr{P}$ with a zero expectation, and it describes a stationary process on $\mathscr{P}$.

Our next step is to formulate an "invariance principle" for the process ΔA_n as briefly sketched in the introduction. We begin by setting $n = [Kt]$, $t \in [0, 1]$ and replace $\Delta A_{[Kt]}$ by (1.2). Using the notation $\xi_K(t, \omega)$ defined in (0.6) the main result of this paper anticipated by (0.7) can now be restated as:

Theorem 1.1. *We have the following limit:*

$$\lim_{K \to \infty} \mathbb{V}\left(\frac{1}{\sigma \sqrt{K}} \sum_{s=0}^{[Kt]-1} g(\theta_s(A, \theta))\right) = \lim_{K \to \infty} \langle (\xi_K(t, \omega))^2 \rangle = t, \tag{1.4}$$

where $\sigma^2 = \mathbb{V}(g) = \frac{1}{3}\pi^2$. $\diamond$

As pointed out in the introduction, this result suggests that the process $\xi_K(t, \omega)$ converges to the normalized Wiener process (and this is called an invariance principle), in the following sense: the measure μ_L induces a family of probability distributions $\tilde{\mu}_K$ on the space $\mathscr{C}$ of continuous functions on $[0, 1]$ with the uniform topology, according to:

$$\tilde{\mu}_K(B) = \mu_L\{\omega \in \mathscr{P}; \xi_K(t, \omega) \in B\}, \tag{1.5}$$

whenever B is a measurable subset of $\mathscr{C}$. We have an invariance principle whenever $\tilde{\mu}_K$ converges weakly to the Wiener measure. To prove such a result, Theorem 1.1 is a necessary condition. A sufficient condition is given by the convergence of all the finite-dimensional distributions of the process. We have not investigated this latter point yet, and we postpone it to a future work.

It is interesting to note that as a by-product of the proof of Theorem 1.1 we can estimate the diffusion coefficient. In fact, if $D_n = \dfrac{\mathbb{V}(\Delta A_n)}{n\sigma^2}$ and D_L are respectively the diffusion coefficient at "time" n and the quasi-linear coefficient we are able to show:

Proposition 1.2. *For* $K > \dfrac{\sqrt{2}}{2}$ *and* $n \geq 0$, *the following estimate holds:*

$$\left|\frac{D_n}{D_L} - 1\right| \leq \frac{288 \, n}{K^2 - 0.5} + \frac{18}{K + 1}. \quad \diamond \tag{1.6}$$

Theorem 1.1 follows immediately from this proposition. The singularity at $K = \dfrac{\sqrt{2}}{2}$ is apparent: it arises to optimize the bound for K large; in fact the key estimate for the proof of Proposition 1.2 is the following lemma:

Lemma 1.3. *We have for* $K \geq 0$: $|\mathbb{E}(g(\theta)g(\theta_s(A, \theta)))| \leq \dfrac{2\pi^2}{a^2}(2 + 22\, c_{s,k})$ *for* $s \geq 2$;

$|\mathbb{E}(g(\theta)g(\theta_1(A, \theta)))| \leq \dfrac{3\pi^2}{K + 1}$, *where* $c_{s,k} = \dfrac{a^s - 2^s}{a^{s-1}(a - 2)}$ *and* $a = \dfrac{K + 2 + \sqrt{K^2 + 4K}}{2}$

is the (constant) positive eigenvalue of the derivative of T. $\diamond$

Proof of Proposition 1.2. By an elementary calculation we get:

$$\left| \dfrac{D_n}{D_L} - 1 \right| \leq \dfrac{2}{\sigma^2} \left| \dfrac{n-1}{n} \mathbb{E}(g(\theta)g(\theta_1(A, \theta))) + n \sum_{s=2}^{n-1} \dfrac{1}{n}\left(1 - \dfrac{s}{n}\right) \mathbb{E}(g(\theta)g(\theta_s(A, \theta))) \right|. \quad (1.7)$$

The term $\displaystyle\sum_{s=2}^{n-1} \dfrac{1}{n}\left(1 - \dfrac{s}{n}\right)$ is bounded by the Riemann integral $\displaystyle\int_0^1 dx(1 - x) = \dfrac{1}{2}$. Using Lemma 1.3, it is easy to conclude. $\diamond$

Here we give a sketch of the proof, that is reduced to estimate the integral of the product $g(\theta)\cdot g(\theta_n(A, \theta))$ only with respect to the angle θ. An important combinatorial role will be played by the graph of the function $\theta_n(A, \theta)\text{-mod } 2\pi$, $\theta\in[-\pi,\pi)$, which we simply call $F_n(\theta)$ keeping A fixed. This graph is made of straight lines with slope $\partial\theta_n$ (namely the derivative $\partial F_n(\theta)/\partial\theta$) which cut (whenever $K \notin \mathbb{N}$) along vertical lines located at the discontinuity points of $F_n(\theta)$. This graph determines a partition of the basic torus $[-\pi, \pi)$ in the following way. Let S_n be the set of θ's in the interval $[-\pi, \pi)$ such that either $F_n(\theta)$ is an odd multiple of π, or θ is a discontinuity points of $F_n(\theta)$. This set divides $[-\pi, \pi)$ into two different kinds of intervals. In the first ones (good intervals or g-typed), both their endpoints are continuity points of $F_n(\theta)$; in the second ones (bad interval, or b-typed) at least one of the endpoints is a discontinuity point of $F_n(\theta)$. We then split the integral:

$$\int_{-\pi}^{\pi} g(\theta)g(\theta_n(A, \theta)) \dfrac{d\theta}{2\pi}$$

over the partition defined by S_n. In each term of this decomposition, $g(\theta)$ can be replaced by a constant, with an error of $O(K^{-n})$. Then, it remains to integrate $g(\theta_n(A, \theta))$. If we have a g-typed interval, the integral vanishes. Whereas on b-typed interval, the integral can be written in one of the two following forms:

(i) $\dfrac{1}{\partial F_n} \displaystyle\int_{-\pi}^{\alpha_n} g(y)\dfrac{dy}{2\pi}$;   (ii) $\dfrac{1}{\partial F_n} \displaystyle\int_{\beta_n}^{\pi} g(y)\dfrac{dy}{2\pi}$,

where α_n (respectively β_n) is different from π (respectively $-\pi$), and the slope ∂F_n is of order K^n. We will prove that, up to an error of order K^{-2}, one can group these integrals in approximately K^{n-2} terms, each of which being the sum of $l_n(K) = O(K)$ integrals of type (i) and (ii) above, where now the boundaries α_J and β_J $(1 \leq J \leq l_n(K))$, are distributed on $\mathbb{T}$ according to a well-defined rotation of the circle. The contribution of each sum can be estimated by Denjoy's inequality and gives a bound of order 1 (and not K). Since each integral is multiplied by the

inverse of the slope of $F_n(\theta)$, we finally get a bound of order K^{-2}, which is enough to prove Lemma 1.3.

2. Structures of the Discontinuities

In this section we prepare some technical tools needed to prove Lemma 1.3. Our major interest is for non-integer K. It is easy to prove the following lemma by induction:

Lemma 2.1. *For positive n's the following relations hold:*

$$A_n(A, \theta) = A + K \sum_{s=0}^{n-1} g(\theta_s(A, \theta)),$$

$$\theta_n(A, \theta) = \theta + nA + K \sum_{s=0}^{n-1} (n - s)g(\theta_s(A, \theta)),$$

and similar relations hold for negative n's. ◇

Let $F_n(\theta)$ be the function on $\mathbb{R}$ defined as $F_n(\theta) = \theta_n(A, \theta)$ where A is kept fixed once and for all. Since $F_n(\theta + 2\pi) = F_n(\theta) + 2\pi$, F_n is itself the lift of a function on the torus $\mathbb{T}$. So that it is sufficient to consider the restriction of $F_n(\theta)$ on the interval $[-\pi, \pi)$. We also remark that the function $\tilde{F}_n(\theta) \equiv F_n(\theta) - (\theta + nA)(\theta \in \mathbb{R})$ is 2π-periodic. Now, the function $g(\theta)$, $\theta \in \mathbb{R}$, is monotone increasing everywhere but on the set $(2\mathbb{Z} + 1)\pi$ of odd multiples of π, where the mapping T becomes discontinuous for $K \notin \mathbb{N}$. Let Disc $\{F_n(\theta)\}$ denote the set of points in $[-\pi, \pi)$, where $F_n(\theta)$ is discontinuous (including the end points). By definition, we clearly have Disc $\{F_n(\theta)\}$ = Disc $\{\tilde{F}_n(\theta)\}$. Therefore:

$$\text{Disc}\,\{F_n(\theta)\} \subset \bigcup_{s=0}^{n-1} \text{Disc}\,\{g \circ F_s(\theta)\}. \tag{2.1}$$

We then remark that $g \circ F_s(\theta)$ is discontinuous at points θ such that either Disc $\{F_n(\theta)\} \neq \varnothing$ or $F_n(\theta) = \pi \,(\text{mod}\, 2\pi)$. Together with (2.1), this gives:

Lemma 2.2. *For $n > 1$ we get the recursion formula for the discontinuities of F_n:*

$$\text{Disc}\,\{F_n(\theta)\} \subset \bigcup_{s=0}^{n-1} (\text{Disc}\,\{F_s(\theta)\}) \cup \{\theta \in [-\pi, \pi); F_{n-1}(\theta) = \pi \,(\text{mod}\, 2\pi)\}. \quad ◇$$

The previous lemma also shows that Disc $\{F_n(\theta)\}$ is a finite set. We will give later on an upper bound on its cardinality. The derivative ∂F_n at the points of continuity is independent of θ (and of A too) and is given by:

Proposition 2.3. *In the points where it is defined, the partial derivative of $F_n(\theta)$ with respect to θ is given by:*

$$\partial F_n = 1 + K \sum_{j=0}^{n-1} (n - j)\partial F_j = \frac{a^{2n+1} + 1}{(a + 1)a^n}, \quad n \geq 0. \quad ◇ \tag{2.2}$$

Proof. Starting from the formulæ in Lemma 2.1, since $\dfrac{\partial g}{\partial \theta} = 1$ we get immediately:

$$\partial F_n = 1 + K \sum_{j=0}^{n-j} (n-j)\partial F_j.$$

The generating function is the formal power series in X given by:

$$F(X) = \sum_{m=0}^{\infty} X^m \cdot \partial F_m.$$

It follows immediately that F satisfies the following equation:

$$F(X) = \frac{1}{1-X} + K \frac{X}{(1-X)^2} F(X),$$

and therefore

$$F(X) = \frac{1-X}{1-(K+2)X + X^2}.$$

Remembering the definition of the constant a in Lemma 1.3, one can expand easily F in power series using $1 - (K+2)X + X^2 = (a - X)(a^{-1} - X)$. An elementary calculation leads to the result.   ◇

Corollary 2.4. *We get the following estimates:*

$$K + 2 - \frac{1}{K+1} < a < K + 2 - \frac{1}{K+2}, \quad \frac{a}{a+1} a^n \leq \partial F_n \leq a^n. \quad ◇ \qquad (2.3)$$

Proof. Since $a + a^{-1} = K + 2$, it follows that $K > 0 \Rightarrow 1 < a < K + 2$. Thus $\dfrac{1}{K+2} < a^{-1} < 1 \Rightarrow K + 1 < a < K + 2 - \dfrac{1}{K+2}$. Hence $\dfrac{1}{K+2} < a^{-1} < \dfrac{1}{K+1}$ giving the first estimate in (2.3). On the other hand, from (2.2) we get $\partial F_n = \dfrac{a^{2n+1} + 1}{(a+1)a^n} \leq \dfrac{a^{2n+1} + x}{(a+x)a^n}$ for $0 \leq x \leq 1$. Setting $x = 0$ we get the upper bound in (2.3). Using $a^{2n+1} + 1 \geq a^{2n+1}$, we immediately get lower bound.   ◇

We want now to get a more precise description of Disc $\{F_n(\theta)\}$. For $n = 1$, $F_1(\theta) = \theta + A + Kg(\theta)$, and we get Disc $\{F_1(\theta)\} = \{-\pi\}$. For $n = 2$, $F_2(\theta) = \theta + 2A + 2Kg(\theta) + Kg(F_1(\theta))$, so that beside $\theta = \{-\pi\}$, the discontinuities of $F_2(\theta)$ are given by the points θ_{J_1} such that $F_1(\theta_{J_1}) = (2J_1 + 1)\pi$. We then remark that $F_1(-\pi + 0) = A - (K+1)\pi$ and $F_1(\pi - 0) = A + (K+1)\pi$, so that the total variation of F_1 is $2\pi(K+1)$ and we get for any fixed A:

$$\theta_{N_1^- + 1} - \frac{2\pi}{\partial F_1} < -\pi < \theta_{N_1^- + 1} < \theta_{N_1^- + 2} < \cdots < \theta_{J_1} < \cdots < \theta_{N_1^+} < \pi < \theta_{N_1^+} + \frac{2\pi}{\partial F_1}$$

with

$$\theta_{J+1} - \theta_J = \frac{2\pi}{\partial F_1}, \quad N_1^- + 1 \leq J_1 \leq N_1^+ - 1.$$

The preceding equations define the set of indices $N_1^\pm$; moreover we also set $\theta_{N_1^-} = -\pi$ and $\theta_{N_1^+ + 1} = \pi$, so that:

$$\theta_{N_1^+} - \theta_{N_1^- + 1} = (N_1^+ - N_1^- - 1)\frac{2\pi}{\partial F_1} \leq 2\pi \leq (N_1^+ - N_1^- + 1)\frac{2\pi}{\partial F_1},$$

namely,

$$(N_1^+ - N_1^- - 1) \leq \partial F_1 \leq (N_1^+ - N_1^- + 1).$$

We denote by $I_J = [\theta_J; \theta_{J+1})$, for $N_1^- \leq J \leq N_1^+$, the intervals defined by this partition of $[-\pi, \pi)$. In each such interval, F_1 is continuous, affine and takes on values in $[(2J+1)\pi, (2J+3)\pi]$. In particular, $g \circ F_1$ and therefore F_2 also, is continuous and affine there, and the discontinuities of F_2 are exactly given by the endpoints of these intervals. So:

$$\text{Disc}\{F_2\} = \{\theta_J; N_1^- \leq J \leq N_1^+ + 1\}.$$

In much the same way we get recursively a family of intervals (called intervals of type $s-1$), giving a partition of $[-\pi, \pi]$, and defined by:

$$I_{J_1,\ldots,J_{s-1}} = [\theta_{J_1,\ldots,J_{s-1}}; \theta_{J_1,\ldots,J_{s-1}+1}),$$

where:

$$N_1^- \leq J_1 \leq N_1^+, \quad N_p^-(J_1,\ldots,J_{p-1}) \leq J_p \leq N_p^+(J_1,\ldots,J_{p-1}) \quad 2 \leq p \leq s-1,$$

$$\theta_{J_1,\ldots,J_{s-2},N_{s-1}^+ + 1} \equiv \theta_{J_1,\ldots,J_{s-2}+1} = \theta_{J_1,\ldots,J_{s-2}+1,N_{s-1}^-},$$

having also defined: $\theta_{J_1,\ldots,J_{s-2}} = \theta_{J_1,\ldots,J_{s-2},N_{s-1}^-}$, and:

$$\theta_{N_1^+,N_2^+,\ldots,N_{s-2}^+,N_{s-1}^+ + 1} = \pi.$$

Moreover the indices $N_s^\pm(J_1,\ldots,J_{s-1})$ are defined in such a way as to satisfy:

$$(2N_s^- + 1)\pi \leq F_s(\theta_{J_1,\ldots,J_{s-1}} + 0) < (2N_s^- + 3)\pi,$$
$$(2N_s^+ + 1)\pi \leq F_s(\theta_{J_1,\ldots,J_{s-1}+1} - 0) < (2N_s^+ + 3)\pi,$$

and the new interval at the sth step are defined by:

$$F_s(\theta_{J_1,\ldots,J_{s-1},J_s}) = (2J_s + 1)\pi \quad \text{whenever} \quad N_s^- < J_s \leq N_s^+$$

We call "good" the intervals of type $(s-1)$ with $J_{s-1} \neq N_{s-1}^\pm$. Actually the function $F_{s-1}(\theta)$ is monotone increasing with variation 2π on them, so that their width is $2\pi/\partial F_{s-1}$. On a "bad" intervals with $J_{s-1} = N_{s-1}^-$ (respectively N_{s-1}^+), the function $F_{s-1}(\theta)$ is discontinuous at the left (respectively right) endpoint, or at both.

This latter case may happen if there is a bad interval of type $(s-2)$ with length less than or equal to $\dfrac{2\pi}{\partial F_{s-1}}$. Such an intervals can be written as $I_{J_1,\ldots,J_{s-2}} = [\theta_{J_1,\ldots,J_{s-2}}; \theta_{J_1,\ldots,J_{s-2}+1})$, where

$$\theta_{J_1,\ldots,J_{s-2},N_{s-1}^- + 1} = \theta_{J_1,\ldots,J_{s-2}+1}$$

or

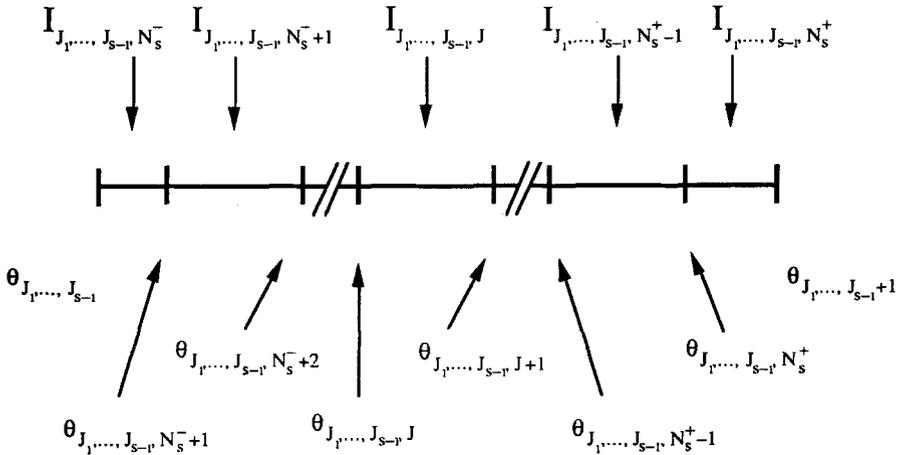$$\theta_{J_1,\ldots,J_{s-2},N_{s-1}^+} = \theta_{J_1,\ldots,J_{s-2}} = \theta_{J_1,\ldots,J_{s-2},N_{s-1}^-}.$$

Fig. 1. The discontinuities of F_s and the corresponding partition

This means that $I_{J_1,\ldots,J_{s-2}}$ generates only one interval of type $(s-1)$ which we can also denote by $I_{J_1,\ldots,J_{s-2},N^{\pm}_{s-1}}$ where $N^-_{s-1}(J_1,\ldots,J_{s-2})=N^+_{s-1}(J_1,\ldots,J_{s-2})$.

On the other hand, it is easy to see that in any good interval of type $(s-1)$, there are approximately $\partial F_s/\partial F_{s-1}$ good intervals of type (s). More precisely, if $N^{\pm}_{s-1}=N^{\pm}_{s-1}(J_1,\ldots,J_{s-2})$ with $J_{s-1}\neq N^{\pm}_{s-1}$, we get:

$$N^+_s - N^-_s - 1 \leq [\partial F_s/\partial F_{s-1}] \leq N^+_s - N^-_s + 1. \qquad (2.4)$$

For large K this number can be estimated using Proposition 2.3:

Lemma 2.5. For $s \geq 1$, we get $a - \dfrac{1}{a^{2(s-1)}} \leq \dfrac{\partial F_s}{\partial F_{s-1}} \leq a.$ $\diamond$

Proof. From Proposition 2.3, we get $\dfrac{\partial F_s}{\partial F_{s-1}} = \dfrac{a^{2s+1}+1}{a(a^{2s-1}+1)} \leq \dfrac{a^{2s+1}+x}{a(a^{2s-1}+x)}$ for $0 \leq x \leq 1$. Setting $x = 0$, we get the upper bound. For the lower bound,

$$a - \dfrac{\partial F_s}{\partial F_{s-1}} = \dfrac{a^2-1}{a(a^{2s-1}+1)} \leq \dfrac{1}{a^{2(s-1)}}. \qquad \diamond$$

Let us call $\mathscr{P}_{s-1}$ the partition of $[-\pi,\pi]$ given by the intervals of type $(s-1)$ and let P_{s-1} be its cardinality. To get an upper bound of P_{s-1} we note that each interval I of $\mathscr{P}_{s-2}$ gives at most two (possibly bad) intervals of $\mathscr{P}_{s-1}$ having at least one endpoint in common with an endpoint of I. So they give a contribution of $2P_{s-2}$ to P_{s-1}. If we now subtract these latter intervals, all the remaining intervals of $\mathscr{P}_{s-1}$ are good and their number is simply bounded by $2\pi/(2\pi/\partial F_{s-1})$. Then we get: $P_{s-1} \leq 2P_{s-2} + \partial F_{s-1}$. We also observe that the cardinality of Disc $F_s(\theta)$ is surely bounded by $2P_{s-1}$; iterating the preceding inequality for P_{s-1} with $P_0 = 1$, and using (2.2) we get:

Lemma 2.6.

$$\#\mathrm{Disc}\, F_s(\theta) \leqq 2P_{s-1} \leqq 2\sum_{j=0}^{s-1} 2^j \partial F_{s-1-j} \leqq 2c_{s,k}a^{s-1}, \quad s \geqq 1,$$

where $c_{s,k}$ was defined in Lemma 1.3, Sect. 1. $\diamondsuit$

In the next section we have to compare $P_{s-\alpha}$ to ∂F_s for $\alpha = 1, 2, 3$.
Collecting Lemma 2.6 and Proposition 2.3 we easily get:

Lemma 2.7. *For $s \geqq \alpha$ we have* $\dfrac{P_{s-\alpha}}{\partial F_s} \leqq \dfrac{c_{s,k}}{a^\alpha}\left(1 + \dfrac{1}{a^{2s}}\right) \leqq \dfrac{2c_{s,k}}{a^\alpha}.$ $\diamondsuit$

3. Proof of Lemma 1.3

Now we have all the ingredients to prove Lemma 1.3 following the strategy
indicated at the end of Sect. 1. We will use the notation:

$$E_s = \mathbb{E}(g(\theta)g(\theta_s(A,\theta))) \leqq \int_{-\pi}^{\pi} dv\, g(\theta)g(\theta_s(A,\theta)), \quad s \geqq 1, \tag{3.1}$$

where $dv = \dfrac{d\theta}{2\pi}$.

Let us decompose (3.1) into a sum over the intervals $I_{J_1,\ldots,J_s}$ defined in Sect. 2.
Let $\theta^{(0)}_{J_1,\ldots,J_s}$ be a point in this interval. We can then write $g(\theta) = g(\theta^{(0)}_{J_1,\ldots,J_s}) + (\theta - \theta^{(0)}_{J_1,\ldots,J_s})$, and therefore we immediately get:

$$|E_s| = \left|\sum_{J_1\ldots J_s} \int_{I_{J_1,\ldots,J_s}} dv\, g(\theta)g(F_s(\theta))\right|$$

$$\leqq \left|\sum_{J_1\ldots J_s} g(\theta^{(0)}_{j_1,\ldots,J_s}) \int_{I_{J_1,\ldots,J_s}} dv\, g(F_s(\theta))\right|$$

$$+ \sum_{J_1\ldots J_s} \left|\int_{I_{J_1,\ldots,J_s}} dv(\theta - \theta^{(0)}_{J_1,\ldots,J_s})g(F_s(\theta))\right| = E(1) + E(2). \tag{3.2}$$

Using Lemma 2.4, we prove that the second term is bounded by:

$$E(2) \leqq \max_{J_1,\ldots,J_s}\, (|I_{J_1,\ldots,J_s}|) \sum_{J_1\ldots J_s} v(I_{J_1,\ldots,J_s})\max_\theta |g(\theta)| \leqq \frac{2\pi}{\partial F_s}\pi \leqq \frac{4\pi^2}{a^s}. \tag{3.3}$$

Now let us consider the first term $E(1)$ in (3.2). The good intervals in the sum give
no contribution for F_s is linear and varies from $-\pi$ to π, whereas g has mean
zero. Therefore only bad intervals must be kept, namely those for which $J_s = N_s^\pm$.
We now get $E(1) \leqq E(11) + E(12)$, where $E(11)$ is obtained by restricting the sum
in $E(1)$ to those intervals for which $J_{s-1} = N_{s-1}^\pm$, and $E(12)$ is obtained in the same
way with now $J_{s-1} \neq N_{s-1}^\pm$. An estimate similar to (3.3) using Lemma 2.7 will give
easily:

$$E(11) \leqq \max_\theta |g^2(\theta)| \sum_{J_1\ldots J_{s-2},\pm,\pm} v(I_{J_1,\ldots,J_{s-2},N_{s-1}^\pm,N_s^\pm}) \leqq 4\pi^2\frac{P_{s-2}}{\partial F_s} \leqq \frac{8c_{s,k}\pi^2}{a^2}. \tag{3.4}$$

Then we have to estimate the sums $E(12)$. Let now $\theta^{(0)}_{J_1,\ldots,J_{s-2}}$ be a point in the interior of $I_{J_1,\ldots,J_{s-2}}$ which contains $I_{J_1,\ldots,J_{s-1},N_s^{\pm}}$. A shift of the argument in $E(12)$ gives:

$$E(12) \leq \left| \sum_{J_1\ldots J_{s-2}} g(\theta^{(0)}_{J_1,\ldots,J_{s-2}}) \sum_{J_{s-1}\neq N_{s-1}^{\pm}} \sum_{\varepsilon = \pm} \int_{I_{J_1,\ldots,J_{s-1},N_s^{\varepsilon}}} dv\, g(F_s(\theta)) \right|$$

$$+ \left| \sum_{J_1\ldots J_{s-2}} \sum_{J_{s-1}\neq N_{s-1}^{\pm}} \sum_{\varepsilon = \pm} (\theta^{(0)}_{J_1,\ldots,J_{s-1},N_s^{\varepsilon}} - \theta^{(0)}_{J_1,\ldots,J_{s-2}}) \right.$$

$$\left. \cdot \int_{I_{J_1,\ldots,J_{s-1},N_s^{\varepsilon}}} dv\, g(F_s(\theta)) \right| = E(121) + E(122). \tag{3.5}$$

Using again the results of Sect. 2, we easily see that $E(122)$ is bounded by:

$$E(122) \leq \frac{4\pi^2 P_{s-1}}{\partial F_{s-2}\partial F_s} \leq \frac{16 c_{s,k}\pi^2}{a^s}. \tag{3.6}$$

Once again we split the sums involved in $E(121)$ into two pieces as follows:

$$E(121) \leq \left| \sum_{J_1\ldots J_{s-3}} \sum_{J_{s-2}=N_{s-2}^{\pm}} (\cdots) \right|$$

$$+ \left| \sum_{J_1\ldots J_{s-3}} \sum_{J_{s-2}\neq N_{s-2}^{\pm}} (\cdots) \right| = E(1211) + E(1212). \tag{3.7}$$

The first give a contribution which is easy to estimate:

$$E(1211) \leq 4\pi^2 \frac{\partial F_{s-1}}{\partial F_{s-2}} \frac{P_{s-3}}{\partial F_s} \leq \frac{8 c_{s,k}\pi^2}{a^2}. \tag{3.8}$$

Now we are ready to estimate the most difficult piece $E(1212)$. Since g is bounded by π, we can estimate this piece as follows:

$$E(1212) \leq \pi \sum_{J_1\ldots J_{s-3}} \sum_{J_{s-2}\neq N_{s-2}^{\pm}} \sum_{J_{s-1}\neq N_{s-1}^{\pm}} \sum_{\varepsilon = \pm} \int_{I_{J_1,\ldots,J_{s-1},N_s^{\varepsilon}}} dv\, g(F_s(\theta)) \Bigg|. \tag{3.9}$$

The same kind of estimate as above produces only a bound of order $O(K^{-1})$ which is not enough. So we have to use a more sophisticated technique. In order to do so, let us consider the partial sum:

$$S_{J_1,\ldots,J_{s-2}} = \sum_{J_{s-1}\neq N_{s-1}^{\pm}} \left(\int_{I_{J_1,\ldots,J_{s-1},N_s^-}} dv\, g(F_s(\theta)) + \int_{I_{J_1,\ldots,J_{s-1},N_s^+}} dv\, g(F_s(\theta)) \right).$$

$$\tag{3.10}$$

To deal with it we will need some more notations, namely:

(i) $\eta_0^+ = F_s(\theta_{J_1,\ldots,J_{s-2},N_{s-1}^-+1} + 0); \quad \eta_0^- = F_s(\theta_{J_1,\ldots,J_{s-2},N_{s-1}^-+2} - 0)$

(ii) $\Gamma_\delta(x) = \int_x^\pi g(y)\frac{dy}{2\pi} + \int_{-\pi}^{x+\delta} g(y)\frac{dy}{2\pi}, \quad \delta = \eta_0^- - \eta_0^+, \quad x \in \mathbb{T},$

(iii) $R_s(x) = x + 2\pi\alpha_s(K)(\mathrm{mod}\ 2\pi)$. This is actually the rotation on the circle $\mathbb{T}^1$ with rotation number:

$$\alpha_s(K) = \frac{1 + K\sum_{t=0}^{s-2}(s-t)\partial F_t}{\partial F_{s-1}} - 1 = \frac{\partial F_s}{\partial F_{s-1}} - K - 1. \qquad (3.11)$$
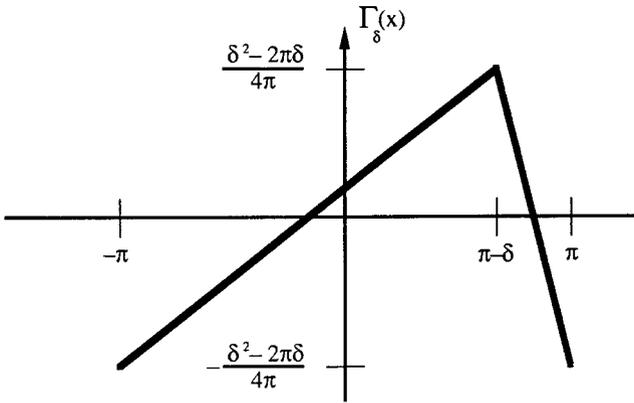

Fig. 2. Graph of Γ_δ

Lemma 3.1. *If $T_{s-1} = N_{s-1}^+ - N_{s-1}^- - 1$, the following identity hold:*

$$S_{J_1,\dots,J_{s-2}} = \frac{1}{\partial F_s}\sum_{t=0}^{T_{s-1}-1}\Gamma_\delta(\eta_0^+ + 2\pi t\alpha_s(K)) = \frac{S'_{J_1,\dots,J_{s-2}}}{\partial F_s}. \qquad \Diamond \qquad (3.12)$$

Proof. See Appendix 1. $\Diamond$

By inspection we can check that $\Gamma_\delta(x)$ is continuous, piecewise linear, with zero average on the circle, and a bounded total variation. The standard way to estimate the sum in (3.12) consists in using Denjoy's inequality [13, 14], giving an upper bound $\mathrm{Var}(\Gamma_\delta)$ provided the rotation number $\alpha_s(K)$ is sufficiently "well approximated" by a rational numbers of the form p/T_{s-1}. But we get:

Lemma 3.2. *For $s \geq 2$, we get the following estimate:*

$$\left|\alpha_s(K) - 1 + \frac{1}{T_{s-1}}\right| \leq \frac{3}{(T_{s-1})^2}. \qquad \Diamond \qquad (3.13)$$

Proof. From Lemma 2.5, we get $\alpha_s(K) - 1 + \dfrac{1}{T_{s-1}} \leq a - K - 2 + \dfrac{1}{T_{s-1}} = \dfrac{1}{T_{s-1}} - \dfrac{1}{a}.$

In much the same way $\alpha_s(K) - 1 + \dfrac{1}{T_{s-1}} \geq a - \dfrac{1}{a^{2(s-1)}} - K - 2 + \dfrac{1}{T_{s-1}} = \dfrac{1}{T_{s-1}} -$

$\frac{1}{a} - \frac{1}{a^{2(s-1)}} \geqq -\frac{1}{a^{2(s-1)}}$. Now we know (see Eq. (2.4)) that $T_{s-1} < a < T_{s-1} + 2 +$

$\frac{1}{a^{2(s-2)}} < T_{s-1} + 3$. A simple calculation leads then to the result. $\diamond$

Lemma 3.3 (Denjoy's Inequality). *Let h be a continuous function on the circle with bounded variation. Let $\alpha \in \mathbb{R}$ and $p/q \in \mathbb{Q}$ (p and q prime to each other) be numbers such that $|\alpha - p/q| \leqq \varepsilon$ for some $\varepsilon > 0$. If we denote by $[\![X]\!]$ the smallest integer larger than or equal to X, then for any $x \in \mathbb{T}$, we have:*

$$\left| \sum_{t=0}^{q-1} h(x + 2\pi\alpha t) - q \int_{\mathbb{T}} h(y) \frac{dy}{2\pi} \right| \leqq [\![q^2\varepsilon]\!] \cdot \mathrm{Var}(h). \quad \diamond \tag{3.14}$$

Proof. See Appendix 2. $\diamond$

Therefore we can now bound $S'_{J_1,\dots,J_{s-2}}$ using (3.12) and Lemmas 3.2 and 3.3:

$$|S'_{J_1,\dots,J_{s-2}}| \leqq 3 \,\mathrm{Var}(\Gamma_\delta) \leqq 6\pi, \tag{3.15}$$

for Γ_δ has a zero mean over the circle and its variation is bounded by 2π. Replacing in (3.12) and then in (3.9), one gets for $s \geqq 2$:

$$E(1212) \leqq \pi \sum_{J_1 \dots J_{s-3}} \sum_{J_{s-2} \neq N_{s-2}^{\pm}} \frac{S'_{J_1,\dots,J_{s-2}}}{\partial F_s} \leqq \pi \frac{P_{s-2}}{\partial F_s} 6\pi \leqq \frac{12 c_{s,k} \pi^2}{a^2}. \tag{3.16}$$

Thus for $s \geqq 2$ and $K \geqq 0$, we get:

$$|\mathbb{E}(g(\theta)g(\theta_s(A,\theta)))| \leqq E(2) + E(11) + E(122) + E(1211) + E(1212)$$

$$\leqq \frac{4\pi^2}{a^s} + \frac{8 c_{s,k}\pi^2}{a^2} + \frac{16 c_{s,k}\pi^2}{a^s} + \frac{8 c_{s,k}\pi^2}{a^2} + \frac{12 c_{s,k}\pi^2}{a^2} \leqq (2 + 22 c_{s,k}). \tag{3.17}$$

We can estimate $|\mathbb{E}(g(\theta)g(\theta_1(A,0)))|$ in much the same way. A direct calculation gives:

$$|\mathbb{E}(g(\theta)g(\theta_1(A,\theta)))| \leqq \frac{3\pi^2}{K+1}, \tag{3.18}$$

which achieves the proof of the main lemma. $\diamond$

Appendices

Appendix 1. Proof of Lemma 3.1

We consider the formula (3.10) which gives the main term to be estimated:

$$S_{J_1,\dots,J_{s-2}} = \sum_{J_{s-1} \neq N_{s-1}^{\pm}} \left(\int_{I_{J_1,\dots,J_{s-1},N_s^-}} dv\, g(F_s(\theta)) + \int_{I_{J_1,\dots,J_{s-1},N_s^+}} dv\, g(F_s(\theta)) \right). \tag{A1.1}$$

Since in this sum, $J_{s-2} \neq N^{\pm}_{s-2}$, the interval $I_{J_1,\ldots,J_{s-2}}$ is good and its length is $\dfrac{2\pi}{\partial F_{s-2}}$. This interval is subdivided into the intervals $I_{J_1,\ldots,J_{s-2},J_{s-1}}$, $N^-_{s-1}+1 \leq J_{s-1} \leq N^+_{s-1}$. In the sum (A1.1) only the values $J_{s-1} \neq N^{\pm}_{s-1}$ are taken into account, so that the number T_{s-1} of permitted values of the index J_{s-1} is $T_{s-1} = N^+_{s-1} - N^-_{s-1} - 1$, and we have $T_{s-1} < \dfrac{\partial F_{s-1}}{\partial F_{s-2}} < T_{s-1} + 2$. We then set:

(i) $t = J_{s-1} - (N^+_{s-1}+1) \Rightarrow 0 \leq t \leq T_{s-1} - 1$,
(ii) $\eta^-_t = F_s(\theta_{J_1,\ldots,J_{s-2},J_{s-1}} + 0) - (2N^-_s + 1)\pi$,
(iii) $\eta^+_t = F_s(\theta_{J_1,\ldots,J_{s-2},J_{s-1}+1} - 0) - (2N^+_s + 1)\pi$.

Note that $N^{\pm}_s$ may depend upon J_{s-1}. We know that

$$F_s(\theta) = \theta + sA + sKg(\theta) + (s-1)Kg(F_1(\theta)) + \cdots + 2Kg(F_{s-2}(\theta)) + Kg(F_{s-1}(\theta)),$$

is piecewise affine with a constant slope and that its discontinuities are precisely the endpoints of the intervals $I_{J_1,\ldots,J_{s-2},J_{s-1}}$. These discontinuities are due to the last term only, for F_{s-1} is continuous on $I_{J_1,\ldots,J_{s-2}}$. It follows that:

$$\eta^-_{t+1} - \eta^-_t = (\theta_{J_1,\ldots,J_{s-2},J_{s-1}+1} - \theta_{J_1,\ldots,J_{s-2},J_{s-1}})\left\{1 + K\sum_{j=0}^{s-2}(s-j)\partial F_j\right\}(\text{mod } 2\pi),$$

for indeed $g(F_{s-1}(\theta_{J_1,\ldots,J_{s-2},J_{s-1}} + 0)) = \pi$. Since $I_{J_1,\ldots,J_{s-2},J_{s-1}}$ is good, using the formula (2.2) we get:

$$\eta^-_{t+1} - \eta^-_t = \{\partial F_s - K\partial F_{s-1}\}\frac{2\pi}{\partial F_{s-1}} = 2\pi\alpha_s(K)(\text{mod } 2\pi). \qquad (\text{A1.2})$$

The same formula holds for η^+_t.

Let us now set $y = F_s(\theta)$ in (A1.1). Then $dy = \partial F_s d\theta$ and y varies from $\eta^{\pm}_t$ to $\pm\pi$, depending upon which interval is considered. This gives

$$S_{J_1,\ldots,J_{s-2}} = \frac{1}{\partial F_s}\sum_{t=0}^{T_{s-1}-1}\left\{\int_{\eta^+_t}^{\pi}\frac{dy}{2\pi}g(y) + \int_{-\pi}^{\eta^-_t}\frac{dy}{2\pi}g(y)\right\}.$$

Using the definition of Γ_δ given in Eq. (3.11(ii)) we get immediately the result. $\diamond$

Appendix 2. Denjoy's Inequality

Let h be a periodic function of period 1 with bounded variation on the circle. Let $\alpha \in \mathbb{R}$ and $p/q \in \mathbb{Q}$ (p and q prime to each other) be numbers such that $|\alpha - p/q| \leq \varepsilon$ for some $\varepsilon > 0$. If we denote by $[\![X]\!]$ the smallest integer larger than or equal to X, then for any $x \in \mathbb{T}$, we have:

$$\left|\sum_{t=0}^{q-1}h(x+\alpha t) - q\int_{\mathbb{T}}h(y)dy\right| \leq [\![q^2\varepsilon]\!]\,\text{Var}(h). \quad \diamond \qquad (3.14)$$

Without loss of generality we can assume that $0 \leq \alpha < 1$, $p/q \leq \alpha$, $x = 0$.

Then if $0 \leq t < q$ is an integer, let p_t be the integer in $[0, q)$ such that $p_t \equiv t_p$ (mod q). It follows that $t\alpha \in [p_t/q, p_t/q + q\varepsilon]$(mod 1). Since h has bounded variation,

there is a finite signed measure μ on $[0,1]$ such that $h(y) - h(x) = \mu([x,y])$, and $\mathrm{Var}(h) = \|\mu\|$. It follows that if $x \in [p_t/q, p_t/q + 1/q)$, we get:

$$|h(t\alpha) - h(x)| = \left| \int_x^{t\alpha} \mu(dy) \right| \leq \int_{(p_t/q)}^{(p_t/q) + q\varepsilon} |\mu(dy)| \leq \sum_{j=0}^{\llbracket q^2\varepsilon \rrbracket - 1} \int_{(p_t+j)/q}^{(p_t+j+1)/q} |\mu(dy)|. \quad (A2.1)$$

Note that (A2.1) is true also for $1/q \leq \varepsilon$, replacing the interval $[p_t/q, p_t/q + q\varepsilon]$ with the unit interval.

Integrating over x on both sides we get (all the p_t are distinct since p/q is an irreducible fraction):

$$\left| \sum_{t=0}^{q-1} h(\alpha t) - q \int_{\mathbb{T}} h(y)dy \right| \leq \sum_{t=0}^{q-1} q \int_{(p_t/q)}^{(p_t+1)/q} |h(\alpha t) - h(x)| dx$$

$$\leq \sum_{j=0}^{\llbracket q^2\varepsilon \rrbracket - 1} \sum_{s=0}^{q-1} q \int_{(s+j)/q}^{(s+j+1)/q} |\mu(dy)| \int_{(p_t/q)}^{(p_t+1)/q} dx \leq \llbracket q^2\varepsilon \rrbracket \, \mathrm{Var}(h). \quad \diamond$$

References

1. Chirikov, B. V.: A universal instability of many-dimensional oscillator systems. Phys. Rep. **52**, 263–379 (1979)
2. Lichtenberg, A. J., Lieberman, M. A.: Regular and stochastic motion, pp. 499. Berlin, Heidelberg, New York: Springer 1982
3. Greene, J. M., Mac Kay, R. S., Vivaldi, F., Feigenbaum, M. J.: Universal behavior in families of area preserving maps. Physica **D3**, 468–485 (1981)
4. Mac Kay, R. S., Meiss, J. D., Percival, I. C.: Transport in Hamiltonian Systems. Physica **D13**, 55–81 (1984) (and references therein)
5. Meiss, J. D., Ott, E.: Markov tree model of intrinsic transport in hamiltonian systems. Phys. Rev. Lett. **55**, 2741–2744 (1985)
6. Meiss, J. D., Ott, E.: Markov tree model of transport in area preserving maps. Physica **D20**, 387–402 (1986)
7. Lichtenberg, A. J., Lieberman, M. A.: Diffusion in two-dimensional mappings. Physica **D33**, 211–239 (1988)
8. Aubry, S., Abramovici, G.: Chaotic trajectories in the standard map. The concept of anti-integrability. Physica **D43**, 199–219 (1990)
9. Meiss, J. D., Cary, J. R., Grebogi, C., Crawford, J. D., Kaufman, A. N., Abarbanel, H. D. I.: Correlations of periodic area-preserving maps. Physica **D6**, 375–384 (1983)
10. Cary, J. R., Meiss, J. D.: Rigorously diffusive deterministic map. Phys. Rev. **A24**, 2624–2668 (1981)
11. Cary, J. R., Meiss, J. D., Bhattacharee, A.: Statistical characterization of periodic area preserving mappings. Phys. Rev. **A33**, 2744 (1981)
12. Billingsley, P.: Convergence of Probability measures, pp. 253 (New York: Wiley 1968)
13. Cornfeld, J. P., Fomin, S. V., Sinai, Ya. G.: Ergodic Theory, pp. 486 Berlin, Heidelberg, New York: Springer 1982
14. Hermann, M.: Sur la conjuguaison différentiable des difféomorphisms du cercles à une rotation, Publ. Math. IHES **49**, 5–234 (1979)
15. Vaienti, S.: Ergodic properties of the discontinuous sawtooth map, to appear in Journal of Statistical Physics

16. Percival, I., Vivaldi, F.: A linear code for the sawtooth and cat maps. Physica **D27**, 373–386 (1987)
17. Vivaldi, F.: Arithmetic theory of Anosov diffeomorphisms. Proc. R. Soc. **A413**, 97–107 (1987)
18. Bird, N., Vivaldi, F.: Periodic orbits of the sawtooth maps. Physica **D30**, 164 (1988)
19. Dana, I.: Hamiltonian transport on unstable periodic orbits. Physica **D39**, 205–230 (1989)
20. Dana, I., Murray, N. W., Percival, J. C.: Resonances and diffusion in periodic Hamiltonian maps. Phys. Rev. Lett. **62**, 233–236 (1989)
21. Dana, I.: Organization of Chaos in area-preserving maps. Phys. Rev. Lett. **64**, 2339–2342 (1990)
22. Chen, Q., Dana, I., Meiss, J. D., Murray, N. W., Percival, I. C.: Resonances and transport in the sawtooth map. Physica **46D**, 217–240 (1990)
23. Rechester, A. B., White, R. B.: Calculation of turbulent diffusion for the Chirikov–Taylor model. Phys. Rev. Lett. **44**, 1586–1589 (1980)
24. Rechester, A. B., Rosenbluth, M. N., White, R. B.: Fourier-space paths applied to the calculation of diffusion in the Chirikov–Taylor model. Phys. Rev. **A23**, 2664–2672 (1981)
25. Abarbanel, H. D. J.: Non-linear diffusion in hamiltonian systems exhibiting chaotic motion. Physica **D4**, 89–104 (1981)
26. Karney, C. F. F., Rechester, A. B., White, R. B.: Effect of noise on the standard mapping. Physica **D4**, 425–438 (1982)
27. Murray, N. W., Lichtenberg, A. J., Lieberman, M. A.: Corrections to quasilinear diffusion in area preserving maps. Phys. Rev. **A32**, 2413–2424 (1985)
28. Bunimovich, L. A., Sinai, Ja. G.: Markov Partitions for dispersed billiards. Commun. Math. Phys. **78** (2), 247–280 (1980/81)
29. Bunimovich, L. A., Sinai, Ja. G.: Statistical Properties of Lorentz Gas with Periodic Configuration of Scatterers. Commun. Math. Phys. **78**, 479–497 (1981)
30. Bunimovich, L. A., Sinai, Ja. G., Chernov, N. I.: Markov partitions for two dimensional hyperbolic Billiards. Russ. Math. Surv. **45** (3), 105–152 (1990)
31. Levy, Y.: A note on Sinai and Bunimovich's Markov partition for billiards. J. Stat. Phys. **45** (1/2), 63–68 (1986)
32. Krüger, T., Troubetzkoy, S.: Markov partitions and shadowing for non-uniformly hyperbolic systems with singularities, Preprint BiBoS n° 465/91

Communicated by T. Spencer