Journal of Statistical Physics, Vol. 92, Nos. 5/6, 1998

Finite Range of Large Perturbations in Hamiltonian Dynamics

D. Bénisti[1] and D. F. Escande[1]

Received January 21, 1998; final June 1, 1998

The dynamics defined by the Hamiltonian $H = p^2/2 + A \sum_{m=-M}^{M} \cos(q - mt + \varphi_m)$, where the φ_m are fixed random phases, is investigated for large values of A, and for $M \gg A^{2/3}$. For a given P^* and for $\Delta v \geqslant A^{2/3}$, this Hamiltonian is transformed through a rigorous perturbative treatment into a Hamiltonian where the sum of all the nonresonant terms, having a Q dependence of the kind $\cos(kQ - nt + \varphi_m)$ with $|n/k - P^*| > \Delta v$, is a random variable whose r.m.s. with respect to the φ_m is exponentially small in the parameter $\varepsilon = A/\Delta v^{3/2}$. Using this result, a rationale is provided showing that the statistical properties of the dynamics defined by H, and of the reduced dynamics including at each time t only the terms in H such that $|m - p(t)| \leqslant \alpha A^{2/3}$, can be made arbitrarily close by increasing α. For practical purposes α close to 5 is enough, as confirmed numerically. The reduced dynamics being nondeterministic, it is thus analytically shown, without using the random-phase approximation, that the statistical properties of a chaotic Hamiltonian dynamics can be made arbitrarily close to that of a stochastic dynamics. An appropriate rescaling of momentum and time shows that the statistical properties of the dynamics defined by H can be considered as independent of A, on a finite time interval, for A large. The way these results could generalize to a wider class of Hamiltonians is indicated.

KEY WORDS: Perturbation theory; Hamiltonian dynamics; wave–particle interaction: transport properties; chaos; plasma turbulence.

1. INTRODUCTION

Perturbation theory has proven its efficiency for dealing with small perturbations in Hamiltonian Systems. This is exemplified in the Kolmogorov–

[1] Equipe Turbulence Plasma du Laboratoire de Physique des Interactions Ioniques et Moléculaires, UMR 6633 CNRS-Université de Provence, Centre universitaire de Saint-Jérôme, F-13397 Marseille Cedex 20, France; present address: Consorzio RFX, Corso Stati Uniti, 4, 35127 Padova, Italy; e-mail: benisti@igi.pd.cnr.it, escande@igi.pd.cnr.it.

0022-4715/98/0900-0909$15.00/0

Arnold–Moser[1,2] (KAM), and the Nekhoroshev[3,4,5] theorems. As yet perturbation techniques have not been used to deal with Hamiltonians far enough from integrable ones for large scale chaos to be dominant in the corresponding dynamics. This paper yields the first example showing that this may be done in a rigorous way, by dealing with the dynamics defined by

$$H = p^2/2 + A \sum_{m=-M}^{M} \cos(q - mt + \varphi_m) \tag{1}$$

where the φ_m's are fixed random phases. In the case when all the φ_m's are 0, and when M goes to infinity, the Hamiltonian H defines the same dynamics as the standard map with parameter $K = 4\pi^2 A$. In the general case H defines the dynamics of a particle acted upon by a broadband spectrum of electrostatic waves.

Using perturbation theory we are able to identify in (1) the terms which give non-negligible contributions when studying the statistical properties of the dynamics defined by H. Intuitively, it seems clear that the term $\cos(q - mt + \varphi_m)$ (from a physical point of view a wave with phase velocity m), which is resonant for $p = m$, acts as a fast, and therefore negligible, perturbation for large enough values of $|m - p(t)|$. In this paper, this intuitive idea is clarified, and we analytically show that the statistical properties of the dynamics defined by (1) can be made arbitrarily close to that of the reduced dynamics defined by $dq/dt = p$, $dp/dt = A \sum_{|m-p(t)| \leqslant \Delta v_R} \sin(q - mt + \varphi_m)$, where $\Delta v_R = \alpha A^{2/3}$, simply by increasing α. This reduced dynamics can be seen as deriving from the effective Hamiltonian

$$H' = \frac{p^2}{2} + A \sum_{|m-p(t)| \leqslant \Delta v_R} \cos(q - mt + \varphi_m) \tag{2}$$

Therefore, large perturbations have a finite range on the p axis; in physical language, wave–particle interaction occurs locally in phase velocity.

Perturbation theory thus makes it possible to extract from (1) the relevant, and more simple, dynamics to be studied in order to derive the statistical properties of (1). Now, we stress that, while the dynamics defined by (1) is deterministic, the dynamics of (2) is not in the case where A is large enough to allow changes in p larger than unity. Indeed, as p varies by more than unity, the definition of the dynamical system in (2) changes, and changes in a random fashion because the φ_m's are random phases. Therefore, this paper analytically shows, without ever making any random phase approximation, that the statistical properties of a chaotic Hamiltonian dynamics can be made arbitrarily close to that of a stochastic dynamics.

The property of locality in phase velocity implies that when $M \gg A^{2/3}$, and as long as $|M \pm p(t)|$ remains large enough, the statistical properties of the dynamics defined by (1) can be considered as independent of M. Moreover, it will be shown that locality implies that on a finite time interval independent of A, in the variables $(q_r = q,\ p_r = A^{-2/3}p,\ t_r = A^{2/3}t)$, the statistical properties of (1) can be considered as independent of A. Therefore, the dynamics of H exhibits universal behaviors with respect to both the number of waves, and the wave amplitudes.

Our perturbation analysis is the central part of this paper. It is formally similar to the one used to prove the KAM, or Nekhoroshev theorems, but with two important differences: first we consider the case of large perturbations, and second we do not attempt to derive a result about the details of a particle's dynamics, but only about the statistical properties of the dynamics. The fact that we deal with large perturbations makes it impossible to remove, as in KAM theorem, all the resonances present in (1). But this does not mean that none of them can be removed. Actually, our result is closer to the one obtained in the Nekhoroshev theory. This theory focuses on Hamiltonians which are close to integrable ones, and therefore on the form: $H^{(N)}(q, p) = h^{(N)}(p) + \varepsilon f^{(N)}(q, p)$, where $h^{(N)}(p)$ is integrable, ε is a small parameter, and $f^{(N)}(q, p)$ is bounded for example $|f^{(N)}(q, p)| \leqslant 1$. In the so-called analytic part of Nekhoroshev theorem, one defines a canonical change of variables $(q, p) \mapsto (Q, P)$ which transforms $H^{(N)}(q, p)$ in the Hamiltonian

$$\hat{H}^{(N)}(Q, P) = h^{(N)}(P) + \varepsilon K^{(N)}(Q, P) + R^{(N)}(Q, P, \varepsilon) \tag{3}$$

In the case where are only considered values of p close to a value p^* such that $\omega(p^*) = (dh/dp)_{p=p^*}$ is a rational, then the Fourier expansion of $K^{(N)}$, $K^{(N)}(Q, P) = \sum_\nu K_\nu^{(N)}(P)\, e^{i\nu \cdot Q}$, only contains values of ν such that $\nu \cdot \omega(p^*) = 0$ (see ref. 5). Therefore, in the case of Nekhoroshev theory, one gives up removing all the resonances present in the initial Hamiltonian. Moreover, in the case where the Hamiltonian $H^{(N)}$ is of the same form as (1), then the condition $\nu \cdot \omega(p^*) = 0$ simply amounts to saying that $K^{(N)}$ only contains terms whose Q dependence is of the kind $\cos(kQ - \omega t + \varphi_m)$, with $\omega/k = p^*$. Using a physical language we will say that $K^{(N)}$ only contains waves whose phase velocities v_φ are such that $v_\varphi = p^*$. As long as the remainder $R^{(N)}$ is negligible, only those waves play an important role for the dynamics of $H^{(N)}$. As for the remainder $R^{(N)}$, it is proven an estimate of the form $|R^{(N)}(Q, P)| \leqslant E \exp(-1/\varepsilon^\delta)$, where δ mainly depends on the dimension of the dynamics.

In this paper we show that it is possible to define a canonical change of variables $(q, p) \mapsto (Q, P)$ transforming the Hamiltonian (1) in the Hamiltonian

$$\hat{H}(Q, P) = \frac{P^2}{2} + \varepsilon K(Q, P) + R(Q, P, \varepsilon) \tag{4}$$

which is thus of the same type as (3). In the case of the Hamiltonian (1), if we focus on values of P close to a given P^*, then imposing the condition $\nu \cdot \omega(P^*) = 0$ for K would imply that K only contains waves whose phase velocity is exactly P^*. However, this cannot be reached in the case of large perturbations, and one has to include in K all the waves whose phase velocities v_φ are such that $|v_\varphi - P^*| \leqslant \Delta v$, where Δv is chosen so that the remainder $R(Q, P, \varepsilon)$ would be small. Keeping these nearby waves in K insures that during the perturbation analysis the denominators $1/(v_\varphi - P)$ would be large enough to compensate the large value of the wave amplitude A.

As for the remainder $R(Q, P, \varepsilon)$ we do not try, as in Nekhoroshev theory, to find an upper bound for the norm of $R(Q, P, \varepsilon)$. Indeed, since we are only interested in the statistical properties of the dynamics defined by (1), we only provide a statistical estimate of the remainder $R(Q, P, \varepsilon)$, and we prove that $\sqrt{\langle R^2 \rangle} < 5\varepsilon^{[1/(2^{9/8}\varepsilon^{1/4})]}$, where $\langle \cdot \rangle$ means an averaging over the initial phases φ_m's. Here ε is the small parameter we have to introduce in order to perform a perturbation analysis, and is defined by

$$\varepsilon = \frac{A}{\Delta v^{3/2}} \tag{5}$$

Therefore, for ε and the remainder R to be indeed small regardless of the value of A, Δv has to be proportional to $A^{2/3}$. This explains why the range of interaction in phase velocity between the particle and the waves is proportional to $A^{2/3}$.

The property of locality is then deduced from the results of perturbation theory by showing that the Hamiltonian H' (2) can also be transformed in a Hamiltonian $\hat{H}'$ of the same type as (4), and by showing that, whatever the number $(2M+1)$ of waves present in (1), the statistical properties of $\hat{H}$ and $\hat{H}'$ can be made arbitrarily close by choosing $\Delta v_R = \alpha A^{2/3}$, with α large enough. The principle of the derivation of locality is sketched in Fig. 1, which will be further discussed in the next section. A value $\alpha \approx 5$ is analytically estimated and numerically checked to give a good approximation of the full chaotic transport.

We already noticed that the property of locality implied that the statistical properties of the dynamics defined by (1) could be derived from the non-deterministic dynamics of (2). This result is actually the cornerstone of a theory explaining the origin of diffusion in Hamiltonian dynamics. This theory, though not completely rigorous, is supported by very detailed numerical computations. The corresponding developments are far too lengthy to be reported in this paper, but can be found in refs. 6–8. Basically, the idea of the derivation of a diffusion equation rests on the fact that, as far as the chaotic motion of the dynamics defined by (2) is concerned, a particle may be considered as acted upon by a series of statistically independent sets of waves, each with an extension $2\alpha A^{2/3}$ in phase velocity. This allows the use of a central limit argument to describe the evolution of the particle velocity distribution function.

The property of locality presented here is reminiscent of that obtained in the socalled Resonance Broadening Theory[9] derived in the framework of plasma physics. However, this theory assumes the dynamics to be diffusive, and the diffusion to occur immediately. This second assumption has been shown in refs. 6–8 to be wrong in the case of the dynamics defined by (1). In the present paper no assumption is made on the statistical properties of this dynamics. Moreover, we show in Section 8 that the scaling given in ref. 9 for the range of perturbations is in general not correct.

Locality was assumed in ref. 10 in order to recover analytically the superquasilinear regime numerically observed in the chaotic dynamics defined by (1) for intermediate values of A. Numerical calculations performed in ref. 11 confirmed this assumption, but gave indications neither about the origin of locality, nor about the range of interaction in phase velocity between the particle and the waves.

The paper is organised as follows. Section 2 states the mathematical results, as well as their physical interpretation, in order to provide the reader with an overview of the new perturbation theory, and of the derivation of locality. Section 3 explains the major steps of the new perturbation theory, and provides the scaling $A^{2/3}$. The technical details are carried over to appendices of the paper. Section 4 derives locality by using the abovementioned perturbation theory. Section 5 provides a heuristic analytical calculation of the minimal order of magnitude of Δv_R. Section 6 confirms through numerical calculations the property of locality and the minimal value of Δv_R analytically estimated. Section 7 proves that A may be scaled out in the reduced dynamics on a finite time for A large enough. Section 8 sketches how the results derived for (1) could generalize to a wider class of Hamiltonians. Section 9 summarizes the results and concludes the paper.

2. MAIN MATHEMATICAL RESULTS AND PHYSICAL PICTURES

In this section we state the major steps of the derivation of locality, without proving them, in order to provide the reader with an overview of the derivation. Our main mathematical result, on which the derivation of locality is based, is the following

Theorem 2.1. For any value of $\Delta v \geqslant (2^{3/8}/e)$, and for any $A \leqslant \Delta v^{3/2}$, there exists a canonical change of variables $(q, p) \mapsto (Q, P)$, that can be defined around any value P^* of the particle's velocity, which transforms the Hamiltonian $H = p^2/2 + A \sum_{m=-M}^{M} \cos(q - mt + \varphi_m)$ into $\hat{H}(Q, P) = p^2/2 + \varepsilon K(Q, P) + R(Q, P, \varepsilon)$, where $\varepsilon = A/\Delta v^{3/2}$, and where K only contains waves with phase velocities v_φ such that $|v_\varphi - P^*| \leqslant \Delta v$, or terms which do not depend on Q and which oscillate with an angular frequency less than Δv in absolute value. As for the remainder $R(Q, P, \varepsilon)$, it can be estimated by $\sqrt{\langle R^2 \rangle} < 5\varepsilon^{[1/(2^{9/8}\varepsilon^{1/4})]}$, where $\langle \cdot \rangle$ means an averaging over the initial phases φ_m's.

This theorem implies that if, for any given A, one chooses Δv proportional to $A^{2/3}$, in the new variables (Q, P) the dynamics defined by (1) can be considered as a reduced dynamics which only includes waves with phase velocities v_φ such that $|v_\varphi - P^*| \leqslant \Delta v$, as is schematized in the top part of Fig. 1. This figure is schematic because A is large enough to allow an overlap of the primary resonances. Moreover, because the initial phases φ_m's are not all the same, the primary resonances are actually shifted the ones with respect to the others. However, Fig. 1 clearly shows that $\hat{H}$ has not been obtained by simply eliminating the waves of H whose phase velocities m are such that $|m - P^*| > \Delta v$, but that $\hat{H}$ actually results from the nonlinear interaction of all the waves present in H. This is illustrated in Fig. 1 by the presence of chains of 2 and 4 islands in the resonances of $\hat{H}$. Therefore, the previous theorem proving that the Hamiltonian H can be transformed in $\hat{H}$ does not prove by itself the property of locality.

In order to prove the property of locality by using the results of the perturbation analysis, we first study the dynamics defined by (1) in the variables (Q, P). We consider the dynamics of (1) on a time interval $[0, t]$ which we divide in smaller time intervals I_j on each of which we actually perform the change of variables $(q, p) \mapsto (Q, P)$ about a given value P_j. The proximity of $P(t)$ to P_j can be controlled on any of the I_j's for a set of phases whose normalized measure can be made arbitrarily close to 1 by decreasing ε. We then consider that, as regards the statistical properties of the dynamics, this result is enough to proceed as though the proximity of $P(t)$ to P_j could be controlled for any phase realization. This is the only

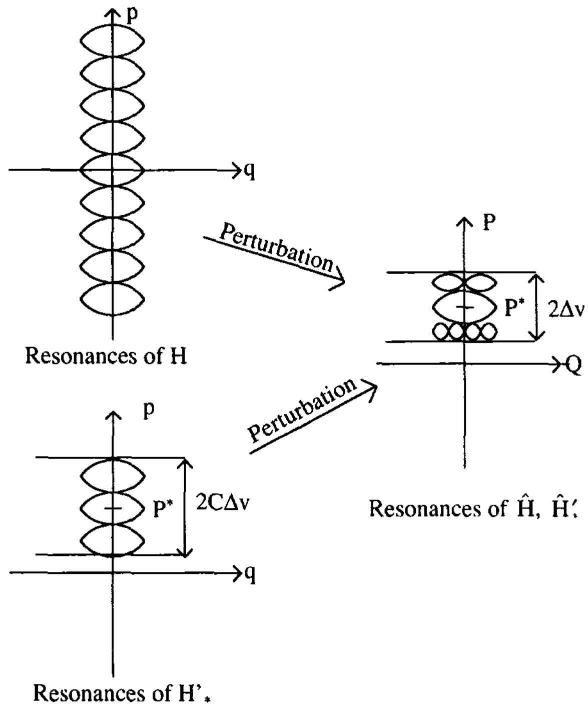

Fig. 1. Sketch of the derivation of the property of locality.

non-rigorous point of the derivation of locality. On each I_j the dynamics of (1) is given in the variables (Q, P) by a Hamiltonian of the kind of (4) where, as regards the statistical properties, the remainder is proven to be negligible for ε small enough and as long as $P(t)$ is close enough to P_j. In other words, the statistical properties of the dynamics of (1) in the variables (Q, P) can be deduced from the study of a sequence of Hamiltonians of the type

$$\hat{H}_j = P^2/2 + \varepsilon K_j(Q, P) \tag{6}$$

where K_j only contains waves with phase velocities v_φ such that $|v_\varphi - P_j| \leqslant \Delta v$.

We then proceed in a similar way for the dynamics defined by (2). We divide the time interval $[0, t]$ in smaller intervals I'_j on each of which $p(t)$ remains close to a given value p_j. Moreover, it is always possible to choose the intervals I_j and I'_j, in the case of (1) and (2) respectively, such that $I_j = I'_j$, which we actually do. On each I_j we perform the change of variables

$(q, p) \mapsto (Q, P)$ about $P = p_j$, and we can prove that the normalized measure of the set of phases such that $P(t)$ remains close to p_j on any interval I_j can be made arbitrarily, close to 1 by decreasing ε. As in the case of H, we proceed as though the proximity of $P(t)$ to p_j could be controlled for any phase realization. Then, on each I_j, the dynamics of (2) is described in the variables (Q, P) by a Hamiltonian of the same kind as (4). Moreover, as in the case of H, when ε is small enough, and when $P(t)$ is close enough to p_j, it is proven that the remainder can be neglected. Therefore, on each I_j the statistical properties of the dynamics of (2) in the variables (Q, P) can be deduced from a Hamiltonian of the type $\hat{H}'_j = P^2/2 + \varepsilon K'_j(Q, P)$, where K'_j only contains waves with phase velocities v_φ such that $|v_\varphi - p_j| \leqslant \Delta v$. Moreover, we prove that when Δv_R in (2) is chosen such that $\Delta v_R = C\,\Delta v$, the statistical properties of K'_j and K_j can be made arbitrarily close simply by increasing C, regardless of the values of A and M in (1), as long as $|M \pm p(t)|$ remains large enough. This implies that in the variables (Q, P) the statistic properties of the dynamics defined by (1) and (2) can be made arbitrarily close. This corresponds to the right part of Fig. 1. Moreover, we prove that the changes of variables $(q, p) \mapsto (Q, P)$ defined from (1) and (2) can also be made arbitrarily close by increasing C, which implies that in the physical variables (q, p) the statistical properties of (1) and (2) can be made arbitrarily close by choosing $\Delta v_R = \alpha A^{2/3}$ with α large enough. This concludes the derivation of locality.

3. THE THEORY OF LARGE PERTURBATIONS

In this section we prove that the Hamiltonian (1) can be reduced into the Hamiltonian (4), i.e., Theorem 2.1.

3.1. Reduction of the Width in Phase Velocity of the Wave-Spectrum

As a first step of the derivation of Theorem 2.1, we prove that it is possible to transform (1) into

$$\hat{H}(Q, P, \tau) = \frac{P^2}{2} + \sum_{i=1}^{n} \varepsilon^i h_i(Q, P, \tau) + R_n(Q, P, \tau) \tag{7}$$

where each h_i has the reduced width $2\,\Delta v$ in phase velocity. Moreover, we prove the existence of a constant, ρ_n, independent of A and M, such that $\sqrt{\langle R_n \rangle} \leqslant \varepsilon^{n+1} \rho_n$. Therefore, we do show that for ε small enough $\hat{H}$ can be considered as having a reduced spectrum, but we do not give yet an exponential estimate for the remainder.

The possibility to reduce the width in phase velocity of the wave spectrum of H to $2\,\Delta v$ indicates that Δv is the relevant scale of velocity in our problem. Therefore, it is natural to measure velocities its units Δv. Hence, we perform the change of variable $p_1 = p/\Delta v$. Since we keep $q_1 = q$, the time scale has to be changed in $\tau = t\,\Delta v$ for q_1 and p_1 to be conjugates. In the variables q_1, p_1, τ, the dynamics of (1) is given by the Hamiltonian $H_1 = p_1^2/2 + (A/\Delta v^2) \sum_{m=-M}^{M} \cos(q_1 - m\tau/\Delta v + \varphi_m)$. Using definition (5) of ε, H_1 can be written

$$H_1 = \frac{p_1^2}{2} + \varepsilon f(q_1, \tau) \tag{8}$$

where

$$f(q_1, \tau) = \frac{1}{\sqrt{\Delta v}} \sum_{m=-M}^{M} \cos\left(q_1 - \frac{m\tau}{\Delta v} + \varphi_m\right) \tag{9}$$

Starting from the Hamiltonian (7), we perform a canonical change of variables, $(q_1, p_1) \mapsto (Q, P)$, about a given value P^*, using the generating function $\Phi(q_1, P, \tau) = Pq_1 + \Phi'(q_1, P, \tau)$, where Φ' is expanded in power series of ε

$$\Phi' = \sum_{i=1}^{n} \varepsilon^i \Phi_i \tag{10}$$

The power series of Φ' stops at a finite order n, because we expect this series to be divergent as in Nekhoroshev theory. As we are only interested in the statistical properties of the dynamics of (1), we estimate the terms of the series of Φ' by their variances, as a function of the phases φ_m's. Therefore, $\varepsilon^{2i}\langle \Phi_i^2 \rangle$ is expected to diverge when i goes to infinity. In the next subsection, we will actually find an upper bound for the variances of the terms of the perturbation series (10) in terms of a Gevrey series.[12] The new variables (Q, P) are defined by

$$P = p_1 - \frac{\partial \Phi'(q_1, P, \tau)}{\partial q_1}, \qquad Q = q_1 + \frac{\partial \Phi'(q_1, P, \tau)}{\partial P} \tag{11}$$

and in the variables (Q, P, τ) H_1 is transformed in

$$\hat{H} = H_1 + \frac{\partial \Phi'(q_1, P, \tau)}{\partial \tau} \tag{12}$$

The change of variables (11) is aimed at transforming the Hamiltonian (8) in the Hamiltonian (7) where the h_i's only contain waves whose phase velocities v_φ, are such that $|v_\varphi - P^*| \leqslant 1$ (we now measure velocities in units Δv), or terms which do not depend on Q and which oscillate with an angular frequency which is less than 1 in absolute value. Using (11) and (12), we find that in order to transform (8) in (7) one has to solve the equation

$$\frac{1}{2}\left(\frac{\partial \Phi'}{\partial q_1}\right)^2 + P\frac{\partial \Phi'}{\partial q_1} + \varepsilon f(q_1, \tau) + \frac{\partial \Phi'}{\partial \tau} = \sum_{i=1}^{n} \varepsilon^i h_i\left(q_1 + \frac{\partial \Phi'}{\partial P}, P, \tau\right) + R_n(Q, P, \tau) \tag{13}$$

By expanding $(\partial \Phi'/\partial q_1)^2$ and $\sum_{i=1}^{n} \varepsilon^i h_i(q_1 + \partial \Phi'/\partial P, P, \tau)$ in power series of ε, and by equating the terms of same power of ε in (13), one finds that for any i such that $1 \leqslant i \leqslant n$,

$$P\frac{\partial \Phi_i}{\partial q_1} + \frac{\partial \Phi_i}{\partial \tau} = h_i(q_1, P, \tau) + X_i \tag{14}$$

where

$$X_1 = -f(q_1, \tau)$$

$$X_i = \sum_{l=1}^{i-1} \sum_{m=1}^{i-l} \frac{\partial^m h_l}{\partial q_1^m} \frac{1}{m!} \sum_{i_1} \cdots \sum_{\{1, i-l\}} \sum_{i_m} \prod_{j=1}^{m} \frac{\partial \Phi_{i_j}}{\partial P} - \frac{1}{2} \sum_{j=1}^{i-1} \frac{\partial \Phi_j}{\partial q_1} \frac{\partial \Phi_{i-j}}{\partial q_1}, \qquad i \geqslant 2 \tag{15}$$

where in (15), the symbol $\{1, i-l\}$ means that the i_j's are such that $i_1 \geqslant 1, \ldots, i_m \geqslant 1$ and $\sum_{i=1}^{l} i_m = i - l$.

Equation (14) is solved by choosing h_i equal to the sum of the terms in $-X_i$ whose phase velocities are such that $|v_\varphi - P^*| \leqslant 1$, or which do not depend on q_1 and which oscillate with an angular frequency less than 1 in absolute value. By construction, h_i is of the desired form. As for the other terms of X_i, they are eliminated by an appropriate choice of Φ_i, which is always possible when P is close enough to P^*.

In order to illustrate our procedure, we explicitly calculate the first order terms. When $i = 1$, (14) writes

$$P\frac{\partial \Phi_1}{\partial q_1} + \frac{\partial \Phi_1}{\partial \tau} = h_1(q_1, P, \tau) - \frac{1}{\sqrt{\Delta v}} \sum_{m=-M}^{M} \cos\left(q_1 - \frac{m\tau}{\Delta v} + \varphi_m\right) \tag{16}$$

We thus choose

$$h_1(Q, P, \tau) = \frac{1}{\sqrt{\Delta v}} \sum_{|m/\Delta v - P^*| \leqslant 1} \cos(Q - m\tau/\Delta v + \varphi_m) \tag{17}$$

and then, solving for Φ_1 yields

$$\Phi_1(Q, P, \tau) = \frac{1}{\sqrt{\Delta v}} \sum_{|m/\Delta v - P^*| > 1} \frac{\sin(Q - m\tau/\Delta v + \varphi_m)}{m/\Delta v - P} \tag{18}$$

We estimate the terms of the perturbation series through their root mean square as a function of the φ_m's. We calculate here $\sqrt{\langle h_i^2(Q, P, \tau)\rangle}$ and $\sqrt{\langle \Phi_i^2(Q, P, \tau)\rangle}$ for fixed values of Q and P. At first order, in the case when $\Delta v \geqslant 1$, (which is the only relevant case), one finds

$$\sqrt{\langle h_1^2\rangle} = (1 + 1/2\,\Delta v)^{3/2} \leqslant \sqrt{3/2} \tag{19}$$

Therefore, when $\Delta v \geqslant 1$, one can find an upper bound for $\sqrt{\langle h_1^2\rangle}$ which is independent of A, M, and Δv. As for Φ_1, we find

$$\langle \Phi_1^2\rangle = \frac{\Delta v}{2} \sum_{|m - \Delta vP^*| > \Delta v} \frac{1}{(m - \Delta vP)^2} \tag{20}$$

In the next subsection, where we give an exponential estimate for the remainder, we need to consider complex values of P. Therefore, we give an upper bound for $\sqrt{|\langle \Phi_1^2\rangle|}$ on the disk D_r, centered on P^*, and of radius $(1-r)$, $0 < r < 1$. On D_r

$$\begin{aligned} |\langle \Phi_1^2\rangle| \leqslant \frac{\Delta v}{2}\Bigg(&\sum_{m = \mathrm{Int}(\Delta vP^*) + \Delta v + 1}^{M} \frac{1}{(m - \Delta vP^* - \Delta v(1-r))^2} \\ &+ \sum_{m=-M}^{\mathrm{Int}(\Delta vP^*) - \Delta v - 1} \frac{1}{(m - \Delta vP^* + \Delta v(1-r))^2}\Bigg) \end{aligned} \tag{21}$$

where $\mathrm{Int}(x)$ denotes the integer part of x. Now, for any $\mu \geqslant 1$,

$$\sum_{m=0}^{M} \frac{1}{(m+\mu)^2} \leqslant \frac{\pi^2}{6} \int_0^{+\infty} \frac{dx}{(x+\mu)^2} \tag{22}$$

This is due to the fact that the discrepancy between the sum and the integral decreases with μ and that the sum is equal to the integral when $\mu = 1$ and $M = +\infty$. Using (21) and (22), we find that if

$$r\,\Delta v \geqslant 1 \tag{23}$$

(which is always possible if $\Delta v > 1$), then on the disk D_r

$$\sqrt{|\langle \Phi_1^2 \rangle|} \leqslant \frac{\pi}{\sqrt{6r}} \tag{24}$$

Therefore, it is also possible to find an upper bound for $\sqrt{|\langle \Phi_1^2 \rangle|}$ which is independent of A, M, and Δv. This leads to prove the

Proposition 3.1. On the disk D_r, r fulfilling condition (23), and for any i such that $1 \leqslant i \leqslant n$, $\sqrt{|\langle \Phi_i^2 \rangle|}$ and $\sqrt{|\langle h_i^2 \rangle|}$ can be bounded from above by quantities independent of A, M, and Δv.

Proof. Here are only given the main steps of the proof. The details are carried over in Appendix A. It is easy to show, by induction, that the generating functions Φ_i are of the form

$$\Phi_i(Q, P, \tau) = \sum_{l=1}^{i} S_l(Q, P, \tau) \tag{25}$$

were

$$S_l = \frac{1}{\Delta v^{i/2}} \sum_{\varepsilon_1} \cdots \sum_{\varepsilon_i} \sum_{m_1} \cdots \sum_{m_i} \sum_{j=1}^{n_l} \frac{s_{j,l} \cos[\sum_{k=1}^{i} \varepsilon_k \xi_k + v_{j,l}(\pi/2)]}{\prod_{k=1}^{l} (m_k/\Delta v - P)^{\alpha_{j,k,l}} \Delta_{j,l}(\overline{m_{1,i}}/\Delta v, \overline{\varepsilon_{1,i}}, P)} \tag{26}$$

where

$\alpha_{j,k,l} \geqslant 1$

$\xi_k = Q - m_k \tau/\Delta v + \varphi_{m_k}$

$\varepsilon_1 = 1$ and $\varepsilon_k \in \{-1; 1\}$ if $2 \leqslant k \leqslant i$

$\overline{m_{1,i}}$ stands for the i-tuple $(m_1, \ldots, m_i)$, and $\overline{m_{1,i}}/\Delta v$ stands for the i-tuple $(m_1/\Delta v, \ldots, m_i/\Delta v)$,

the m_k's are such that $|m_k| \leqslant M$, and $|m_k/\Delta v - P^*| > 1$ if $1 \leqslant k \leqslant l$, and $|m_k/\Delta v - P^*| \leqslant 1$ if $l+1 \leqslant k \leqslant i$,

the $\Delta_{j,l}$ are polynomials in P which are defined in Appendix A. They can be bounded from below on D_r, independently of A, M, and Δv, whatever $\overline{m_{1,i}}/\Delta v$ and $\overline{\varepsilon_{1,i}}$.

The form (25)–(26) can be easily understood. Indeed, Φ_i is a sum of terms which are i-linear in f (9). f being a sum of cosines, f^i is an "i-multiple" sum of products of cosines. This explains the origin of the sum $\sum_{m_1} \cdots \sum_{m_i}$ in (26). Because a product of cosines can be expressed as the sum and the difference of the arguments, it is natural to find in Φ_i all the possible sums and differences of the ξ_k's, which explains the presence of the ε_k's in (26). Finally, when calculating Φ_1, we showed that treating perturbatively a wave with phase velocity v_φ yields the denominator $1/(v_\varphi - P)$. These very denominators can be found in (26). According to (A3), denominators of the form $(\sum_{k_2=1}^{i} \mu_{k_2} m_{k_2}/\Delta v)^{\chi_{n_2}}$ can also be found in Φ_i. They arise when treating perturbatively terms independent of q_1 and oscillating with the angular frequency $\sum_{k_2=1}^{i} \mu_{k_2} m_{k_2}/\Delta v$. By construction, all the denominators present in Φ_i are strictly larger than unity in absolute value when $P = P^*$. Indeed, the terms whose phase velocities v_φ are such that $|v_\varphi - P^*| \leqslant 1$, or which do not depend on q_1 and which oscillate with an angular frequency lower than unity in absolute value, are not treated perturbatively but are kept in the Hamiltonians h_i. These Hamiltonians are also of the form (25)–(26), except h_1 which actually only contains the term $l = 0$. Let us now define

$$g_{j,l}(\overline{m_{1,i}}/\Delta v, \overline{\varepsilon_{1,i}}, P) = \frac{s_{j,l}}{\prod_{k=1}^{l} (m_k/\Delta v - P)^{\alpha_{j,k,l}} \Delta_{j,l}(\overline{m_{1,i}}/\Delta v, \overline{\varepsilon_{1,i}}, P)} \tag{27}$$

then in order to estimate $\langle h_i^2 \rangle$ when $i \geqslant 2$, or $\langle \Phi_i^2 \rangle$, one needs to calculate

$$\begin{aligned} \langle S_{l_1} S_{l_2} \rangle = {} & \frac{1}{\Delta v^i} \sum_{\varepsilon_1} \cdots \sum_{\varepsilon_{2i}} \sum_{m_1} \cdots \sum_{m_{2i}} \sum_{j_1=1}^{n_{l_1}} \sum_{j_2=1}^{n_{l_2}} g_{j_1, l_1} g_{j_2, l_2} \\ & \times \left\langle \cos \left[\sum_{k=1}^{2i} \varepsilon_k \xi_k + (\nu_{j_1, l_1} + \nu_{j_2, l_2}) \pi/2 \right] \right\rangle \end{aligned} \tag{28}$$

Because the φ_k's are random phases, the only terms which may be non-zero in (28) are such that $\sum_{k=1}^{2i} \varepsilon_k \varphi_{m_k} = 0$. This last condition is fulfilled only if $\sum_{k=1}^{2i} \varepsilon_k = 0$, and if there is a one-to-one relation between the indices m_k such that $\varepsilon_k = 1$, and the indices $m_{k'}$ such that $\varepsilon_{k'} = -1$. In the sum (28) there are thus only i independent m_k's. By renumbering the indices, one can choose the independent m_k's to be $m_1, \ldots, m_i$. Therefore, there exists a function $a_{l_1, l_2}(\overline{m_{1,i}}/\Delta v, P)$ such that

$$|\langle S_{l_1} S_{l_2} \rangle| \leqslant \frac{1}{\Delta v^i} \sum_{\varepsilon_1} \cdots \sum_{\varepsilon_{2i}} \sum_{m_1} \cdots \sum_{m_i} |a_{l_1, l_2}(\overline{m_{1,i}}/\Delta v, P)| \tag{29}$$

Because the polynomials $\Delta_{j,l}$ can be bounded from below on the disk D_r, one can show that there exists a function $C(r)$ such that for any $P \in D_r$

$$|a_{l_1, l_2}(\overline{m_{1,i}}/\Delta v, P)| \leqslant \frac{C(r)}{\prod_{k=1}^{(l_1+l_2)/2} (m_k/\Delta v - P)^{2\alpha_k}} \tag{30}$$

The sum $1/\Delta v^i \sum_{\varepsilon_1} \cdots \sum_{\varepsilon_{2i}} \sum_{m_1} \cdots \sum_{m_i} |a_{l_1, l_2}(\overline{m_{1,i}}/\Delta v, P)|$ is thus less than a series which converges when $M \to +\infty$. This implies that $|\langle S_{l_1} S_{l_2} \rangle|$ can be bounded from above by a quantity independent of M. Actually, the sum (29) is nothing but a Riemann sum. Comparing this sum to its integral implies that for Δv large enough, for example $\Delta v > 1$, (29) is less than a quantity which does not depend on Δv. Finally, because (29) is independent of A, $|\langle S_{l_1} S_{l_2} \rangle|$, and thus $\langle \Phi_i^2 \rangle$ and $\langle h_i^2 \rangle$, are bounded from above by a quantity independent of A, M and Δv. This proves Proposition 3.1.

The optimality of the estimates for $\langle \Phi_i^2 \rangle$ and $\langle h_i^2 \rangle$ is simple to show. Indeed, if Δv is large enough, the Riemann sum is close to its integral, and if $|M/\Delta v - P^*|$ is large enough, the sum (29) is close to its limit when $M \to +\infty$. Therefore, when Δv and $|M/\Delta v - P^*|$ are large, $\langle \Phi_i^2 \rangle$ and $\langle h_i^2 \rangle$ weakly depend on M and Δv. This implies that the terms of order i in the perturbation series are of order ε^i and that $\varepsilon = A/\Delta v^{3/2}$ is indeed the right small parameter to make the perturbation analysis with.

Actually, we proved a result a bit stronger than Proposition 3.1, because we actually proved that the modulus of the variance of any function of the form (25)–(26) is less than a quantity independent of A, M and Δv. This allows us to prove the

Proposition 3.2. On the interval $[P^* - (1-r), P^* + (1-r)]$, r fulfilling condition (23), the root mean square of the remainder $R_n(Q, P, \tau)$ in (7) is such that $\sqrt{\langle R_n^2 \rangle} \leqslant \varepsilon^{n+1} \rho_n$, where ρ_n is independent of A, M and Δv.

Proof. By construction, the variance of R_n is analytic in ε, and its series expansion in terms of ε begins by ε^{2n+2}. This implies that if $\varepsilon \in [0; \varepsilon_0]$, then

$$\langle R_n^2(\varepsilon) \rangle \leqslant (\varepsilon/\varepsilon_0)^{2n+2} \sup_{\varepsilon' \leqslant \varepsilon_0} \langle R_n^2(\varepsilon') \rangle \tag{31}$$

Note that in order for the remainder to be analytic in ε, ε must be varied independently of Δv. Hence, when performing the perturbation analysis, Δv is assigned a fixed value, while ε is varied by varying A. Then, if a result is valid only when $\varepsilon \leqslant \varepsilon_0$, this implies that the result holds only when $A \leqslant \varepsilon_0(\Delta v)^{3/2}$. This does not mean that we are restricted to small values of A, because in order to derive a result for given value A_0, one has to choose, *a priori* $\Delta v \geqslant (A_0/\varepsilon_0)^{2/3}$.

In order to prove Proposition 3.2, one only has to prove that on the interval $[P^* - (1-r), P^* + (1-r)]$, and whatever the value of ε, $\sqrt{\langle R_n^2 \rangle}$ is less than a quantity ρ_n independent of A, M and Δv. This part of the proof is detailed in Appendix B. We explain however the most difficult step. Using (13) and (14), one can see that the remainder R_n can actually be written as

$$R_n(Q, P, \tau) = \mathfrak{R}_\alpha(q_1, P, \tau) + \mathfrak{R}_\beta(Q, P, \tau) \tag{32}$$

where

$$\mathfrak{R}_\alpha(q_1, P, \tau) = \frac{1}{2}\left(\frac{\partial \Phi'}{\partial q_1}\right)^2 + \sum_{i=1}^{n} \varepsilon^i h_i(q_1, P, \tau) + \sum_{i=2}^{n} \varepsilon^i X_i(q_1, P, \tau) \tag{33}$$

$$\mathfrak{R}_\beta(Q, P, \tau) = -\sum_{i=1}^{n} \varepsilon^i h_i(Q, P, \tau) \tag{34}$$

Therefore R_n is the sum of a function of the old variable q_1 and of a function of the new variable Q. If the expression of q_1 as a function of Q, given by (11), is used in (33)–(34), then one does not obtain a function of the form (25)–(26) for R_n. Therefore, Proposition 3.1 cannot be used directly to prove Proposition 3.2. However, it can be shown that there exists a function of the form (25)–(26) whose root mean square is larger than R_n's one.

Equation (26) shows that $\mathfrak{R}_\alpha(Q - \partial\Phi'/\partial P, P, \tau)$ can be written in the symbolic way

$$\mathfrak{R}_\alpha(Q - \partial\Phi'/\partial P, P, \tau) = \sum r_\alpha \cos(\tilde{\xi} - \tilde{\varepsilon}\, \partial\Phi'/\partial P) \tag{35}$$

where we have used the notation $\tilde{\varepsilon} = \sum_{k=1}^{i} \varepsilon_k$, and $\tilde{\xi} = \sum_{k=1}^{i} \varepsilon_k \xi_k$. Expanding the cosines in (35) and isolating the terms having the same values of $\tilde{\varepsilon}$ leads to

$$R_n = \mathfrak{R}_\beta + \sum_{i=0}^{2n} \cos\left(i \frac{\partial \phi'}{\partial P}\right) \sum_{|\tilde{\varepsilon}| = i} r_\alpha \cos \tilde{\xi} + \sum_{i=1}^{2n} \sin\left(i \frac{\partial \phi'}{\partial P}\right) \times \left[\sum_{\tilde{\varepsilon} = 1} r_\alpha \sin \tilde{\xi} - \sum_{\tilde{\varepsilon} = -i} r_\alpha \sin \tilde{\xi}\right] \tag{36}$$

Using twice the inequality $xy \leqslant (x^2+y^2)/2$, it is shown in Appendix B that

$$\langle R_n^2(Q, P, \tau)\rangle \leqslant (6n+3)\langle \mathfrak{R}_\alpha^2(Q, P, \tau)\rangle + 6n\langle \mathfrak{R}_\alpha'^2(Q, P, \tau)\rangle + 3\langle \mathfrak{R}_\beta^2(Q, P, \tau)\rangle \tag{37}$$

where

$$\mathfrak{R}_\alpha'^2 = \sum_{i=1}^{2n} \left[\sum_{\tilde{\xi}=i} r_\alpha \sin \tilde{\xi} - \sum_{\tilde{\xi}=-i} r_\alpha \sin \tilde{\xi} \right]^2 \tag{38}$$

Note that the inequality $xy \leqslant (x^2+y^2)/2$ is only true if x and y are real numbers, which explains why we only give an upper bound of $\langle R_n^2\rangle$ on the interval $[P^*-(1-r), P^*+(1-r)]$ and not on the whole disk D_r.

All the terms on the right hand side of (37) are of form (25)–(26). Therefore, from Proposition 3.1 it follows that on the interval $[P^*-(1-r), P^*+(1-r)]$, $\sqrt{\langle R_n^2\rangle}$ is less than a constant ρ_n independent of A, M and Δv, which proves Proposition 3.2.

This proposition implies that whatever the number of waves in (1) and whatever their amplitude A, by choosing *a priori* Δv such that $\varepsilon = A/\Delta v^{3/2}$ would be small enough, one can make the variance of the remainder arbitrarily small. This implies that regardless of the width in phase velocity in (1), through the change of variables $(q_1, p_1) \mapsto (Q, P)$, one can effectively reduce the width in phase velocity of the Hamiltonian H_1 to 2, in units Δv.

At this point, one would like to have a more explicit estimate for the remainder than the one given in Proposition 3.2, in order to have an idea of how the remainder actually decreases with ε. This is done in the next section, where, as announced in Theorem 2.1, we prove that the variance of the remainder decreases exponentially with ε. We thus give an estimate *à la* Nekhoroshev[3,4,5] for the remainder.

3.2. Exponential Estimate of the Remainder

3.2.1. Upper Bounds for the Variances. In order to derive an exponential estimate for the remainder, one needs to estimate the variances of the terms of the perturbation series in a more precise way than in Appendix A. Because h_i and Φ_i are expressed in terms of the h_l's and Φ_l's, with $l \leqslant i-1$, it is natural to try to estimate the variances of the terms of the perturbation series by induction. This is however not straightforward because the variances of the sum or products of the h_l's and Φ_l's cannot be simply expressed in terms of $\langle h_l^2\rangle$ and $\langle \Phi_l^2\rangle$. Therefore, the knowledge of an upper bound for the variances of the h_l's or the Φ_l's does not give, *a priori*, any indication of how to derive an upper bound for h_i and Φ_i. This is why, for any function B of the form (25)–(26), we introduce the quantity $\overline{\langle B^2 \rangle}$ which is such that $\overline{\langle B^2 \rangle} \geqslant \langle B^2 \rangle$, and which has simpler properties regarding arithmetic calculations than the variances themselves.

Let B be a function of the form (25)–(26), $B = \sum_{l=1}^{i} B_l(q_1, P, \tau)$. Then

$$\langle B^2 \rangle = \sum_{l_1=1}^{i} \sum_{l_2=1}^{i} \langle B_{l_1} B_{l_2} \rangle \tag{39}$$

As shown in Appendix A, in (39) only the terms such that $l_1 + l_2$ is even are non-zero. Let n_i be the number of these terms,

$$\begin{cases} n_i = i^2/2, & \text{if } i \text{ is even} \\ n_i = (i^2+1)/2, & \text{if } i \text{ is odd} \end{cases} \tag{40}$$

Using the same kind of notations as in Appendix A, one finds

$$\langle B_{l_1} B_{l_2} \rangle = \frac{1}{\Delta v^i} \sum_{\varepsilon_1} \cdots \sum_{\{0\}} \sum_{\varepsilon_{2i}} \sum_{m \in \overline{m_+}} \sum_{\sigma_1} \sum_{\sigma_2} f_{l_1}[\sigma(\overline{m_{1,i}}, P)]\, f_{l_2}[\sigma(\overline{m_{1+i,2i}}, P)] \tag{41}$$

where

$$f_{l_1}[\sigma(\overline{m_{1,i}}, P)] = \sum_{j_1=1}^{n_{l_1}} f_{j_1, l_1}[\sigma(\overline{m_{1,i}}, P)] \tag{42}$$

$$f_{j_1, l_1}[\sigma(\overline{m_{1,i}}, P)] = \frac{b_{j_1, l_1}}{\sqrt{2} \prod_{k_1=1}^{l_1} [\sigma(m_{k_1})/\Delta v - P]^{\alpha_{j_1, k_1, l_1}} \Delta_{j_1, l_1}[\sigma(\overline{m_{1,i}}, P)]} \tag{43}$$

and f_{l_2} is defined in a similar way. For any $k \leqslant n$, we denote by $D_{k,a}$ the disk centered on P^* and of radius $1 - k/\sqrt{a}$, $(a > n^2)$. It is then clear that the functions f_{l_1} et f_{l_2} are analytic on $D_{k,a}$. For any function which is analytic on this disk we define

$$\|f\|_k = \sup_{p \in D_{k,a}} |f(P)| \tag{44}$$

From (A4) we know that there exists an integer $\eta_{j,l}$ such that whatever $\overline{m_{1,2i}}$ and $\overline{\varepsilon_{1,2i}}$,

$$\left\| \frac{1}{\Delta_{j,l}} \right\|_k \leqslant \left(\frac{\sqrt{a}}{k} \right)^{\eta_{j,l}} \tag{45}$$

Therefore

$$\|f_{j_1,l_1}\|_k \leqslant \left(\frac{\sqrt{a}}{k}\right)^{\eta_{j,l}+\sum_{n=1}^{l_1}(\alpha_{j_1,n,l_1}-1)} \frac{|b_{j_1,l_1}|}{\sqrt{2}} \prod_{n=1}^{l_1} \left\|\frac{1}{\sigma(m_n)/\Delta v - P}\right\|_k \tag{46}$$

Let us now define $\Lambda = \{\lambda \geqslant 0/\forall \overline{m_{1,2i}},\ \|f_{j_1,l_1}\|_k \leqslant \lambda \prod_{n=1}^{l_1} \|1/(\sigma(m_n)/\Delta v - P)\|_k\}$ and, for a given realization of the ε_i's, let us denote $\lambda_{j_1,l_1}^{(k)}(\overline{\varepsilon_{1,i}}) = \inf(\Lambda)$. A similar notation is used for f_{l_2}. Then, by renumbering the indices so that $\overline{m_+} = \{m_1,..., m_i\}$, one obtains

$$\begin{aligned}\|\langle B_{l_1}B_{l_2}\rangle\|_k \leqslant \frac{1}{\Delta v^i} \sum_{\varepsilon_1} \cdots \sum_{\{0\}\ \varepsilon_{2i}} \sum_{m\in\overline{m_+}} \sum_{\sigma_1} \sum_{\sigma_2} \sum_{j_1=1}^{l_1} \sum_{j_2=1}^{l_2} \lambda_{j_1,l_1}^{(k)} \\ \times (\overline{\varepsilon_{1,i}})\, \lambda_{j_2,l_2}^{(k)}(\overline{\varepsilon_{1+i,2i}}) \prod_{n=1}^{(l_1+l_2)/2} \left\|\frac{1}{(m_n/\Delta v - P)^2}\right\|_k\end{aligned} \tag{47}$$

We then define

$$\lambda_l^{(k)} = \max_{\overline{\varepsilon_{1,i}}} \left(\sum_{j=1}^{l} \lambda_{j,l}^{(k)}(\overline{\varepsilon_{1,i}})\right) \tag{48}$$

$$\lambda_B^{(k)} = \max_l \lambda_l^{(k)} \tag{49}$$

$$\mu_{l_1,l_2} = \left(\sum_{|m/\Delta v - P^*| \leqslant 1} 1\right)^{i-(l_1+l_2)/2} \left(\sum_{|m/\Delta v - P^*|>1} \left\|\frac{1}{(m_k/\Delta v - P)^2}\right\|_k\right)^{(l_1+l_2)/2} \tag{50}$$

$$\mu_i^{(k)} = \max_{l_1,l_2} \mu_{l_1,l_2} \tag{51}$$

Then, because

$$\begin{aligned}\sum_{\varepsilon_1} \cdots \sum_{\{0\}\ \varepsilon_{2i}} \sum_{\sigma_1} \sum_{\sigma_2} 1 &= \binom{2i-(l_1+l_2)}{i-(l_1+l_2)/2} \binom{l_1+l_2-1}{(l_1+l_2)/2} \left(\frac{l_1+l_2}{2}\right)! \left(i - \frac{l_1+l_2}{2}\right)! \\ &\leqslant \binom{2i-1}{i} i!\end{aligned}$$

we finally find that, for any (l_1, l_2)

$$\|\langle B_{l_1}B_{l_2}\rangle\|_k \leqslant \frac{(\lambda_B^{(k)})^2 \mu_i^{(k)}}{\Delta v^i} \binom{2i-1}{i} i! \tag{52}$$

We then define

$$\overline{\langle B^2 \rangle_k} = n_i \binom{2i-1}{i} i! \frac{(\lambda_B^{(k)})^2}{\Delta v^i} \mu_i^{(k)} \tag{53}$$

It is then clear that for any P in $D_{i,a}$

$$|\langle B^2 \rangle| \leqslant \overline{\langle B^2 \rangle_i} \tag{54}$$

Estimating $\langle B^2 \rangle$ by $\overline{\langle B^2 \rangle_k}$ may seem, at first sight, a rough estimate. Actually, in (53), the term $n_i\binom{2i-1}{i}$ cannot be avoided, so that the less accurate estimate comes from $(\lambda_B^{(k)})^2 \mu_i^{(k)}$. However, because we evaluate the $\overline{\langle B^2 \rangle_k}$'s by induction, $(\lambda_B^{(k)})^2 \mu_i^{(k)}$ will never be explicitly calculated; except when $i=1$, in which case we obtain the estimates (19) and (24) which are quite accurate. Because h_1 is not of the form (25)–(26), we have to define

$$\overline{\langle h_1^2 \rangle_k} = \langle h_1^2 \rangle \leqslant 1 + 1/2\, \Delta v \tag{55}$$

As already mentioned, the advantage of the $\overline{\langle B^2 \rangle_k}$'s comes from the fact that it is very easy to make arithmetic calculations with them. Indeed, in Appendix C are shown the following properties

Property 1.

$$\forall l \geqslant 1, \qquad \overline{\left\langle \left(\frac{\partial B}{\partial P}\right)^2 \right\rangle_{k+l}} \leqslant \frac{a}{l^2} \overline{\langle B^2 \rangle_k} \tag{56}$$

The relation (56) is actually derived using the Cauchy inequalities for analytic functions (13), which write $\|\partial f/\partial P\|_{k+1} \leqslant \|f\|_k \sqrt{a}/1$.

Property 2. If $B_1, B_2, \ldots, B_n$ are of the form (25)–(26), then

$$\overline{\langle (B_1 + B_2 + \cdots + B_n)^2 \rangle_k} \leqslant \left(\sqrt{\overline{\langle B_1^2 \rangle_k}} + \sqrt{\overline{\langle B_2^2 \rangle_k}} + \cdots + \sqrt{\overline{\langle B_n^2 \rangle_k}}\right)^2 \tag{57}$$

This is shown by simply using the triangular inequality which implies that $\lambda_{B_1+B_2+\cdots+B_n} \leqslant \lambda_{B_1} + \lambda_{B_2} + \cdots + \lambda_{B_n}$. Therefore $\lambda^2_{B_1+B_2+\cdots+B_n} \leqslant (\sqrt{\lambda^2_{B_1}} + \sqrt{\lambda^2_{B_2}} + \cdots + \sqrt{\lambda^2_{B_n}})^2$, which yields (57).

Property 3. If C is of the form (25)–(26), i being replaced by j, or if C is one of the derivatives of h_1 with respect to q_1 (in which case $j=1$),

$$\overline{\langle (BC)^2 \rangle_k} \leqslant \frac{(i+j)^2}{[\max(i,j)]^2} \frac{\binom{2(i+j)-1}{i+j}(i+j)!}{\binom{2i-1}{i}\binom{2j-1}{j} i!\, j!} \overline{\langle B^2 \rangle_k \langle C^2 \rangle_k} \tag{58}$$

Moreover, it is clear that if $k > k'$, $\overline{\langle B^2 \rangle_k} \leqslant \overline{\langle B^2 \rangle_{k'}}$, therefore, if $k \geqslant i$ and $k \geqslant j$, then

$$\overline{\langle (BC)^2 \rangle_k} \leqslant \frac{(i+j)^2}{[\max(i,j)]^2} \frac{\binom{2(i+j)-1}{i+j}(i+j)!}{\binom{2i-1}{i}\binom{2j-1}{j} i!\, j!} \overline{\langle B^2 \rangle_i \langle C^2 \rangle_j} \tag{59}$$

The first factor in (59) only comes from the fact that the number of non-zero terms in $\langle (BC)^2 \rangle$ is larger than the product of non-zero terms in $\langle B^2 \rangle$ and $\langle C^2 \rangle$. Indeed the nonzero terms in $\langle B^2 \rangle$, and $\langle C^2 \rangle$, are of the form $\langle B_{l_1} B_{l_2} \rangle$, $l_1 + l_2$ even, respectively $\langle C_{l_3} C_{l_4} \rangle$, $l_3 + l_4$ even, and in $\langle (BC)^2 \rangle$ they are of the form $\langle B_{l_1} B_{l_2} C_{l_3} C_{l_4} \rangle$, $l_1 + l_2 + l_3 + l_4$ even. Now it is clear that $l_1 + l_2 + l_3 + l_4$ can be even while $l_1 + l_2$ and $l_3 + l_4$ are not. The second factor in (59) is obtained by replacing in (53) i by $(i+j)$. The last factor simply comes from the inequality $\|fg\|_k \leqslant \|f\|_k \|g\|_k$.

Property 4.

$$\overline{\left\langle \left(\frac{\partial^m B}{\partial Q^m} \right)^2 \right\rangle_k} \leqslant i^{2m} \overline{\langle B^2 \rangle_k} \tag{60}$$

Thanks to the equations (56) to (60) one can show the Gevrey nature of the perturbation series, from which the exponential estimate for the remainder follows.

3.2.2. Gevrey Nature of the Perturbation Series.

A Gevrey series is a series whose term of order i is less than, or of the order of, $\varepsilon^i (i!)^\alpha$. In order to illustrate the link between a Gevrey series and an exponentially small remainder, let us consider the function $g(\varepsilon) = \int_0^{+\infty} e^{-t/\varepsilon}/(1+t)$ and let us denote

$$g_n(\varepsilon) = \varepsilon - 1!\, \varepsilon^2 + 2!\, \varepsilon^3 + \cdots + (-1)^{n-1} (n-1)!\, \varepsilon^n \tag{61}$$

$$R_n(\varepsilon) = (-1)^n \int_0^{+\infty} \frac{t^n e^{-t/\varepsilon}}{1+t} dt \tag{62}$$

It is easy to see that $g(\varepsilon) = g_n(\varepsilon) + R_n(\varepsilon)$ and that $|R_n(\varepsilon)| \leqslant n!\,\varepsilon^{n+1}$. Therefore, the absolute value of the remainder R_n is less than the first omitted term in the expansion (61). The smallest term in the series (61) is for $n^* \approx 1/\varepsilon$ and its order of magnitude is, using the Stirling formula, $\sqrt{2\pi/\varepsilon}\, e^{-1/\varepsilon}$. Therefore, if the expansion (61) is stopped at the smallest term, then the remainder is exponentially small in ε. An exponentially small remainder for a Gevrey series is thus obtained by stopping the series at the smallest term, if one can prove that the remainder is of the order of the first omitted term.

In this subsection, in order to obtain an exponential estimate for the remainder R in (4), we prove that if $\Phi_i(Q, P, \tau)$ writes $\Phi_i(Q, P, \tau) = \Phi_i^{(1)}(Q, P, \tau) + \Phi_i^{(2)}(P, \tau)$, then

Proposition 3.3. There exist two constants, F and σ, depending only on the order n up to which the perturbation series is led, such that

$$\forall i \leqslant n, \quad \overline{\left\langle \left(\frac{\partial \Phi_i^{(1)}}{\partial Q} + \Phi_i^{(2)} \right)^2 \right\rangle_i} \leqslant F\sigma^i (i!)^3 \binom{2i-1}{i} \tag{63}$$

$$\forall i \leqslant n, \quad \overline{\langle h_i^2 \rangle_i} \leqslant F\sigma^i (i!)^3 \binom{2i-1}{i} \tag{64}$$

Using Property 5 of Appendix C, it follows from (63) that

$$\forall i \leqslant n, \quad \overline{\langle \Phi_i^2 \rangle_i} \leqslant F\sigma^i (i!)^3 \binom{2i-1}{i} \tag{65}$$

$$\forall i \leqslant n, \quad \overline{\left\langle \left(\frac{\partial \Phi_i}{\partial Q} \right)^2 \right\rangle_i} \leqslant F\sigma^i (i!)^3 \binom{2i-1}{i} \tag{66}$$

Proof. The proof is made by induction. Let us suppose that for any $j \leqslant i-1$, $(i \geqslant 2)$, the inequalities (63) and (64) are satisfied. It is then proven in Appendix D that if

$$F \leqslant \frac{1}{4a^2 \exp(2)} \tag{67}$$

then

$$\overline{\langle X_i^2 \rangle_i} \leqslant \frac{i^2}{a} F\sigma^i \binom{2i-1}{i} (i!)^3 \tag{68}$$

where X_i is defined by (15). Let us choose, for example

$$a = 8n^2 \exp(-2) \tag{69}$$

then $a > n^2$, and let us choose

$$F = \frac{1}{32n^4} \tag{70}$$

so that (67) is satisfied. Because h_i is obtained by keeping only certain terms of $-X_i$, those whose phase velocities are such that $|v_\varphi - P^*| \leqslant 1$ or those which oscillate with an angular frequency less than unity in absolute value, $\overline{\langle h_i^2 \rangle_i} \leqslant \overline{\langle X_i^2 \rangle_i} \leqslant F\sigma^i \binom{2i-1}{i} (i!)^3$. Using a notation analogous to (27), let us now write X_i under the form

$$X_i = \frac{1}{\Delta v^{i/2}} \sum_{l=1}^{i} \sum_{\varepsilon_1, \ldots, \varepsilon_i} \sum_{m_1, \ldots, m_i} g_l(P) \cos\left(\sum_{k=1}^{2i} \varepsilon_k \xi_k \right) \tag{71}$$

Then, solving (14) for Φ_i yields

$$\frac{\partial \Phi_i^{(1)}}{\partial Q} = \frac{1}{\Delta v^{i/2}} \sum_{l=1}^{i} \sum_{\substack{\varepsilon_1, \ldots, \varepsilon_i \\ \varepsilon_1 + \cdots + \varepsilon_i \neq 0}} \sum_{m_1, \ldots, m_i} \frac{g_l(P) \cos(\sum_{k=1}^{2i} \varepsilon_k \xi_k)}{P - \sum_{k=1}^{i} \varepsilon_k m_k / (\Delta v \sum_{k=1}^{i} \varepsilon_k)} \tag{72}$$

$$\Phi_i^{(2)} = \frac{1}{\Delta v^{i/2}} \sum_{l=1}^{i} \sum_{\substack{\varepsilon_1, \ldots, \varepsilon_i \\ \varepsilon_1 + \cdots + \varepsilon_i = 0}} \sum_{m_1, \ldots, m_i} \frac{g_l(P) \cos(\sum_{k=1}^{2i} \varepsilon_k \xi_k)}{\sum_{k=1}^{i} \varepsilon_k m_k / (\Delta v \sum_{k=1}^{i} \varepsilon_k)} \tag{73}$$

Now, by construction, in (72)

$$\left| \sum_{k=1}^{i} \varepsilon_k m_k / \Delta v \right| > 1 \tag{74}$$

and in (73) $|P^* - \sum_{k=1}^{i} \varepsilon_k m_k / (\Delta v \sum_{k=1}^{i} \varepsilon_k)| > 1$, which implies that on the disk $D_{i,a}$

$$\left| P - \sum_{k=1}^{i} \varepsilon_k m_k \middle/ \left(\Delta v \sum_{k=1}^{i} \varepsilon_k \right) \right| > i/\sqrt{a} \tag{75}$$

Then, from (67), (72), (73), (74) and (75) it follows that

$$\overline{\left\langle \left(\frac{\partial \Phi_i^{(1)}}{\partial Q} + \Phi_i^{(2)} \right)^2 \right\rangle_i} \leqslant \frac{a}{i^2} \overline{\langle X_i^2 \rangle_i} \leqslant F\sigma^i \binom{2i-1}{i} (i!)^3 \tag{76}$$

Because $\overline{\langle(\partial\Phi_1^{(1)}/\partial Q+\Phi_1^{(2)})^2\rangle_1}=\overline{\langle\Phi_1^2\rangle_1}$ and $\overline{\langle h_1^2\rangle_1}=\langle h_1^2\rangle$, in order to conclude the proof of Proposition 3.3 one only has to check that $\overline{\langle\Phi_1^2\rangle_1}\leqslant F\sigma$ and that $\langle h_1^2\rangle\leqslant F\sigma$. Through the same kind of calculations as the ones made to derive (24), one finds that if

$$\Delta v\geqslant\sqrt{a} \tag{77}$$

then

$$\overline{\langle\Phi_1^2\rangle_1}\leqslant\frac{\pi^2}{6}\sqrt{a} \tag{78}$$

Condition (77) will be discussed later on. According to (78), $\overline{\langle\Phi_1^2\rangle_1}\leqslant F\sigma$ if $\sigma\geqslant(\pi^2/6)(\sqrt{a}/F)$ which, according to (69) and (70), is satisfied if one chooses

$$\sigma=64n^5 \tag{79}$$

Then because $a>n\geqslant 1$, $\langle h_1^2\rangle\leqslant\sqrt{3/2}<(\pi^2/6)\sqrt{a}\leqslant F\sigma$, which concludes the proof of Proposition 3.2. Using the inequality $\binom{2i-1}{i}\leqslant 4^{i-1}$, Proposition 3.3 implies that

$$\overline{\left\langle\left(\frac{\partial\Phi_i^{(1)}}{\partial Q}+\Phi_i^{(2)}\right)^2\right\rangle_i}\leqslant\frac{F}{4}(4\sigma)^i(i!)^3 \tag{80}$$

$$\overline{\langle h_i^2\rangle_i}\leqslant\frac{F}{4}(4\sigma)^i(i!)^3 \tag{81}$$

In order to obtain an exponential estimate for the remainder, one has to stop the perturbation series at the smallest term. Actually, because the values of the terms of the perturbation series are not exactly known, we evaluate the rank of the smallest term given by the right hand side of (81). Hence, we estimate

$$\frac{\varepsilon^{2n}\langle h_n^2\rangle}{\varepsilon^{2(n-1)}\langle h_{n-1}^2\rangle}\approx 4\varepsilon^2\sigma n^3=256n^8\varepsilon^2 \tag{82}$$

From (82), we deduce that the rank of the smallest term is $n^*\approx 1/(2\varepsilon^{1/4})$. However, as shown in Appendix D, for technical reasons we have to impose $n\leqslant 1/2^{9/8}\varepsilon^{1/4}$. We thus choose to stop the perturbation expansion at

$$n^*=\mathrm{Int}\left(\frac{1}{2^{9/8}\varepsilon^{1/4}}\right) \tag{83}$$

We can now discuss condition (77): $\Delta v \geqslant \sqrt{a}$. Using the values (69) and (83) for a and n respectively, and using the fact that $\Delta v = (A/\varepsilon)^{2/3}$, one finds that (77) is satisfied if

$$\Delta v \geqslant (2^{9/8}\varepsilon^{5/12}/e) \tag{84}$$

Therefore, the inequalities (63) and (64) are true only if Δv is large enough. Note however that, as shown in Subsection 3.1.1, no restriction needs to be imposed on Δv in order to prove that the dynamics of (1) can be considered as reduced in the variables (Q, P). Therefore, condition (84) is only a technical one that needs to be imposed in order to derive an exponential estimate for the remainder. If are only considered values of ε less than unity, (84) translates to $\Delta v \geqslant (2^{3/8}/e)$.

3.2.3. Exponential Estimate for the Remainder. In order to derive an exponential estimate for the remainder, we use (31) with $\varepsilon_0 = 1$ and n given by (83), which yields

$$\sqrt{\langle R^2(\varepsilon)\rangle} \leqslant \varepsilon^{1/2^{9/8}\varepsilon^{1/4}} \sup_{\varepsilon' \leqslant 1} \sqrt{\langle R^2(\varepsilon')\rangle} \tag{85}$$

We then use (37) to estimate $\sup_{\varepsilon' \leqslant 1} \sqrt{\langle \mathrm{R}^2(\varepsilon')\rangle}$ and we prove, in Appendix E, that on the interval $[P^* - 1 + n/\sqrt{a}, P^* + 1 - n/\sqrt{a}]$

$$\sup_{\varepsilon' \leqslant 1} \sqrt{\langle \mathrm{R}^2(\varepsilon')\rangle} \leqslant 5 \tag{86}$$

which concludes the proof of Theorem 2.1.

4. FINITE RANGE OF THE PERTURBATIONS

Theorem 2.1 proved that, as regards the statistical properties, the dynamics defined by (1) can be considered as having a reduced range in phase velocity, in the variables (Q, P). This is however not enough to show the property of locality because, as soon as the perturbation analysis is led to an order larger than 2, all the terms present in (1) give a contribution to the h_i's in (7). Therefore, Theorem 2.1 does not prove by itself that only the terms such that $|m - p(t)| \leqslant \Delta v_R$, Δv_R proportional to $A^{2/3}$, are relevant to describe the statistical properties of the dynamics defined by (1). However, by using the results of Theorem 2.1, we are able to provide a rationale showing that the statistical properties of the dynamics defined by

(1) and (2) can be made arbitrarily close. This rationale is not a rigorous proof, but it is however very detailed and goes through rigorous steps.

4.1. Scheme of the Derivation

In order to take advantage of the results of the previous section, we work in the variables (Q, P) in order to conclude about the statistical properties of the dynamics of (1) and (2) in the physical variables (q, p). To do so, we study the dynamics of (1) on a time interval $[0, t]$ which we divide in smaller intervals I_j such that on each I_j, and for each phase realization, the instantaneous velocity $p(t)$ remains close to a given value $\Delta v P_j$. On each I_j we perform the change of variables $(q, p) \mapsto (Q, P)$ about the value P_j, transforming (1) into the Hamiltonian $\hat{H}_j + R_n = P^2/2 + \sum_{i=1}^{n} \varepsilon^i h_i + R_n$, where each h_i has a width in phase velocity equal to two, and R_n is the remainder. However, the results of Section 3 hold only if $P(t)$ is close to P_j, which is true if $P(t)$ remains close enough to $p(t)/\Delta v$. We actually are able to prove that $\langle [p(t)/\Delta v - P(t)]^2 \rangle$ can be made arbitrarily small by decreasing ε. This implies that the relative measure of the phase realizations such that $P(t)$ is close to $p(t)/\Delta v$ can be made arbitrarily close to one. Then, without rigorously showing it, we proceed as though, to estimate the statistical properties of the dynamics of (1), we could consider that for all the phase realizations $P(t)$ remains close enough to $p(t)/\Delta v$ so that on each time interval I_j, $P(t) \in [P_j - (1 - \delta), P_j + (1 - \delta)]$, $1/\Delta v < \delta < 1$. Under such an assumption, we can prove that the contribution of the remainder to the statistical properties of $\hat{H}_j + R_n$ can be made arbitrarily small by decreasing ε. Therefore, when the remainder is negligible, the statistical properties of the dynamics of (1), in the variables (Q, P), can be deduced from a sequence of Hamiltonians $\hat{H}_j = P^2/2 + \sum_{i=1}^{n} \varepsilon^i h_i$ whose potential parts only contain waves with phase velocities v_φ such that $|v_\varphi - P_j| \leqslant 1$.

We then consider the dynamics defined by (2) with $\Delta v_R = C \Delta v$, C being a constant larger than 1. As in the case of (1) we divide $[0, t]$ in smaller intervals I_j such that on each I_j $P(t)$ remains close to a given value P'_j. Then, in the variables (Q, P) the statistical properties of dynamics of (2) can be deduced from a sequence of Hamiltonians $\hat{H}'_j = P^2/2 + \sum_{i=1}^{n} \varepsilon^i h'_i$ which only contain waves with phase velocities v_φ such that $|v_\varphi - P'_j| < 1$. Therefore, in order to compare the statistical properties of the dynamics of (1) and (2) in the variables (Q, P), one only has to compare those of $\hat{H}_j$ and $\hat{H}'_j$.

These Hamiltonians are actually a sum of terms of the form (26) where P has to be replaced by $P(t)$, which depends on the phase realization, and where the integers m_k also depend on the phase realizations because the change of variables is defined about P_j, which depends on the phases. This implies that it is impossible to derive the statistical properties of $\hat{H}_j$ and $\hat{H}'_j$ directly in the variables (Q, P), because, for example, the variations of $(Q - m_k\tau/\Delta v)$ cannot be controlled when the φ_m's vary. Hence, we define the new variables $\Delta Q = Q - P_j\tau$ and $\Delta P = P - P_j$. In (26), the cosine in the numerator can be written as $\cos(\Delta Q)\cos[\sum_{k=1}^{l} \varepsilon_k\zeta_k + v_{j,l}(\pi/2)] - \sin(\Delta Q)\sin[\sum_{k=1}^{l} \varepsilon_k\zeta_k + v_{j,l}(\pi/2)]$, where $\zeta_k = (P_j - m_k/\Delta v)\,\tau + \varphi_{m_k}$. Now, if we choose P_j such that the fractional part of $\Delta v P_j$ is a constant, for example 1/2, then $x_k \equiv (P_j - m_k/\Delta v)$ does not depend on the phase realization. The denominator in (26) can be expanded in Taylor series about $P = P_j$ to yield a sum of functions of ΔP multiplied by functions of the x_k's. This implies that, in the variables $(\Delta Q, \Delta P)$, the dynamics of $\hat{H}_j$ is given by the Hamiltonian $\Delta\hat{H}_j$ which can be written in the symbolic form $\Delta\hat{H}_j = \sum F(\Delta Q, \Delta P)\,\mathrm{G}(\bar{x}, \bar{\varphi})$, where $\bar{x}$ stands for the set of the phase independent x_k's and $\bar{\varphi}$ denotes the set of phases present in H. Similarly, in the variables $(\Delta Q, \Delta P)$, the dynamics of $\hat{H}'_j$ is given by the Hamiltonian $\Delta\hat{H}'_j = \sum F(\Delta Q, \Delta P)\,\mathrm{G}'(\bar{x}, \bar{\varphi})$. Since the functions of $(\Delta Q, \Delta P)$ are the same in $\Delta\hat{H}_j$ and $\Delta\hat{H}'_j$, in order to compare the statistical properties of the dynamics defined by (1) and (2) in the variables $(\Delta Q, \Delta P)$ (and thus in the variables (Q, P)), one only has to compare the statistical properties of the functions $G(\bar{x}, \bar{\varphi})$ to those of $G'(\bar{x}, \bar{\varphi})$. Clearly, the functions G and G' exactly have the same dependence in terms of the x_k's, only the range of values assumed by the x_k's are not the same. In the case of G, $|x_k| \leqslant 1$ or $1 < |x_k| < |M/\Delta v \pm P_j|$, while in the case of G', $|x_k| \leqslant 1$ or $1 < |x_k| < \mathrm{C}$. Since the x_k's are phase independent it is easy to show, following the same lines as in Section 3, that the l-time correlation function of G' is the Riemann sum of an integral, I_C which converges when $C \to \infty$ towards the integral I_∞.

Evaluating the l-time correlation function of G is more difficult because the maximum value reached by $|x_k|$ depends on the phase realizations through P_j. However, the contribution of the terms involving the x_k's such that $|x_k| > \Delta_m$, where Δ_m is a large enough constant, are negligible, not only to evaluate the statistical properties of the dynamics, but also as regards the details of the dynamics. If only those terms existed, then the situation would be similar to that of Nekhoroshev theorem. Moreover, it can be readily seen that in (26) the absolute value of the sum of the terms involving the x_k's such that $|x_k| > \Delta_m$ can be made arbitrarily small by increasing Δ_m. Therefore, in order to estimate the statistical properties of G, we only take into account the x_k's such that $|x_k| \leqslant \Delta_m$. By doing so we find that the l-time correlation function of G is the Riemann sum of an integral, I_{Δ_m}, which converges when $\Delta_m \to \infty$ towards the same integral I_∞ as in the case of G', since G and G' have the same functional dependence in terms of the x_k's. Therefore, the l-time correlation function of G and G' can be made arbitrarily close simply by increasing C as long as $\Delta_m \geqslant C$. This implies that the statistical properties of the dynamics defined by (1) and (2) can be made arbitrarily close in the variables $(\Delta Q, \Delta P)$, and therefore also in the variables (Q, P).

Now, since the generating functions Φ_j used to define the change of variables $(q, p) \mapsto (Q, P)$ have exactly the same form as the Hamiltonians h_i, we deduce that the statistical properties of the change of variables $(q, p) \mapsto (Q, P)$, defined from (1) or from (2) can be made arbitrarily close. Hence, if we choose $\Delta v_R = \alpha A^{2/3}$, by increasing α, and as long as $|M \pm p(t)|$ remains large enough (at least larger than Δv_R), the statistical properties of (1) and (2) can be made arbitrarily close in the physical variables (q, p). This shows the property of locality.

We are now going to develop the previous mathematical arguments.

4.2. Derivation of the Property of Locality

Since we are going to work in the variables (Q, P) in order to conclude about the statistical properties of the dynamics of (1) and (2) in the physical variables (q, p), it is therefore necessary to make clear some of the properties of the change of variables $(q, p) \mapsto (Q, P)$. We are in particular going to study the values of P solutions to the equation $p/\Delta v = P + \varepsilon\, \partial\Phi'(q_1, P, \tau)/\partial q_1$ where p may depend on the phases φ_m's. For any phase realization there exists a given P_j such that the fractional part of $(\Delta v P_j)$ is $1/2$, and such that $|p/\Delta v - P_j| \leqslant 1/(2\,\Delta v)$. Then, although P_j depends on the phase realization, the terms of the kind $m_k/\Delta v - P^*$, present in (26), are independent of the φ_m's when P^* is replaced by P_j. What we actually need to prove in order to derive the property of locality is that when $|p/\Delta v - P_j| \leqslant 1/(2\,\Delta v)$ there always is a value of P solution to $p/\Delta v = P + \varepsilon\, \partial\Phi'(q_1, P, \tau)/\partial q_1$ such that $P \in [P_j - (1-\delta), P_j + (1-\delta)]$, where $0 < \delta < 1$ and $(1-\delta) > 1/(2\,\Delta v)$. This is true if, when P varies over the whole interval $[P_j - (1-\delta), P_j + (1-\delta)]$, $p/\Delta v = P + \varepsilon\, \partial\Phi'(q_1, P, \tau)/\partial q_1$ varies over an interval including $[P_j - 1/(2\,\Delta v), P_j + 1/(2\,\Delta v)]$. In order to prove this last result, it is enough to prove that $|P - p/\Delta v| < 1 - \delta - 1/(2\,\Delta v)$, since p is a continuous function of P when P varies from $P_j - (1-\delta)$ to $P_j + (1-\delta)$. Here, we prove that the normalized measure of the initial phases φ_m such that $|P - p/\Delta v| \geqslant 1 - \delta - 1/2\,\Delta v$ can be made arbitrarily small by decreasing ε. To do so, we prove in Appendix F that $\langle (p/\Delta v - P)^2 \rangle$ can be made arbitrarily small by decreasing ε. Then, in the remaining of this section, we are going to proceed as if it were possible to find a solution P to the equation $p/\Delta v = P + \varepsilon\, \partial\Phi'(q_1, P, \tau)/\partial q_1$ in the interval $P \in [P_j - (1-\delta), P_j + (1-\delta)]$ for any phase realization, although we do not provide here an estimate of the influence on the statistical properties of the dynamics of the phase realizations for which such a solution cannot be found. This last point is the only one which is missing in order to obtain a rigorous result regarding the property of locality.

Let us now consider the dynamics defined by (1) on a given time interval $[0, t]$ which we divide in smaller intervals I_j such that on each I_j, and for any phase realization, $P(\tau)$ remains in an interval of the type $[P_j-(1-\delta),\ P_j+(1-\delta)]$ where $0<\delta<1$, $\delta\Delta v>1$ and $(1-\delta)\,\Delta v>1/2$ (this is possible because $dP/d\tau$ remains finite). The second condition on δ is necessary for the estimates on the remainders made in the previous section to be valid, while the third condition insures that it is possible to choose the P_j's such that the fractional part of $\Delta v P_j$ is 1/2. Then, the terms of the kind $m_k/\Delta v-P_j$, present in (26), are independent of the phase realization.

On each time interval I_j, in the variables (Q, P), the dynamics of (1) is given by a Hamiltonian of the same type as (7). When the remainder is negligible, in order to derive the statistical properties of (1) in the variables (Q, P), it is enough to study on each I_j the Hamiltonian

$$\hat{H}_j=\frac{P^2}{2}+\sum_{i=1}^{n}\varepsilon^i h_i(Q, P, \tau) \tag{87}$$

where the h_i's only contain waves with phase velocities v_φ such that $|v_\varphi-P_j|\leqslant 1$. Let us now define $\Delta Q=Q-P_j\tau$ and $\Delta P=P-P_j$. In the variables $(\Delta Q, \Delta P)$ the dynamics of (87) is defined by

$$\Delta\hat{H}_j=\frac{\Delta P^2}{2}+\sum_{i=1}^{n}\varepsilon^i h_i(\Delta Q+P_j\tau, \Delta P+P_j, \tau) \tag{88}$$

Now, according to the form (26) of the h_i's, one can write (88) under the form

$$\begin{aligned}\Delta\hat{H}_j=\frac{\Delta P^2}{2}+\sum_{k=-n}^{n}&[F_k(P_j\tau, \Delta P+P_j, \tau, \bar{\varphi}, \varepsilon)\cos[k\,\Delta Q]\\&+G_k(P_j\tau, \Delta P+P_j, \tau, \bar{\varphi}, \varepsilon)\sin[k\,\Delta Q]]\end{aligned} \tag{89}$$

where $\bar{\varphi}$ denotes the set of phases present in H. The functions F_k and G_k will not be explicitly calculated here, but can be easily deduced from (26). From (26) one sees that these functions only depend on τ through the $(m_i/\Delta v-P_j)\,\tau$'s, where, as already mentioned, P_j is such that $m_i/\Delta v-P_j$ is independent of time and of the phase realization. As long as $\Delta P<1$, the functions F_k and G_k can be expanded in Taylor series so that (89) can be written as

$$\Delta\hat{H}_j = \frac{\Delta P^2}{2} + \sum_{l=0}^{+\infty} \frac{(\Delta P)^l}{l!} \sum_{k=-n}^{n} \left[\cos[k\,\Delta Q] \left.\frac{\partial^l F_k}{\partial P^l}\right|_{P=P_j} + \sin[k\Delta Q] \left.\frac{\partial^l G_k}{\partial P^l}\right|_{P=P_j} \right] \tag{90}$$

From (90) it is clear that the statistical properties of the dynamics of (1) in the variables $(\Delta Q, \Delta P)$ are totally determined by the statistical properties of the derivatives of the F_k and G_k evaluated at $P = P_j$. These derivatives are functions of the variables $x_i = m_i/\Delta v - P_j$ which are independent of time and of the phase realization. Therefore, although the Hamiltonian (90) describes the dynamics of (1) only locally in time, whatever the time interval I_j, $\Delta\hat{H}_j$ is expressed through the same functions of the same variables x_i. Only the range of values assumed by the x_i's depend on the time interval I_j as the x_i's are such that $-M/\Delta v - P_j \leqslant x_j < -1$, or $1 < x_j \leqslant M/\Delta v - P_j$, or $|x_i| \leqslant 1$.

Let us now consider the Hamiltonian

$$H'_j = \frac{P^2}{2} + A \sum_{|m/\Delta v - P_j| \leqslant C} \cos(q - mt + \varphi_m) \tag{91}$$

where $C > 1$ and where the fractional part of $\Delta v P_j$ is 1/2. Theorem 2.1 applies to H'_j so that one can define the change of variables $(q, p) \mapsto (Q, P)$ transforming H'_j in a Hamiltonian of the same kind as (7), where the variance of the remainder can be made exponentially small in ε. When the remainder is negligible, the statistical properties of (91) in the variables (Q, P) are deduced from

$$\hat{H}'_j = \frac{P^2}{2} + \sum_{i=1}^{n} \varepsilon^i h'_i(Q, P, \tau) \tag{92}$$

where in (92) the h'_i's only contain waves with phase velocities v_φ, such that $|v_\varphi - P_j| \leqslant 1$. Actually the h'_i's have exactly the same form (25)–(26) as the h_i's. The only difference between the h_i's and the h'_i's comes from the domain of variations of the integers m_k in (26). In the case of the h_i's the m_k's are such that $|m_k| \leqslant M$ while in the case of the h'_i's they are such that $|m_k/\Delta v - P_j| \leqslant 1$. Therefore, as in the case of the Hamiltonian (87), as long as $P(\tau) \in [P_j - (1-\delta), P_j + (1-\delta)]$, in the variables $(\Delta Q, \Delta P)$ the dynamics of (92) is defined by a Hamiltonian of the type

$$\Delta \hat{H}'_j = \frac{\Delta P^2}{2} + \sum_{l=0}^{+\infty} \frac{(\Delta P)^l}{l!} \sum_{k=-n}^{n} \left[\cos[k\,\Delta Q] \left.\frac{\partial^l F'_k}{\partial P^l}\right|_{P=P_j} + \sin[k\Delta Q] \left.\frac{\partial^l G'_k}{\partial P^l}\right|_{P=P_j} \right] \tag{93}$$

Moreover, the derivatives of the functions F'_k and G'_k have exactly the same dependence in terms of the variables $x_i = m_i/\Delta v - P_j$ as the derivatives of the functions F_k and G_k.

The only difference between these derivatives comes from the domain of variation of the x_i's. When $|x_i| > 1$ then, in the case of F_k and G_k, $-M/\Delta v - P_j \leqslant x_j < -1$ or $1 < x_j \leqslant M/\Delta v - P_j$, while in the case of F'_k and G'_k $-C \leqslant x_i < -1$ or $1 \leqslant x_i < C$. Therefore, in the case of $\Delta \hat{H}'_j$, the domain of variations of the x_i's is independent of P_j. This implies that the statistical properties of the dynamics defined by $\Delta \hat{H}'_j$ are independent of P_j.

Let us now consider the dynamics defined on the interval $[0, t]$, and in the variables (Q, P), in the following way. For each phase realization $P(0)$ is assigned a given value P_1, such that the fractional part of $\Delta v P_1$ is 1/2, and, on the time interval $I'_1 = [0, \tau_1]$ such that $P(\tau) \in [P_1 - (1-\delta), P_1 + (1-\delta)]$ whatever the phase realization, we consider the dynamics defined by (91) with P_j replaced by P_1. Since $(1-\delta)\,\Delta v > 1/2$, there exists a value P_2 such that the fractional part of $\Delta v P_2$ is 1/2 and such that for any phase realization $|P(\tau_1) - P_2| \leqslant \delta_1 < (1-\delta)$. Then, on the time interval I'_2 such that $P(\tau) \in [P_2 - (1-\delta), P_2 + (1-\delta)]$, regardless of the phase realization, we consider the dynamics defined by (91) with P_j replaced by P_2. Similarly, we can divide the whole interval $[0, t]$ in time intervals I'_j such that on each I'_j $P(\tau)$ lies in an interval of the type $P(\tau) \in [P_j - (1-\delta),\ P_j + (1-\delta)]$ and on each I'_j we consider the dynamics defined by (91). It is always possible to choose I'_j and I_j, defined in the case of H'_j and H respectively, such that $I_j = I'_j$, which we actually do.

In the variables $(\Delta Q, \Delta P)$, and when the remainders are negligible, the statistical properties of the dynamics defined by the sequence of Hamiltonians H'_j are actually completely determined by the statistical properties of the derivatives of the F'_k and G'_k evaluated at $P = P_j$. Therefore, in order to compare the statistical properties of the dynamics defined by H and by the sequence of Hamiltonians H'_j, in the variables $(\Delta Q, \Delta P)$, we compare the m-time correlation functions of $\partial^l K_k/\partial P^l|_{P=P_j}$ and $\partial^l G_k/\partial P^l|_{P=P_j}$ to those of $\partial^l F'_k/\partial P^l|_{P=P_j}$ and $\partial^l G'_k/\partial P^l|_{P=P_j}$. We actually restrict to the case when the time interval $[0, t]$ is such that for almost all the phase realizations $|M/\Delta v \pm P_j| \geqslant \Delta_m \geqslant C$ for the dynamics of H. We then write F_k under the form $F_k = F_k^{(1)} + F_k^{(2)}$ where in $F_k^{(2)}$ the x_i's are larger than Δ_m in absolute value, and make a similar decomposition for G_k.

From (26) it is clear that the absolute value of $F_k^{(2)}$ and $G_k^{(2)}$ can be made arbitrarily small by increasing the value of Δ_m, whatever the value of the phase realization. Hence, their contribution to the determination of the statistical properties of the dynamics defined by the $\Delta\hat{H}_j$'s can be made arbitrarily small. The value of Δ_m to impose in order to make $F_k^{(2)}$ and $G_k^{(2)}$ small may depend on M. However, one can see that when all the cosines in (26) are set equal to 1, the sum may at most diverge logarithmically with M. Therefore, Δ_m is at most proportional to $\ln(M)$, and thus only weakly depends on M.

Hence, in order to compare the statistical properties of the dynamics defined by H and by the sequence of Hamiltonians H'_j, we only need to compare the m-time correlation functions of $\partial^l F_k^{(1)}/\partial P^l|_{P=P_j}$ and $\partial^l G_k^{(1)}/\partial P^l|_{P=P_j}$, which we denote by $\langle F_{l,m}^{(1)}\rangle$ and $\langle G_{l,m}^{(1)}\rangle$ respectively, to the m-time correlation functions of $\partial^l F'_k/\partial P^l|_{P=P_j}$ and $\partial^l G'_k/\partial P^l|_{P=P_j}$, which we denote by $\langle F'_{l,m}\rangle$ and $\langle G'_{l,m}\rangle$ respectively. Since F_k and F'_k only depend on time through some cosines, one can always find an upper bound for $|\langle F_{l,m}^{(1)}\rangle - \langle F'_{l,m}\rangle|$ by setting the m times to 0. The same property holds for $|\langle G_{l,m}^{(1)}\rangle - \langle G'_{l,m}\rangle|$. We thus restrict to the case when the m times are 0. Then, as in the section 3, one can show that $\langle F_{l,m}^{(1)}\rangle$ and $\langle F'_{l,m}\rangle$ are Riemann sums of integrals which converge when Δ_m, or C, goes to infinity. Therefore, for any $\eta > 0$, there exists a value Δv_η, and a sum of multiple integrals, which we simply denote by I_{Δ_m}, such that if $\Delta v > \Delta v_\eta$, then

$$|\langle F_{l,m}^{(1)}\rangle - I_{\Delta_m}| < \eta/4 \tag{94}$$

Similarly, there exists a value $\Delta v'_\eta$ and a sum of multiple integrals I'_C such that if $\Delta v > \Delta v'_\eta$, then

$$|\langle F'_{l,m}\rangle - I'_C| < \eta/4 \tag{95}$$

Note that it is always possible to choose $\Delta v_\eta = \Delta v'_\eta$, which we actually do. Moreover, since $F_k^{(1)}$ and F'_k have exactly the same dependence in terms of the x_i's, I_{Δ_m} and I'_C are integrals of the same function, and these integrals converge when Δ_m or C goes to infinity. Therefore, when Δ_m and C go to infinity, I_{Δ_m} and I'_C converge towards the same limit I_∞. Hence, there exists a value C_η, independent of Δv_η, such that if $C > C_\eta$,

$$|I'_C - I_\infty| < \eta/4 \tag{96}$$

Now, as $\Delta_m \geqslant C$, when (96) is fulfilled then

$$|\mathrm{I}_{\Delta_m} - \mathrm{I}_\infty| < \eta/4 \tag{97}$$

Using Eq. (94) to (97), one then finds

$$|\langle F_{l,m}^{(1)}\rangle - \langle F'_{l,m}\rangle|$$
$$\leqslant |\langle F_{l,m}^{(1)}\rangle - I_{\Delta m}| + |I_{\Delta m} - I_{\infty}| + |I_{\infty} - I'_C| + |I'_C - \langle F'_{l,m}\rangle| < \eta \qquad (98)$$

An inequality similar to (98) can also be derived for $|\langle G_{l,m}^{(1)}\rangle - \langle G'_{l,m}\rangle|$. Therefore, on the interval $[0, t]$, and as long as the remainders obtained from perturbation theory are negligible, in the variables $(\Delta Q, \Delta P)$ the statistical properties of the dynamics defined by H and those of the dynamics defined by the sequence of Hamiltonians H'_j can be made arbitrarily close, simply by increasing the value of C in (91), regardless of the number of terms in (1).

Now, if we choose the same distribution function of P_1 for both dynamics, then, since in the time interval I_1 the statistical properties of the dynamics defined by H and H'_1 can be made arbitrarily close in the variables $(\Delta Q, \Delta P)$, the statistical properties of the dynamics defined by H and H'_1 can also be made arbitrarily close in the variables (Q, P). Then, the statistical properties of P_2 can also be made arbitrarily close for both dynamics, and by induction, the statistical properties of H and of the dynamics defined by the sequence of Hamiltonians H'_j can be made arbitrarily close in the variables (Q, P).

Since the functions Φ_j which define the change of variables $(q_1, p_1) \mapsto (Q, P)$ have exactly the same form as the Hamiltonians h_i and h'_i in (87) and (92) respectively, one can use the same kind of reasoning as previously to show that the statistical properties of changes of variables defined from H or from the H'_j's can be made arbitrarily close. Therefore, H and the sequence of Hamiltonians H'_j define in physical variables (q, p) dynamics which can be made arbitrarily close on the statistical point of view.

Let us now explain how this result can be used to show that the statistical properties of the dynamics defined by H and H' (2) (with $\Delta v_R = C\,\Delta v$) can be made arbitrarily close. Note first of all that in (2) $p(t)$ can be replaced by $\Delta v P_j(t)$, where the fractional part of $\Delta v P_j = 1/2$, and where $|p(t) - \Delta v P_j| < 1/2$. Then, if $|p - \Delta v P_j| < 1/2$ implies that $|P(t) - P_j| < (1 - \delta)$, (which is true if $|P - p/\Delta v| < 1 - \delta - 1/2\,\Delta v$), the dynamics defined by (2) is equivalent to the dynamics defined by the sequence of Hamiltonians H'_j. Now, the result derived in Appendix F also applies to H' because it was obtained independently of the number of waves in (1). Therefore, $\langle (p/\Delta v - P)^2\rangle$ can be made arbitrarily small by decreasing ε. This implies that the normalized measure of the initial phases φ_m such that when $|p(t) - \Delta v P_j| < 1/2$, $|P(t) - P_j| \geqslant (1 - \delta)$, can be made arbitrarily small by decreasing ε. As in the case of the dynamics of H we consider that this is enough to conclude about the statistical properties of H'. Therefore, the statistical properties of the dynamics defined by (2) can be considered as equivalent to those defined by the sequence of Hamiltonians H'_j. This implies that, when the remainders obtained from perturbation theory are negligible, the statistical properties of the dynamics defined by (1) and (2) can be made arbitrarily close by choosing $\Delta v_R = \alpha A^{/23}$ with α large enough.

Let us now estimate the influence of the remainders when evaluating the l-time force correlation functions. We actually perform the calculation in the case of H, and the same results can be obtained in the case of H'. After performing a perturbation analysis about $P = P_j$ one obtains the Hamiltonian

$$\hat{H}_j + R_n = \frac{P^2}{2} + \sum_{i=1}^{n} \varepsilon^i h_i(Q, P, \tau) + R_n \tag{99}$$

The remainder can be written as $R_n = \Re_\alpha(q_1) + \Re_\beta(Q)$. Therefore the force, which is defined as $F(\tau) = -\partial(\hat{H}_j + R_n)/\partial Q$ can be written as

$$F(\tau) = -\sum_{i=1}^{n} \varepsilon^i \, \partial h_i/\partial Q - (\partial q_1/\partial Q)(\partial \Re_\alpha/\partial q)_1 - \partial \Re_\beta/\partial Q \tag{100}$$

The contribution of the remainder to the l-time correlation function, which we denote by $\langle \Delta F^l \rangle$ is analytic in ε and its power expansion in terms of ε begins by ε^{n+1}. Moreover, by using the inequality $|xy| \leqslant (x^2 + y^2)/2$, one can show that $|\langle \Delta F^l \rangle|$ can be bounded from above in terms of moments of the type $\langle (\partial h_i/\partial Q)^{2j_1} \rangle$, $\langle (\partial q_1/\partial Q)^{2j_2} \rangle$, $\langle (\partial \Re_\alpha/\partial q_1)^{2j_3} \rangle$, and $\langle (\partial \Re_\beta/\partial Q)^{2j_4} \rangle$. Therefore, if one can prove that each of these moments can be bounded from above independently of M and Δv, then one proves an inequality of the type $|\langle \Delta F^l \rangle| \leqslant \varepsilon^{n+l} \Delta F_0$, where ΔF_0 is a constant. $(\partial h_i/\partial Q)^{j_1}$ is a function of the type (25)–(26), therefore, by following the same lines as the ones used to estimate $\langle (\partial \Phi_i/\partial q_1)^2 \rangle$ (see Appendix F) one can prove that $\langle (\partial h_i/\partial Q)^{2j_1} \rangle$ can be bounded from above independently of M and Δv. The same result holds for $\langle (\partial \Re_\beta/\partial Q)^{2j_4} \rangle$ and $\langle (\partial \Re_\alpha/\partial q_1)^{2j_3} \rangle$ for the same reason. Now, $(\partial q_1/\partial Q)^{2j_2} = 1/(1 + \partial^2 \Phi'/\partial q_1 \, \partial P)^{2j_2}$, and $(1 + \partial^2 \Phi'/\partial q^1 \, \partial P)^{2j_2}$ can be written as $(1 + \partial^2 \Phi'/\partial q^1 \, \partial P)^{2j_2} = 1 + \varepsilon \, \Delta \Phi'_{j_2}$, where $\langle |\Delta \Phi'_{j_2}| \rangle$ can be bounded from above in terms of moments of the type $\langle (\partial^2 \Phi^i/\partial q_1 \, \partial P)^{2j} \rangle$. Then, by following the same lines as the ones used to estimate $\langle (\partial \Phi_i/\partial q_1)^2 \rangle$, one can prove that $\langle |\Delta \Phi'_{j_2}| \rangle$ can be bounded from above independently of M and Δv. This implies that when ε is small enough $1 + \varepsilon \langle \Delta \Phi'_{j_2} \rangle \geqslant 1/2$, and therefore $\langle (\partial q_1/\partial Q)^{2j_2} \rangle \leqslant 2$. Therefore, for ε small enough, the l-time force correlation function can be written

$$\langle F(\tau_1) \cdots F(\tau_l) \rangle = (-1)^l \left[\langle \partial \hat{H}_j(\tau_1)/\partial Q \cdots \partial \hat{H}_j(\tau_l)/\partial Q \rangle + \langle \Delta F^l \rangle \right. \tag{101}$$

and there exists a constant ΔF_0 such that $|\langle \Delta F' \rangle| \leqslant \varepsilon^{n+l} \Delta F_0$. Now, $\langle \partial \hat{H}_j(\tau_1)/\partial Q \cdot \partial \hat{H}_j(\tau_l)/\partial Q \rangle$ is analytic in ε and is different from the null function. Therefore, its expansion in powerseries of ε begins by ε^m, with $m \geqslant l$. Hence, as soon as the perturbation analysis is led up to an order larger than $(m-l)$, the contribution of the remainder to the force correlation function is negligible.

Summarizing our results, we found that the statistical properties defined by the Hamiltonians (1) and (2) can be made arbitrarily close if:

(i) the remainders R and R' are negligible, which is true if $\varepsilon = A/\Delta v^{3/2}$ is small enough. This implies that one has to choose $\Delta v = \gamma A^{2/3}$, with γ large enough,

(ii) the changes of variables are statistically close to identity, which implies again that $\varepsilon = A/\Delta v^{3/2}$ has to be small,

(iii) Δv is large enough (and therefore A is large enough) for the Riemann sums $\langle F^{(1)}_{l,m} \rangle$, $\langle G^{(1)}_{l,m} \rangle$, $\langle F'_{l,m} \rangle$ and $\langle G'_{l,m} \rangle$ to be close to their integrals,

(iv) $C = \Delta v_R/\Delta v$ and $|M/\Delta v \pm p(t)|$ are large enough, which physically means that the number of waves acting upon the particle is large enough, and that the particle's orbit remains sufficiently far away from the edge of the chaotic domain.

5. ANALYTICAL ESTIMATE OF THE RANGE OF PERTURBATIONS

The previous sections showed the existence of a finite range in phase velocity, Δv_R, proportional to $A^{2/3}$, for the perturbations. We propose here an analytical calculation, based on physical considerations, in order to yield a more precise estimate of Δv_R. The method consists in statistically comparing the Hamiltonians $\hat{H}_j = P^2/2 + \sum_{i=1}^n \varepsilon^i h_i(Q, P, \tau)$ and $\hat{H}'_j = P^2/2 + \sum_{i=1}^n \varepsilon^i h'_i(Q, P, \tau)$. Since $h_1 = h'_1$, we actually lead the perturbation analysis up to second order in ε and focus on h_2 and h'_2. We gather in h_2 all the terms having the same phase velocity $P_j + v_\varphi$, and the sum $A(v_\varphi)$ of all these terms is considered as a single wave whose amplitude is $\sqrt{\langle A^2(v_\varphi) \rangle}$. We proceed in the same way for h'_2 and obtain the amplitude $\sqrt{\langle A'^2(v_\varphi) \rangle}$. Then, we calculate the value of C such that

$$\Gamma = \left| \frac{\langle A'^2(v_\varphi) \rangle}{\langle A^2(v_\varphi) \rangle} - 1 \right| \tag{102}$$

is less than a given small quantity η for any value of v_φ such that $|v_\varphi| \leqslant 1$. As the variances of velocity for the dynamics of $\hat{H}_j$ and $\hat{H}'_j$ are directly related to the $\langle A^2(v_\varphi)\rangle$'s and the $\langle A'^2(v_\varphi)\rangle$'s, we expect that when $\Gamma = \eta$, the relative discrepancy between the variances of velocity for the dynamics defined by (1) or (2) is of the order of η. In order to estimate Γ we replace the Riemann sums $\langle A^2(v_\varphi)\rangle$ and $\langle A'^2(v_\varphi)\rangle$ by their integrals which we calculate at $P = P_j$. Then, we choose $\eta = 5\%$ and find that $\Gamma \leqslant 5\%$ if

$$C \geqslant 3.4 \tag{103}$$

Since the estimate of C was made from the Hamiltonian H'_j defined with respect to the new variable P_j, we account here for the shift between the old and new variables in order to estimate Δv_R. Hence, we do not simply estimate Δv_R as $\Delta v_R = C\,\Delta v$ but rather as $\Delta v_R = C\,\Delta v + \Delta p$ where $\Delta p = \sqrt{\langle (p - \Delta v P)^2\rangle}$. Calculating Δp to second order in phase velocity, and replacing the Riemann sums by integrals, we obtain $\Delta \mathrm{p} = \Delta \mathrm{v} \times \sqrt{\varepsilon^2/3 + \varepsilon^4(-97/216 + 15\ln(3)/32)}$. Taking into account the fact that $\Delta v = (A/\varepsilon)^{2/3}$ and the value (103) of C we finally find

$$\Delta v_R \approx \varepsilon^{-1/3}(3.4 + \varepsilon\sqrt{1/3 + \varepsilon^2(-97/216 + 15\ln(3)/32)} \tag{104}$$

To conclude the estimate of Δv_R, it remains to find an estimate for ε. To do so we use condition (ii) of the previous section imposing to the change of variables to be statistically close to identity. A way to interpret this condition is to choose ε such that the change of variables is one-to-one for a substantial fraction of the phase realizations. From (11) one sees that $\partial P/\partial p_1$ and $\partial Q/\partial q_1$ keep a constant sign if $|\partial^2\Phi'/\partial q_1\,\partial P| \leqslant 1$. We then choose ε such that $|\langle(\partial^2\Phi'/\partial q_1\,\partial P)^2\rangle| \leqslant 1$. Using a second order calculation, this leads to $\varepsilon \leqslant 0.88$. Then (104) yields the following estimate for Δv_R

$$\Delta v_R \approx 4.2 A^{2/3} \tag{105}$$

When Δv_R assumes the value (105), the relative discrepancy between the variances of velocity for the dynamics defined by (1) and (2) is expected to be of the order of 5%. Actually, as will be shown in the next section, the value of Δv_R leading to such a discrepancy is numerically estimated to be $\Delta v_R \approx 5.4 A^{2/3}$, which is of the order of what is found here. Moreover, by solving $\Gamma = \eta$ (Γ given by (102)) for various values of η, we find that Δv_R varies with η in a way similar to what is numerically observed.

6. NUMERICAL RESULTS

In order to check the property of locality and to measure the range of the perturbations, we numerically integrate the equations of motion derived from the Hamiltonians H (1) and H' (2) for various values of Δv_R and compare the velocity distribution functions thus obtained.

The Hamilton equations are numerically integrated by using a leap-frog scheme.[14, 15] The number $(2M+1)$ of waves in (1) is chosen such that during the whole simulation $|M \pm p(t)| \geqslant \Delta v_R$. The time step is chosen equal to $1/M$ for the dynamics of (1), and equal to $1/\Delta v_R$ for the dynamics of (2). This process is repeated with 9000 different phase realizations.

The velocity distribution functions obtained from (1) and (2) are compared by using a Kolmogorov–Smirnov test.[16, 17] Let $f_1(p)$ be the numerical velocity distribution function obtained from (1), and $f_2(p)$ the one obtained from (2), then the Kolmogorov–Smirnov test yields the probability that these two distribution functions describe the same phenomenon as a function of the maximum value of $|\int_{-\infty}^{p} f(p_1)\,dp_1 - \int_{-\infty}^{p} f(p_2)\,dp_2|$, with respect to p. We actually compare the form of the distribution functions, i.e. we work with $g_1(p)$ and $g_2(p)$ obtained from $f_1(p)$ and $f_2(p)$ respectively by making their average value be zero, and their variance be unity. For values of Δv_R close to 5, the Kolmogorov–Smirnov test indicates that $g_1(p)$ and $g_2(p)$ can be considered as identical. However, this test is not precise enough to provide a criterion yielding a threshold value of Δv_R. Nevertheless, as the numerical simulation proceeds, $f_1(p)$ and $f_2(p)$ become closer and closer to Gaussians with zero means. Therefore, there is only one relevant parameter to characterize these distribution functions: their variance. Actually, it is easier to work on $\langle [p(t)-p(0)]^2 \rangle / 2t$, which is close to a constant, the diffusion coefficient, for times t large enough. The procedure consists in increasing the value of Δv_R until reaching a value denoted by Δv_m such that when $\Delta v_R = \Delta v_m$, the relative discrepancy between the

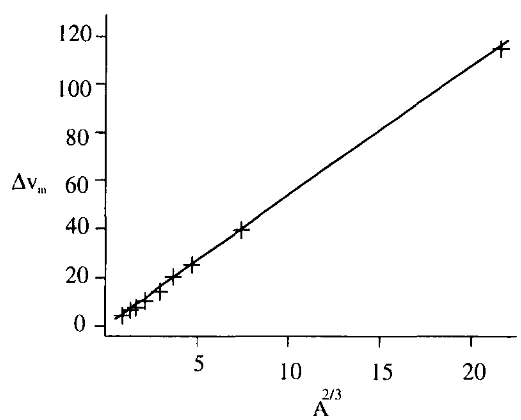

Fig. 2. Δv_m versus $A^{2/3}$ (pluses). The straight line has a slope of 5.4.

diffusion coefficients D and D', for the dynamics defined by H and H' respectively, is close to 5%, which is of the order of the numerical errors. Fig. 2 plots Δv_m versus $A^{2/3}$ for $K = 4\pi^2 A$ varying from 30 to 4000. The curve obtained is close to a straight line with slope 5.4. Therefore the property of locality as well as the scaling $A^{2/3}$ are numerically checked, and the value of Δv_R leading to a relative discrepancy close to 5% between the diffusion coefficients of the dynamics of (1) and (2) is close to $5.4A^{2/3}$.

7. UNIVERSALITY OF THE STATISTICAL PROPERTIES OF THE DYNAMICS

Since the statistical properties of the dynamics defined by (1) and (2) can be made arbitrarily close, we derive here results regarding the statistical properties of dynamics of (1) by studying those of the dynamics defined by (2). Let us first note that, as long as $|M \pm p(t)|$ is large enough, because of locality, the statistical properties of (1) are independent of M. Therefore, during a finite time, and for M large enough, the statistical properties of the dynamics defined by (1) are universal with respect to the number of waves. We are now going to show that during a finite time, they are also independent of the wave amplitudes, in appropriate coordinates. We are actually going to show this result for the dynamics defined by H' and then use the property of locality to conclude about H.

The dynamics defined by H' can actually be seen as defined by a sequence of Hamiltonians of the type

$$H''_j = p^2/2 + A \sum_{|m - p_j| \leqslant \Delta v_R} \cos(q - mt + \varphi_m) \tag{106}$$

where the fractional part of p_j is 1/2 and where $|p(t) - p_j| < 1/2$. Let us now define $q_r = q$, $p_r = A^{-2/3}p$, $t_r = A^{2/3}t$, $\Delta q_r^{(j)} = q_r - A^{-2/3}p_j t_r$ and $\Delta p_r^{(j)} = p_r - A^{-2/3}p_j$. It is clear that if p_1 is chosen independently of the phase realizations, then the statistical properties of the dynamics of H' in the variables (q_r, p_r) are directly related to those obtained in the variables $(\Delta q_r, \Delta pr)$. In $(\Delta q_r, \Delta p_r)$, the dynamics of H''_j is given by

$$\tilde{H}''_j = (\Delta p_r^{(j)})^2/2 + \cos(\Delta q_r)\, C_j(t_r) - \sin(\Delta q_r)\, S_j(t_r) \tag{107}$$

where

$$C_j(t_r) + iS_j(t_r) = A^{-1/3} \sum_{|m - p_j| \leqslant \Delta v_R} \exp[iA^{-2/3}(p_j - m)\, t_r + i\varphi_m] \tag{108}$$

From (108) it is clear that the statistical properties of the dynamics defined by each H_j'', and therefore those of the dynamics defined by H', are completely determined in the variables $(\Delta q_r, \Delta p_r)$ and hence also in the variables (q_r, p_r), by the statistical properties of the functions C_j and S_j. Let us now calculate the 2-time correlation function of C_j

$$\langle C_j(t_r)\, C_j(t_r')\rangle = \frac{A^{-2/3}}{2} \sum_{|m-p_j| \leqslant \Delta v_R} \cos[(p_j - m)\, A^{-2/3}\,(t_r - t_r')] \tag{109}$$

When $(t_r - t_r') \ll A^{2/3}$ then the Riemann sum (109) can be replaced by its integral. Let m_1 be the smallest integer m such that $(m - p_j) \geqslant -\Delta v_R$ and m_2 the largest integer m such that $(m - p_j) \leqslant \Delta v_R$. Then, by replacing in (109) the Riemann sum by its integral, one obtains

$$\langle C_j(t_r)\, C_j'(t_r')\rangle \approx \frac{1}{2} \int_{A^{-2/3}(m_1 - p_j)}^{A^{-2/3}(m_2 - p_j)} \cos[x(t_r - t_r')]\, dx \tag{110}$$

When $A^{2/3} \gg 1$ the limits of the integral (110) are close to $-\Delta v_R / A^{2/3}$ and $\Delta v_R / A^{2/3}$. From Sections 3 and 4 we know that Δv_R is proportional to $A^{2/3}$: $\Delta v_R = \alpha A^{2/3}$. Therefore

$$\begin{aligned} \langle C_j(t_r)\, C_j(t_r')\rangle &\approx \frac{1}{2} \int_{-\alpha}^{\alpha} \cos[x(t_r - t_r')]\, dx \\ &\approx \begin{cases} \dfrac{\sin[\alpha(t_r - t_r')]}{t_r - t_r'}, & \text{if } \mathrm{t_r \neq t_r'} \\ \alpha, & \text{if } \mathrm{t_r = t_r'} \end{cases} \end{aligned} \tag{111}$$

Hence, when $A^{2/3} \gg 1$, and up to times t_r proportional to $A^{2/3}$, the 2-time correlation function of C_j can be considered as independent of A. Similarly, for any l, one can express the l-time correlation functions of C_j and S_j in terms of Riemann sums. Then, up to times proportional to $A^{2/3}$, these Riemann sums can be replaced by their integrals, and one can then see that when $A^{2/3} \gg 1$ the l-time correlation functions of C_j and S_j can be considered as independent of A. This implies that up to a time t_r proportional to $A^{2/3}$, in the variables (q_r, p_r, t_r), the statistical properties of the dynamics defined by H', and hence those of the dynamics defined by H, can be considered as independent of A. Now, as $t_r = A^{2/3} t$, we find that the universality of the statistical properties of the dynamics defined by H with respect to the waves amplitudes, in the variables (q_r, p_r, t_r), is valid up to a physical time t_0 which is independent of A.

8. GENERALIZATION OF THE PROPERTY OF LOCALITY TO A WIDER CLASS OF HAMILTONIANS

In this section we investigate how the results regarding the property of locality, derived in the previous sections for the Hamiltonian (1), can generalize to a Hamiltonian like

$$H_g = p^2/2 + \sum_{m=1}^{M} a_m \cos(k_m q - \omega_m t + \varphi_m) \tag{112}$$

where the φ_m's are still fixed random phases. We will not try here to perform a rigorous perturbation analysis on (112), as we did on (1), nor will we make any detailed numerical calculations. This will be left for future work. Instead, we propose here a heuristic calculation yielding an upper bound for the width of locality, and indicate where lies the mathematical difficulty to obtain a rigorous result.

Let us first proceed as in the case of the Hamiltonian (1) and define dimensionless variables $p_1 = p/\Delta v$, $q_1 = k_{\max} q$ and $\tau = k_{\max} \Delta v t$, where $k_{\max}$ is the wavenumber in (112) having the largest absolute value. Then, in the variables q_1, p_1 and τ, the dynamics of (112) is defined by the Hamiltonian

$$H'_g = \frac{p_1^2}{2} + \frac{1}{\Delta v^2} \sum_{m=1}^{M} a_m \cos(\kappa_m q_1 - \bar{\omega}_m \tau + \varphi_m) \tag{113}$$

where $\kappa_m = K_{\max}$ and $\bar{\omega}_m = \omega_m/(k_{\max} \Delta v)$. As in the case of the Hamiltonian (1) we perform a canonical change of variables $(q_1, p_1) \mapsto (Q, P)$ defined about $P = P^*$, using a generating function Φ, in order to transform (113) into the Hamiltonian $\hat{H}'_g$ having the width 2 in phase velocity (now measured in units Δv). Unlike in the case of (1), we do not try here to specify *a priori* a small parameter for the perturbation analysis. We thus just write the generating function Φ as

$$\Phi = Pq_1 + \sum_{i=1}^{n} \Phi_i \tag{114}$$

and the transformed hamiltonian $\hat{H}'_g$ as

$$\hat{H}'_g = \frac{P^2}{2} + \sum_{i=1}^{n} h_i(Q, P, \tau) + R_n(Q, P, \tau) \tag{115}$$

where the h_i's only contain waves with phase velocities v_φ such that $|v_\varphi - P^*| \leqslant 1$. The Φ_i's and the h_i's are related by (14), where the X_i's are given by (15) when $i \geqslant 2$, and

$$X_1 = \frac{-1}{\Delta v^2} \sum_{m=1}^{M} a_m \cos(\kappa_m q_1 - \bar{\omega}_m \tau + \varphi_m) \tag{116}$$

Solving (14) for h_i and Φ_1, with X_1 given by (116), yields

$$h_1 = \frac{1}{\Delta v^2} \sum_{|\bar{\omega}_m - \kappa_m P^*| \leqslant 1} a_m \cos[\kappa_m q_1 - \bar{\omega}_m \tau + \varphi_m] \tag{117}$$

$$\Phi_1 = \frac{1}{\Delta v^2} \sum_{|\bar{\omega}_m / - \kappa_m P^*| > 1} \frac{a_m \cos[\kappa_m q_1 - \bar{\omega}_m \tau + \varphi_m]}{\bar{\omega}_m - \kappa_m P} \tag{118}$$

At first order, the remainder is

$$R_1 = \frac{1}{2} \left(\frac{\partial \Phi_1}{\partial q_1} \right)^2 + h_1(q_1, \tau) - h_1(Q, \tau) \tag{119}$$

The difficulty in evaluating R_1 lies in the fact that it is expressed in terms of mixed variables. When the wavenumbers are not all the same, the method used in the case of the Hamiltonian (1) no longer holds, and we are no longer able to calculate $\langle R_1^2 \rangle$. At this point, one might want to use the formalism of the Lie transform, instead of a generating function, in order to perform the canonical change of variables. Using a Lie transform would avoid having to deal with mixed variables. However, we found some difficulties in using the Lie transform in the case of large perturbations. We argue that there may be a basic reason for this, coming from the fact that a Lie transform always defines changes of variables close to identity, while, as noted in Section 4, in the case when one deals with large perturbations, the change of variables is not close to identity for all the phase realizations.

Instead of providing a rigorous estimate for R_1, we use a physically reasonable approximation of R_1 given by

$$R_1 \approx \frac{1}{2} \left(\frac{\partial \Phi_1}{\partial q_1} \right)^2 + \frac{\partial h_1}{\partial q_1} (Q - q_1) = \frac{1}{2} \left(\frac{\partial \Phi_1}{\partial q_1} \right)^2 + \frac{\partial h_1}{\partial q_1} \frac{\partial \Phi_1}{\partial P} \tag{120}$$

Using this approximation, one easily finds

$$\langle R_1^2 \rangle = \frac{1}{4\, \Delta v^8} \sum_{|\bar{\omega}_m - \kappa_m P^*| \leqslant 1} (\kappa_m a_m)^2 \sum_{|\bar{\omega}_m - \kappa_m P^*| > 1} \left[\frac{\kappa_m a_m}{(\bar{\omega}_m - \kappa_m P)^2} \right]^2 + \frac{3}{32} \left[\frac{1}{\Delta v^4} \sum_{|\bar{\omega}_m - \kappa_m P^*| > 1} \left(\frac{\kappa_m a_m}{\bar{\omega}_m - \kappa_m P} \right)^2 \right]^2 \tag{121}$$

From (121) one finds $\langle R_1^2 \rangle \leqslant \varepsilon_1^2 \varepsilon_2^2 + 3\varepsilon_2^4/8$ where

$$\varepsilon_1 = \frac{1}{\sqrt{2}\, \Delta v^2} \sqrt{\sum_{|\bar{\omega}_m/ - \kappa_m P^*| \leqslant 1} (\kappa_m a_m)^2} \tag{122}$$

$$\varepsilon_2 = \frac{1}{\sqrt{2}\, \Delta v^2} \sqrt{\sum_{|\bar{\omega}_m - \kappa_m P^*| > 1} \left[\frac{\kappa_m a_m}{\bar{\omega}_m - \kappa_m P} \right]^2} \tag{123}$$

From (117) it is easy to see that $\langle h_1^2 \rangle = (\sum_{|\bar{\omega}_m - \kappa_m P^*| \leqslant 1} a_m^2)/2\, \Delta v^2 \geqslant \varepsilon_1^4$, since for any m $\kappa_m \leqslant 1$. Then, if ε_1 and ε_2 can be made arbitrarily small by increasing Δv, R_1 can be made statistically arbitrarily small compared to h_1. Of course, this result is of interest only if Δv remains smaller than the width in phase velocity of the Hamiltonian (112). In particular ε_1 and ε_2 must remain small quantities also in the limit when the width in phase velocity of (112) goes to infinity. A necessary condition for R_1 to be negligible compared to h_1 is therefore that the sums (122) and (123) converge when the width in phase velocity of (112) goes to infinity.

Going to second order, one can easily see from (14) that h_2 is nothing but the sum of the terms of $(-1/2)(\partial \Phi_1/\partial q_1)^2 - (\partial h_1/\partial q_1)(\partial \Phi_1/\partial P)$ having a phase velocity v_φ such that $|v_\varphi - P^*| \leqslant 1$, or independent of q_1 but oscillating with an angular frequency less than 1 in absolute value. This implies that $\langle h_2^2 \rangle$ is less than $\langle R_1^2 \rangle$, when R_1 is replaced by the left hand side of (120). Therefore, $\langle h_2^2 \rangle \leqslant \varepsilon_1^2 \varepsilon_2^2 + 3\varepsilon_2^4/8$.

Similarly, from (14) one finds that Φ_2 is obtained by dividing each term of $(1/2)(\partial \Phi_1/\partial q_1)^2 + (\partial h_1/\partial q_1)(\partial \Phi_1/\partial P)$ by a term larger than unity in absolute value. Therefore, $\langle \Phi_2^2 \rangle \leqslant \varepsilon_1^2 \varepsilon_2^2 + 3\varepsilon_2^4/8$.

By induction, one can then prove, in a way similar to what has been done in Section 3, that for any $i \geqslant 2$, $\langle h_i^2 \rangle \leqslant \sum_{j=1}^{i} \alpha_j \varepsilon_1^{2(i-j)} \varepsilon_2^{2j}$ and $\langle \Phi_i^2 \rangle \leqslant \sum_{j=1}^{i} \beta_j \varepsilon_1^{2(i-j)} \varepsilon_2^{2j}$, where the α_j's and β_j's are constants. Then, by using the same kind of approximation for R_i as for R_1 one finds $\langle R_i^2 \rangle \leqslant \sum_{j=1}^{i+1} \gamma_j \varepsilon_1^{2(j-1)} \varepsilon_2^{2j}$ where the γ_j's are constants. Therefore, by making ε_1 and ε_2 small enough, one can make, at any order, the remainder negligible.

Hence we provide here an upper bound for the width of locality Δv obtained by making ε_1 (122) and ε_2 (123) small enough. In the case of the Hamiltonian (1), $\varepsilon_1 = (1/\sqrt{2}\,\Delta v^2)\sqrt{\sum_{|m-P^*| \leqslant \Delta v} A^2}$, and $\varepsilon_2 = (1/2\,\Delta v^2) \times \sqrt{\sum_{|m-P^*|>1} 1/(m/\Delta v - P)^2}$. This implies that $A\sqrt{2\,\Delta v - 1/\sqrt{2}\,\Delta v^2} \leqslant \varepsilon_1 \leqslant A\sqrt{2\,\Delta v + 1/\sqrt{2}\,\Delta v^2}$, so that for large enough values of Δv, $\varepsilon_1 \approx A/\Delta v^{3/2}$, while replacing in ε_2 the sum by an integral yields $\varepsilon_2 = A/\Delta v^{3/2}$. Therefore we recover in the case of (1) the scaling $A^{2/3}$ for Δv already obtained in Section 3. In this case, the scaling for Δv corresponds to the one proposed in ref. 9 in the framework of plasma physics and obtained by making the hypothesis that a diffusion in velocity occurs instantaneously. Actually, the scaling proposed in ref. 9 is $\Delta v = (k^2 D)^{1/3}$, where k is the typical wavenumber and D is the diffusion coefficient. This scaling does not, in general, correspond to the one obtained by making ε_1 (122) and ε_2 (123) small. Indeed, if the property of locality is satisfied, the width of locality Δv_R is proportional to Δv: $\Delta v_R = C\,\Delta v$, and the diffusion coefficient $D(p)$ is determined by the a_m's with m such that $|\bar{\omega}_m - \kappa_m p| \leqslant C$. Unlike $D(p)$, ε_2 is defined as a function of the a_m's with m such that $|\bar{\omega}_m - \kappa_m p| > 1$, and is thus in general not related to the diffusion coefficient.

9. CONCLUSION

This paper showed that large perturbations only have a finite range, as regards the statistical properties, in Hamiltonian dynamics. We considered a Hamiltonian describing the dynamics of a particle in a set of electrostatic waves whose initial phases are fixed random variables (this dynamics is a generalization of that defined by the standard map). We introduced a perturbation theory in the case when the waves amplitude A is large enough to allow large variations of the action, by taking advantage of the fact that the perturbations which are far from being resonant can be removed. We proved that it is possible to transform the initial Hamiltonian into one where the terms out of resonance by a mismatch in phase velocity larger than $\Delta v > A^{2/3}$, give an exponentially small contribution in the parameter $\varepsilon = A/\Delta v^{3/2}$ for $\varepsilon \ll 1$.

By taking advantage of the results of the perturbation analysis, we showed that the statistical properties of the Hamiltonian dynamics we considered could be made arbitrarily close to that of a stochastic dynamics including at each time t only the waves with phase velocities m such that $|m - p(t)| \leqslant \Delta v_R$, Δv_R proportional to $A^{2/3}$. By an appropriate rescaling of the coordinates we then showed the universality of the statistical properties of the dynamics with respect to A, on a finite time interval. For practical purposes, we estimated and numerically verified the perturbation due to any wave to have a finite range $\Delta v_R = \alpha A^{2/3}$, of $\alpha \approx 5$ in velocity: wave-particle interaction occurs locally in phase velocity. We indicated how these results could generalize to a wider class of Hamiltonians, and we showed that the range of action of a resonance could have various scalings with the amplitude of the perturbation.

As already alluded to in the introduction, these results are the cornerstone of a (non-rigorous) theory explaining the origin of diffusion in Hamiltonian dynamics. Indeed, in refs. 6–8 we substantiated with numerical checks a series of assertions deduced from these results through physical arguments:

– the decorrelation of the force acting on the particle occurs after a change of its action by $2\,\Delta v_R$

– chaotic diffusion sets in after a change of action by $4\Delta v_R$

– the so-called quasilinear regime of diffusion is due to the crossover between two regimes of diffusion that are of different nature.

APPENDIX A: ESTIMATE OF THE TERMS OF THE PERTURBATION SERIES

The aim of this appendix is to prove Proposition 3.1. To do so, it is necessary to understand the form of the generating functions Φ_i, and of the Hamiltonians h_i. It is easy to show, by induction, that Φ_i writes

$$\Phi_i = \sum_{l=1}^{i} S_l(q_1, P, \tau) \tag{A1}$$

with

$$S_l = \frac{1}{\Delta v^{i/2}} \sum_{\varepsilon_1} \cdots \sum_{\varepsilon_i} \sum_{m_1} \cdots \sum_{m_i} \sum_{j=1}^{n_l} \frac{s_{j,l} \cos\left[\sum_{k=1}^{i} \varepsilon_k \xi_k + v_{j,l} \frac{\pi}{2}\right]}{\prod_{k=1}^{l} \left(\frac{m_k}{\Delta v} - P\right)^{\alpha_{j,k,l}} \Delta_{j,l}(\overline{m_{1,i}}/\Delta v, \overline{\varepsilon_{1,i}}, P)} \tag{A2}$$

where

$v_{j,l}$ is an integer,

$\alpha_{j,k,l} \geqslant 1$,

$\xi_k = Q - m_k \tau/\Delta v + \varphi_m$,

$\overline{\varepsilon_{1,i}}$ stands for the i-tuple $(\varepsilon_1, \ldots, \varepsilon_i)$,

$\varepsilon_1 = 1$ and $\varepsilon_k \in \{-1; 1\}$ if $2 \leqslant k \leqslant i$,

$\overline{m_{1,i}}$ stands for the i-tuple $(m_1,..., m_i)$, and $\overline{m_{1,i}}/\Delta v$ stands for the i-tuple $(m_1/\Delta v,..., m_i/\Delta v)$,

$\Delta_{j,l}$ is of the form

$$\Delta_{j,l} = \prod_{n_1=1}^{\beta_{j,l}} \prod_{n_2=1}^{\gamma_{j,l}} \left(\sum_{k_1=1}^{i} \frac{\lambda_{k_1} m_{k_1}}{\Delta v} - P \right)^{\delta_{n_1}} \left(\sum_{k_2=1}^{i} \frac{\mu_{k_2} m_{k_2}}{\Delta v} \right)^{\chi_{n_2}} \tag{A3}$$

where the λ_{k_1}'s and the μ_{k_2}'s are either -1, 0, or 1, and where δ_{n_1} and χ_{n_2} are positive integers, the m_k's satisfy $|m_k| \leqslant M$, and $|m_k/\Delta v - P^*| > 1$, if $1 \leqslant k \leqslant l$, and $|m_k/\Delta v - P^*| \leqslant 1$, if $l+1 \leqslant k \leqslant i$. Moreover, they are such that all the factors in $\Delta_{j,l}$ are strictly larger than unity in absolute value when evaluated at $P = P^*$. This implies that on the disk D_r, centered on P^* and of radius $(1-r)$, there exists an integer $\eta_{j,l}$ such that

$$|\Delta_{j,l}(P)| \geqslant r^{\eta_{j,l}} \tag{A4}$$

whatever $\overline{\varepsilon_{1,i}}$ and $\overline{m_{1,i}}$.

The h_i's are also of the form (25)–(26), except h_1 which actually only contains the term $l = 0$.

Once the form of the Φ_i's and the h_i's is found, one can calculate their variances. From (A1)–(A2), one gets

$$\langle \Phi_i^2 \rangle = \sum_{l_1=1}^{i} \sum_{l_2=1}^{i} \langle S_{l_1} S_{l_2} \rangle \tag{A5}$$

and

$$\langle S_{l_1} S_{l_2} \rangle = \frac{1}{\Delta v^i} \sum_{\varepsilon_1} \cdots \sum_{\varepsilon_{2i}} \sum_{m_1} \cdots \sum_{m_{2i}} \sum_{j_1=1}^{n_{l_1}} \sum_{j_2=1}^{n_{l_2}} \frac{S_{j_1,l_1} S_{j_2,l_2} \left\langle \cos\left[\sum_{k=1}^{2i} \varepsilon_k \xi_k + (\nu_{j_1,l_1} + \nu_{j_2,l_2})\, \pi/2 \right] \right\rangle}{2 \prod_{k=1}^{l_1+l_2} \left(\frac{m_k}{\Delta v} - P \right)^{\alpha_k} \Delta_{j_1,l_1} \Delta_{j_2,l_2}} \tag{A6}$$

In (A6) the indices have been renumbered so that $|m_k/\Delta v - P^*| > 1$ when $1 \leqslant k \leqslant l_1 + l_2$. Because the φ_k's are random phases, the only terms which may be non-zero in (A6) are such that

$$\sum_{k=1}^{2i} \varepsilon_k \varphi_{m_k} = 0 \tag{A7}$$

For one given realization of the ε_i's, let us denote by $\overline{m_1^+}$ the set of the m_k's such that $|m_k/\Delta v - P^*| > 1$ when $\varepsilon_k = +1$, $\overline{m_1^-}$ the set of the m_k's, such that $|m_k/\Delta v - P^*| > 1$ when $\varepsilon_k = -1$, $\overline{m_2^+}$ the set of the m_k's such that $|m_k/\Delta v - P^*| \leqslant 1$ when $\varepsilon_k = +1$, $\overline{m_2^-}$ the set of the m_k's such that $|m_k/\Delta v - P^*| \leqslant 1$ when $\varepsilon_k = -1$, and finally $\overline{m^+} = \overline{m_1^+} \cup \overline{m_2^+}$. The phases φ_k being independent the ones from the others, the condition (A7) is fulfilled if, and only if

$$\sum_{k=1}^{l_1+l_2} \varepsilon_k = 0 \qquad \text{and} \qquad \sum_{k=1+l_1+l_2}^{2i} \varepsilon_k = 0 \tag{A8}$$

and if there exists a one-to-one relation, σ_1, between $\overline{m_1^-}$ and $\overline{m_1^+}$, and a one-to-one relation, σ_2, between $\overline{m_2^-}$ and $\overline{m_2^+}$:

$$\overline{m_1^-} = \sigma_1(m_1^+),\ \overline{m_2^-} = \sigma_2(m_2^+) \tag{A9}$$

The conditions (A9) can be fulfilled only if $\overline{m_1^-}$ and $\overline{m_1^+}$, and $\overline{m_2^-}$ and $\overline{m_2^+}$, have the same number of elements, which is only possible if $(l_1 + l_2)$ is even. Therefore, if $(l_1 + l_2)$ is odder $\langle S_{l_1} S_{l_2} \rangle = 0$ and if $(l_1 + l_2)$ is even, we define

$$\sigma(m_k) = \begin{cases} m_k, & \text{if } \ m_k \in \overline{m^+} \\ \sigma_1(m_k), & \text{if } \ m_k \in \overline{m_1^-} \\ \sigma_2(m_k), & \text{if } \ m_k \in \overline{m_2^-} \end{cases} \tag{A10}$$

Then

$$\Delta v^i \, |\langle S_{l_1} S_{l_2} \rangle| \leqslant \sum_{\varepsilon_1} \cdots \sum_{\{0\}} \sum_{\varepsilon_{2i}} \sum_{\sigma_1} \sum_{\sigma_2} \sum_{m_k \in \overline{m_+}} \sum_{j_1=1}^{n_{l_1}} \sum_{j_2=1}^{n_{l_2}} \times \frac{|s_{j_1, l_1} s_{j_2, l_2}|}{2 \left| \prod_{k=1}^{l_1+l_2} \left(\frac{\sigma(m_k)}{\Delta v} - P \right)^{\alpha_k} \Delta_{j_1, l_1}[\sigma(\overline{m_{1,i}})] \, \Delta_{j_2, l_2}[\sigma(\overline{m_{1+i, 2i}})] \right|} \tag{A11}$$

where the symbol $\{0\}$ means that the ε_k's fulfil the conditions (A8). Note that the inequality (A11) is independent of the phases $v_{j_1, l_1}\pi/2$ and $v_{j_2, l_2}\pi/2$. Theses phases are actually inessential and will be systematically omitted in the remainder of the paper. From (A4) we know that on the disk D_r there are 2 integers η_{j_1} and η_{j_2} such that $|\Delta_{j_1} \Delta_{j_2}| \geqslant r^{\eta_{j_1} + \eta_{j_2}}$. Then, denoting

$S(l_1, l_2) = \max_{\overline{\varepsilon_{1,2i}}} \{\sum_{j_1=1}^{n_{l_1}} \sum_{j_2=1}^{n_{l_2}} |s_{j_1,l_1} s_{j_2,l_2}|/2r^{\eta_{j_1}+\eta_{j_2}}\}$, one finds that for any P in D_r,

$$|\langle S_{l_1} S_{l_2} \rangle| \leqslant S(l_1, l_2) \prod_{k=1}^{(l_1+l_2)/2} \left(\frac{1}{\Delta v} \sum_{|m/\Delta v - P^*| > 1} \frac{1}{|m/\Delta v - P|^{2\alpha_k}} \right)$$
$$\times \left(\frac{1}{\Delta v} \sum_{|m/\Delta v - P^*| \leqslant 1} 1 \right)^{i-(l_1+l_2)/2} \sum_{\varepsilon_1} \cdots \sum_{\{0\}} \sum_{\varepsilon_{2i}} \sum_{\sigma_1} \sum_{\sigma_2} 1 \qquad \text{(A12)}$$

Now, if $\Delta v \geqslant 1$, $(1/\Delta v) \sum_{|m/\Delta v - P^*| \leqslant 1} 1 \leqslant 3$, and if $r\Delta v \geqslant 1$, from (22),

$$\frac{1}{\Delta v} \sum_{|m/\Delta v - P^*| > 1} \frac{1}{|m/\Delta v - P|^{2\alpha_k}} \leqslant \frac{1}{r^{2\alpha_k - 2} \Delta v} \sum_{|m/\Delta v - P^*| \leqslant 1} \frac{1}{|m/\Delta v - P|^2}$$
$$\leqslant \frac{\pi^2}{3r^{2\alpha_k - 1}} \qquad \text{(A13)}$$

To conclude the estimate of $\langle S_{l_1} S_{l_2} \rangle$, one only has to calculate that the number of 2i-tuples $(\varepsilon_1, ..., \varepsilon_{2i})$ fulfilling (A9) is $\binom{2i-(l_1+l_2)}{i-(l_1+l_2)/2}\binom{l_1+l_2-1}{(l_1+l_2)/2}$, that the number of the one-to-one relations σ_1 is $((l_1+l_2)/2)!$ and that the number of the one-to-one relations σ_2 is $(i-(l_1+l_2)/2)!$. The product of these quantities is maximum when $l_1 = l_2 = i$. Indeed, when (A8) is fulfilled, $\sum_{k=1}^{2i} \varepsilon_k = 0$. Hence, the number of 2i-tuples $(\varepsilon_1, ..., \varepsilon_{2i})$ fulfilling (A8) is maximum when $l_1 = l_2 = i$. Similarly, if $\overline{m_1^-}$ and $\overline{m_1^+}$, and $\overline{m_2^-}$ and $\overline{m_2^+}$, are related by a one-to-one relation, then $\overline{m^+}$ and $\overline{m^-}$ are also related by a one-to-one relation. Hence, the number of one-to-one relation σ_1 multiplied by the number of one-to-one relations σ_2 is less than the number of one-to-one relations σ between $\overline{m^+}$ and $\overline{m^-}$, i.e., $i!$. Finally, we find

$$|\langle S_{l_1} S_{l_2} \rangle| \leqslant S(l_1, l_2)\, 3^{i-(l_1+l_2)/2} \binom{2i-(l_1+l_2)}{i-(l_1+l_2)/2} \binom{l_1+l_2-1}{(l_1+l_2)/2} \left(\frac{l_1+l_2}{2}\right)!$$
$$\times \left(i - \frac{l_1+l_2}{2}\right)! \prod_{k=1}^{(l_1+l_2)/2} \left(\frac{\pi^2}{3r^{2\alpha_k - 1}}\right) \qquad \text{(A14)}$$

Hence, $\langle S_{l_1} S_{l_2} \rangle$ can be bounded from above by a quantity independent of A, M, and Δv. It is thus also true for Φ_i and h_i, which proves Proposition 3.1.

APPENDIX B: UPPER BOUND FOR THE REMAINDER

In this appendix we prove that on the interval $[P^* - (1-r), P^* + (1+r)]$, the remainder R_n, is less than a quantity independent of A,

M, and Δv. Using (7), (8), (11), and (12), one finds that the remainder R_n is

$$R_n(Q, P, \tau) = \frac{1}{2}\left(P + \frac{\partial \Phi'}{\partial q_1}\right)^2 + \varepsilon f(q_1, \tau) + \frac{\partial \Phi'}{\partial \tau} - \frac{P^2}{2} - \sum_{i=1}^{n} \varepsilon^i h_i(Q, P, \tau) \qquad \text{(B1)}$$

Φ' is chosen such that

$$P\frac{\partial \Phi'}{\partial q_1} + \varepsilon f(q_1, \tau) + \frac{\partial \Phi'}{\partial \tau} = \sum_{i=1}^{n} \varepsilon^i h_i(q_1, P, \tau) + \sum_{i=2}^{n} \varepsilon^i X_i \qquad \text{(B2)}$$

Therefore, R_n can be written as

$$R_n(Q, P, \tau) = \mathfrak{R}_\alpha(q_1, P, \tau) + \mathfrak{R}_\beta(Q, P, \tau) \qquad \text{(B3)}$$

where

$$\mathfrak{R}_\alpha(q_1, P, \tau) = \frac{1}{2}\left(\frac{\partial \Phi'}{\partial q_1}\right)^2 + \sum_{i=1}^{n} \varepsilon^i h_i(q_1, P, \tau) + \sum_{i=2}^{n} \varepsilon^i X_i(q_1, P, \tau) \qquad \text{(B4a)}$$

$$\mathfrak{R}_\beta(Q, P, \tau) = -\sum_{i=1}^{n} \varepsilon^i h_i(Q, P, \tau) \qquad \text{(B4b)}$$

In order to calculate the root mean square of the remainder one has to express R_n as a function of the same variable Q. We thus replace in (B3) $\mathfrak{R}_\alpha(q_1)$ by $\mathfrak{R}_\alpha(Q - \partial\Phi'/\partial P)$. We then rewrite $\mathfrak{R}_\alpha(Q - \partial\Phi'/\partial P)$ under the form

$$\mathfrak{R}_\alpha(Q - \partial\Phi'/\partial P, P, \tau) = \sum r_\alpha \cos(\tilde{\xi} - \tilde{\varepsilon}\, \partial\Phi'/\partial P) \qquad \text{(B5)}$$

where $\sum r_\alpha$ stands for the sum $\sum_{i=1}^{2n} (\varepsilon^i/\Delta v^{i/2}) \sum_{l=1}^{i} \sum_{\varepsilon_1} \cdots \sum_{\varepsilon_i} \sum_{m_1} \cdots \sum_{m_i} \times \sum_{j=1}^{n_l} r_{j,l} / \prod_{k=1}^{l} (m_k/\Delta v - P)^{\alpha_k} \Delta_{j,l}$, $\tilde{\varepsilon} = \sum_{\mathbf{k}=1}^{\mathrm{i}} \varepsilon_{\mathbf{k}}$, and $\tilde{\xi} = \sum_{\mathbf{k}=1}^{\mathrm{i}} \varepsilon_{\mathbf{k}} \xi_{\mathbf{k}}$. Using the inequality $xy \leqslant (x^2 + y^2)/2$, we then find

$$(\mathfrak{R}_\alpha + \mathfrak{R}_\beta)^2 \leqslant 3\left[\sum r_\alpha \cos\tilde{\xi} \cos\left(\tilde{\varepsilon}\frac{\partial \Phi'}{\partial P}\right)\right]^2 + 3\left[\sum r_\alpha \sin\tilde{\xi} \sin\left(\tilde{\varepsilon}\frac{\partial \Phi'}{\partial P}\right)\right]^2 + 3\mathfrak{R}_\beta^2 \qquad \text{(B6)}$$

We then isolate in the sum $\sum r_\alpha \cos\tilde{\xi} \cos(\tilde{\varepsilon}\,\partial\Phi'/\partial P)$ the terms having the same values of $|\tilde{\varepsilon}|$:

$$\sum r_\alpha \cos\tilde{\xi} \cos(\tilde{\varepsilon}\,\partial\Phi'/\partial P) = \sum_{i=0}^{2n} \cos(i\,\partial\Phi'/\partial P) \sum_{|\tilde{\varepsilon}|=i} r_\alpha \cos\tilde{\xi} \tag{B7}$$

which yields

$$\left[\sum r_\alpha \cos\tilde{\xi} \cos(\tilde{\varepsilon}\,\partial\Phi'/\partial P)\right]^2 = \sum_{i=0}^{2n} \sum_{j=0}^{2n} \cos(i\,\partial\Phi'/\partial P) \cos(j\,\partial\Phi'/\partial P) \left(\sum_{|\tilde{\varepsilon}|=i} r_\alpha \cos\tilde{\xi}\right)\left(\sum_{|\tilde{\varepsilon}|=j} r_\alpha \cos\tilde{\xi}\right) \tag{B8}$$

Using once again the inequality $xy \leqslant (x^2+y^2)/2$, and the fact that $|\cos(\ )| \leqslant 1$, one finds

$$\cos\left(\frac{i\,\partial\Phi'}{\partial P}\right) \cos\left(\frac{j\,\partial\Phi'}{\partial P}\right)\left(\sum_{|\tilde{\varepsilon}|=i} r_\alpha \cos\tilde{\xi}\right)\left(\sum_{|\tilde{\varepsilon}|=j} r_\alpha \cos\tilde{\xi}\right) \leqslant \frac{\left(\sum_{|\tilde{\varepsilon}|=i} r_\alpha \cos\tilde{\xi}\right)^2 + \left(\sum_{|\tilde{\varepsilon}|=j} r_\alpha \cos\tilde{\xi}\right)^2}{2} \tag{B9}$$

Therefore,

$$\left\langle\left[\sum r_\alpha \cos\tilde{\xi} \cos(\tilde{\varepsilon}\,\partial\Phi'/\partial P)\right]^2\right\rangle \leqslant (2n+1) \sum_{i=0}^{2n} \left\langle\left(\sum_{|\tilde{\varepsilon}|=i} r_\alpha \cos(\tilde{\xi})\right)^2\right\rangle \tag{B10}$$

Now

$$\langle \mathfrak{R}_\alpha^2(Q, P, \tau)\rangle = \sum_{i_1=0}^{2n} \sum_{i_2=0}^{2n} \sum_{|\tilde{\varepsilon}_1|=i_1} \sum_{|\tilde{\varepsilon}_2|=i_2} \langle r_\alpha^{(i_1)} \cos\tilde{\xi}^{(i_1)} r_\alpha^{(i_2)} \cos\tilde{\xi}^{(i_2)}\rangle \tag{B11}$$

and, from (A8), the only nonzero terms in (B11) are those terms such that $|\tilde{\varepsilon}_1| = |\tilde{\varepsilon}_2|$. Therefore $\langle \mathfrak{R}_\alpha^2(Q, P, \tau)\rangle = \sum_{i=1}^{2n} \langle (\sum_{|\tilde{\varepsilon}|=i} r_\alpha \cos\tilde{\xi})^2\rangle$. This implies that

$$\left\langle\left[\sum r_\alpha \cos\tilde{\xi} \cos(\tilde{\varepsilon}\,\partial\Phi'/\partial P)\right]^2\right\rangle \leqslant (2n+1)\langle \mathfrak{R}_\alpha^2(Q, P, \tau)\rangle \tag{B12}$$

Similarly, we estimate

$$\left\langle \left(\sum r_\alpha \sin \tilde{\xi} \sin(\tilde{\varepsilon}\, \partial \Phi' / \partial P) \right)^2 \right\rangle \leqslant 2n \langle \mathfrak{R}'^2_\alpha \rangle \tag{B13}$$

where

$$\mathfrak{R}'^2_\alpha = \sum_{i=1}^{2n} \left[\sum_{\tilde{\varepsilon}=i} r_\alpha \sin \tilde{\xi} - \sum_{\tilde{\varepsilon}=-i} r_\alpha \sin \tilde{\xi} \right]^2 \tag{B14}$$

From (B6), (B10), (B12) and (B13) it follows that

$$\begin{aligned} \langle R_n^2(Q, P, \tau) \rangle \leqslant (6n+3) \langle \mathfrak{R}_\alpha^2(Q, P, \tau) \rangle + 6n \langle \mathfrak{R}'^2_\alpha(Q, P, \tau) \rangle \\ + 3 \langle \mathfrak{R}_\beta^2(Q, P, \tau) \rangle \end{aligned} \tag{B15}$$

Now, all the terms in (B15) are of the form (25)–(26). Hence, using Proposition 3.1, the inequality (B15) implies that $\sqrt{\overline{\langle R_n^2 \rangle}}$ is less than a quantity independent of A, M, and Δv.

APPENDIX C: PROPERTIES OF $\overline{\langle (\ )^2 \rangle_k}$

In Subsection 3.2.1 is introduced, for any function B of the form (25)–(26), $B = \sum_{l=1}^{i} B_l(q_1, P, \tau)$, the quantity $\overline{\langle (B)^2 \rangle_k} = (n_i (\lambda_B^{(k)})^2 / \Delta v^i) \times \binom{2i-1}{i} i!\, \mu_i^{(k)}$. We study here the properties of $\overline{\langle B^2 \rangle_k}$ regarding arithmetic calculations.

Property 1.

$$\forall l \geqslant 1, \quad \overline{\left\langle \left(\frac{\partial B}{\partial P} \right)^2 \right\rangle_{k+l}} \leqslant \frac{a}{l^2} \overline{\langle B^2 \rangle_k} \tag{C1}$$

In order to show (C1), we calculate

$$\left\langle \left(\frac{\partial B}{\partial P} \right)^2 \right\rangle = \sum_{l_1=1}^{i} \sum_{l_2=1}^{i} \left\langle \frac{\partial B_{l_1}}{\partial P} \frac{\partial B_{l_2}}{\partial P} \right\rangle \tag{C2}$$

Using the notations (41)–(43), one finds

$$\sum_{l_1=1}^{i} \sum_{l_2=1}^{i} \left\langle \frac{\partial B_{l_1}}{\partial P} \frac{\partial B_{l_2}}{\partial P} \right\rangle = \frac{1}{\Delta v^i} \sum_{\varepsilon_1} \cdots \sum_{\{0\}\ \varepsilon_{2i}} \sum_{m \in \overline{m_+}} \sum_{\sigma_1} \sum_{\sigma_2} \sum_{j_1=1}^{n_{l_1}} \sum_{j_2=1}^{n_{l_2}} \frac{df_{j_1 l_1}}{dP} \frac{df_{j_2 l_2}}{dP} \tag{C3}$$

For any $l \geqslant 1$ the functions f_{j_1, l_1} and f_{j_2, l_2} being analytic on the disk $D_{i+l,a}$ the Cauchy inequalities[13] yield

$$\left\| \frac{df_{j_1, l_1}}{dP} \right\|_{k+l} \leqslant \frac{\sqrt{a}}{l} \|f_{j_1, l_1}\|_k \tag{C4}$$

and a similar inequality for f_{j_2, l_2}. Let now $\Lambda = \{\lambda \geqslant 0/\forall \overline{m_{1,2i}},\ \|f_{j_1,l_1}\|_k \leqslant \lambda \prod_{n=1}^{l_1} \|1/((\sigma(m_n)/\Delta v - P)\|_k\}$, $\lambda_{j_1,l_1}^{(k)}(\overline{\varepsilon_{1,i}}) = \inf(\Lambda)$, $\Lambda' = \{\lambda \geqslant 0/\forall \overline{m_{1,2i}}$, $\|df_{j_1,l_1}/dP\|_{k+l} \leqslant \lambda \prod_{n=1}^{l_1} \|1/((\sigma(m_n)/\Delta v - P)\|_{k+l}\}$, and $\lambda'^{(k)}_{j_1,l_1}(\overline{\varepsilon_{1,i}}) = \inf(\Lambda')$. Then, from (C4), it is clear that whatever $\overline{m_{1,2i}}$ and $\overline{\varepsilon_{1,i}}$

$$\lambda'^{(k+l)}_{j_1,l_1} \prod_{n=1}^{l_1} \left\| \frac{1}{\sigma(m_n)/\Delta v - P} \right\|_{k+l} \leqslant \frac{\sqrt{a}}{l} \lambda^{(k)}_{j_1,l_1} \prod_{n=1}^{l_1} \left\| \frac{1}{\sigma(m_n)/\Delta v - P} \right\|_k \tag{C5}$$

Therefore, using the definition (48)–(49) for $\lambda_B^{(k)}$, and a similar definition for $\lambda_{dB/dP}^{(k+l)}$, (C5) implies that

$$\lambda^{(k+l)}_{dB/dP} \prod_{n=1}^{l_1} \left\| \frac{1}{\sigma(m_n)/\Delta v - P} \right\|_{k+l} \leqslant \frac{\sqrt{a}}{l} \lambda^{(k)}_{B} \prod_{n=1}^{l_1} \left\| \frac{1}{\sigma(m_n)/\Delta v - P} \right\|_k \tag{C6}$$

Then (C1) follows from the very definitions of $\overline{\langle (\partial B/\partial P)^2 \rangle_{k+l}}$ and $\overline{\langle B^2 \rangle_k}$.

Property 2. Let $B_1, B_2, \ldots, B_n$ be of the form (25)–(26), then $\overline{\langle (\sum_{j=1}^n B_j)^2 \rangle_k}$ is obtained by using in (41) $f_{l_1} = f_{l_1}^{(B_1)} + f_{l_1}^{(B_2)} + \cdots + f_{l_1}^{(B_n)}$, and doing the same kind of substitution for f_{l_2}. The triangular inequality implies that $\|f_{l_1}\|_k \leqslant \sum_{j=1}^n \|f_{l_1}^{(B_j)}\|_k \leqslant (\sum_{j=1}^n \lambda_{B_j}^{(k)}) \prod_{j=1}^{l_1} \|1/((\sigma(m_j)/\Delta v - P)\|_k$. Hence, $\lambda_{B_1+B_2+\cdots+B_n} \leqslant \lambda_{B_1} + \lambda_{B_2} + \cdots + \lambda_{B_n}$, and $\lambda^2_{B_1+B_2+\cdots+B_n} \leqslant (\sqrt{\lambda^2_{B_1}} + \sqrt{\lambda^2_{B_2}} + \cdots + \sqrt{\lambda^2_{B_n}})^2$, which yields

$$\overline{\langle (B_1 + B_2 + \cdots + B_n)^2 \rangle_k} \leqslant \left(\sqrt{\overline{\langle B_1^2 \rangle_k}} + \sqrt{\overline{\langle B_2^2 \rangle_k}} + \cdots + \sqrt{\overline{\langle B_n^2 \rangle_k}}\right)^2 \tag{C7}$$

Property 3. Let C be a function of the form (25)–(26), $C = \sum_{l=1}^{i} C_l(q_1, P, \tau)$, with $j \geqslant i$. Then BC writes

$$BC = \sum_{L=2}^{i+j} D_L \tag{C8}$$

$$D_L = \sum_{l_1=1}^{i} \sum_{\substack{l_2=1 \\ l_1+l_2=L}}^{j} B_{l_1} C_{l_2} \tag{C9}$$

and

$$\langle (BC)^2 \rangle = \sum_{L_1=2}^{i+j} \sum_{L_2=2}^{i+j} D_{L_1} D_{L_2} \tag{C10}$$

The only non-zero terms in (C10) are those such that $(L_1 + L_2)$ is even. The number of such terms is

$$\begin{cases} n_{ij} = (i+j-1)^2/2 + 1/2, & \text{if } i+j \text{ is even,} \\ n_{ij} = (i+j-1)^2/2, & \text{if } i+j \text{ is odd} \end{cases} \tag{C11}$$

In any case, using (40), it is clear that

$$\frac{n_{ij}}{n_i n_j} \leqslant 2\,\frac{(i+j)^2}{i^2 j^2} \tag{C12}$$

Let us now calculate $\langle D_{L_1} D_{L_2} \rangle = (1/\Delta v^{i+j}) \sum_{\varepsilon_1} \underset{\{0\}}{\cdots} \sum_{\varepsilon_{2i+2j}} \sum_{m \in \overline{m_+}} \sum_{\sigma_1} \times \sum_{\sigma_2} f_{L_1} f_{L_2}$, where

$$f_{L_1} = \sum_{l_1=1}^{i} \sum_{\substack{l_2=1 \\ l_1+l_2=L_1}}^{j} \frac{f_{l_1}^{(B)} f_{l_2}^{(C)}}{\sqrt{2}} \tag{C13}$$

$$f_{L_2} = \sum_{l_3=1}^{i} \sum_{\substack{l_4=1 \\ l_3+l_4=L_2}}^{j} \frac{f_{l_3}^{(B)} f_{l_4}^{(C)}}{\sqrt{2}} \tag{C14}$$

In f_{L_1} or f_{L_2} there are at most i terms. Then, because $\|f_{l_1}^{(B)} f_{l_2}^{(C)}\|_k \leqslant \lambda_B^{(k)} \lambda_C^{(k)} \times \prod_{n=1}^{L_1} \|1/((\sigma(m_n)/\Delta v - P)\|_k$, $\|f_{L_1}\|_k \leqslant (i/\sqrt{2})\, \lambda_B^{(k)} \lambda_C^{(k)} \prod_{n=1}^{L_1} \|1/(\sigma(m_n)/\Delta v - P)\|_k$, and a similar inequality holds for f_{L_2}. This implies that

$$\lambda_{BC}^{(k)} \leqslant \frac{i}{\sqrt{2}}\, \lambda_B^{(k)}\, \lambda_C^{(k)} \tag{C15}$$

Moreover, $\mu_{ij}^{(k)} = \max_{L_1, L_2} \sum_{m_1} \cdots \sum_{m_{i+j}} \prod_{n=1}^{(L_1+L_2)/2} \|1/(m_n/\Delta v - P)^2)\|_k$ is

$$\begin{cases} \mu_{ij}^{(k)} = \left(\displaystyle\sum_{|m/\Delta v - P^*| > 1} \left\| \frac{1}{(m/\Delta v - P)^2} \right\|_k \right)^{i+j}, \\ \qquad \text{if } \displaystyle\sum_{|m/\Delta v - P^*| > 1} \left\| \frac{1}{(m/\Delta v - P)^2} \right\|_k \geqslant \sum_{|m/\Delta v - P^*| \leqslant 1} 1 \\ \mu_{ij}^{(k)} = \left(\displaystyle\sum_{|m/\Delta v - P^*| \leqslant 1} 1 \right)^{i+j-2} \left(\displaystyle\sum_{|m/\Delta v - P^*| > 1} \left\| \frac{1}{(m/\Delta v - P)^2} \right\|_k \right)^2 \\ \qquad \text{in the opposite case} \end{cases}$$

In any case

$$\mu_{ij}^{(k)} = \mu_i^{(k)} \mu_j^{(k)} \tag{C16}$$

It is easy to show that

$$\sum_{\varepsilon_1} \cdots \sum_{\{0\}} \sum_{\varepsilon_{2i+2j}} \sum_{\sigma_1} \sum_{\sigma_2} 1 \leqslant \binom{2i+2j-1}{i+j} (i+j)! \tag{C17}$$

Then, using (C11), (C15), (C16) and (C17), one finds

$$\begin{aligned}\overline{\langle (BC)^2 \rangle_k} &= n_{ij} (\lambda_{BC}^{(k)})^2 \frac{\mu_{ij}^{(k)}}{\Delta v^{i+j}} \binom{2i+2j-1}{i+j} (i+j)! \\ &\leqslant \frac{2(i+j)^2}{i^2 j^2} n_i n_j \frac{i^2}{2} (\lambda_B^{(k)})^2 (\lambda_C^{(k)})^2 \frac{\mu_i^{(k)}}{\Delta v^i} \frac{\mu_j^{(k)}}{\Delta v^j} \binom{2i+2j-1}{i+j} (i+j)!\end{aligned} \tag{C18}$$

Using definition (53) for $\overline{\langle B^2 \rangle_k}$ and $\overline{\langle C^2 \rangle_k}$ (C18) yields

$$\overline{\langle (BC)^2 \rangle_k} \leqslant \frac{(i+j)^2}{[\max(i, j)]^2} \frac{\binom{2i+2j-1}{i+j} (i+j)!}{\binom{2i-1}{i} \binom{2j-1}{j} i!\, j!} \overline{\langle B^2 \rangle_k}\, \overline{\langle C^2 \rangle_k} \tag{C19}$$

In order to conclude the proof of Property 3, one has to consider the case where C is a derivative of h_1 with respect to Q:

$$C = \frac{\partial^l h_1}{\partial Q^l} = \frac{1}{\sqrt{\Delta v}} \sum_{|m/\Delta v - P^*| \leqslant 1} \cos(Q - m\tau/\Delta v + l\pi/2) \tag{C20}$$

Because the value of the phase $l\pi/2$ in (C20) does not change the value of $\overline{\langle (\partial h_1^l / \partial Q^l)^2 \rangle_k}$, one can restrict to the case where $C = h_1$. Hence we calculate

$$\langle (Bh_1)^2 \rangle = \sum_{l_1=1}^{i} \sum_{l_2=1}^{i} \langle D_{l_1} D_{l_2} \rangle \tag{C21}$$

$$\langle D_{l_1} D_{l_2} \rangle = \frac{1}{\Delta v^{i+1}} \sum_{\varepsilon_1} \cdots \sum_{\{0\}} \sum_{\varepsilon_{2i+2}} \sum_{\sigma_1} \sum_{\sigma_2} \sum_{m \in \overline{m_+}} \sum_{j_1=1}^{n_{l_1}} \sum_{j_2=1}^{n_{l_2}} \frac{f_{j_1, l_1}^{(B)} f_{j_2, l_2}^{(B)}}{2} \tag{C22}$$

the number of non-zero terms in (C22) is n_l. Moreover, from (C22) it is clear that

$$\lambda_{Bh_1}^{(k)} = \lambda_B^{(k)}/\sqrt{2} \tag{C23}$$

Finally, we evaluate

$$\begin{aligned}
&\max_{l_1, l_2} \sum_{m_1} \cdots \sum_{m_{i+1}} \prod_{j=2}^{1+(l_1+l_2)/2} \left\| \frac{1}{(m_j/\Delta v - P)^2} \right\|_k \\
&= \mu_1 \max_{l_1, l_2} \sum_{m_2} \cdots \sum_{m_{i+1}} \left\| \frac{1}{(m_j/\Delta v - P)^2} \right\|_k \\
&= \mu_1 \mu_i^{(k)}
\end{aligned} \tag{C24}$$

From (C23) and (C24) it follows that

$$\begin{aligned}
\overline{\langle (Bh_1)^2 \rangle_k} &= n_i \binom{2i+1}{i+1} (i+1)! \frac{(\lambda_B^{(k)})^2}{2} \frac{\mu_i}{\Delta v^i} \frac{\mu_1}{\Delta v} \\
&\leqslant \frac{(i+1)^2}{[\max(i, 1)]^2} \frac{\binom{2i+1}{i+1}}{\binom{2i-1}{i}\binom{2i-1}{1}} \frac{(i+1)!}{i!\, 1!} \overline{\langle B^2 \rangle_k}\, \overline{\langle h_1^2 \rangle_k}
\end{aligned} \tag{C25}$$

which is the same as (C19) with $j = 1$.

Property 4. Let B be a function of the form (25)–(26), then $\partial^m B_l / \partial Q^m$ writes

$$\frac{\partial B_l}{\partial Q^m} = \frac{1}{\Delta v^{i/2}} \sum_{\varepsilon_1} \cdots \sum_{\varepsilon_i} \sum_{m_1} \cdots \sum_{m_i} \sum_{j=1}^{n_l} \frac{b_{j,l} \left(\sum_{k=1}^{i} \varepsilon_k \right)^m \cos\left[\sum_{k=1}^{i} \varepsilon_k \xi_k + (v+m)\frac{\pi}{2} \right]}{\prod_{k=1}^{l} \left(\frac{m_k}{\Delta v} - P \right)^{\alpha_{j,k,l}} \Delta_{j,l}(\overline{m_{1,i}}, \overline{\varepsilon_{1,i}}, P)} \tag{C26}$$

$\overline{\langle (\partial^m B_l / \partial Q^m)^2 \rangle_k}$ is thus deduced from $\overline{\langle (B)^2 \rangle_k}$ by changing $b_{j,l}$ in $b_{j,l} (\sum_{k=1}^{i} \varepsilon_k)^m$. Now, because $|b_{j,l} (\sum_{k=1}^{i} \varepsilon_k)^m| \leqslant i^m |b_{j,l}|$, it is clear that $\lambda_{\partial^m B / \partial q_1^m}^{(k)} \leqslant i^m \lambda_B^{(k)}$, which implies

$$\overline{\left\langle \left(\frac{\partial^m B}{\partial Q^m} \right)^2 \right\rangle_k} \leqslant i^{2m}\, \overline{\langle B^2 \rangle_k} \tag{C27}$$

Property 5. Let B be a function of the form (25)–(26) which we write $B(Q, P, \tau) = B^{(1)}(Q, P, \tau) + B^{(2)}(P, \tau)$, with

$$\frac{\partial B^{(1)}}{\partial Q} = \sum_{l=1}^{i} \sum_{\varepsilon_1} \cdots \sum_{\varepsilon_i} \sum_{m_1} \cdots \sum_{m_i} \sum_{j=1}^{n_l^{(1)}} f_{j,l}^{(1)} \cos\left(\sum_{n=1}^{i} \varepsilon_n \xi_n \right) \tag{C28}$$

$$B^{(2)} = \sum_{l=1}^{i} \sum_{\varepsilon_1} \cdots \sum_{\varepsilon_i} \sum_{m_1} \cdots \sum_{m_i} \sum_{j=1}^{n_l^{(2)}} f_{j,l}^{(2)} \cos\left(\sum_{n=1}^{i} \varepsilon_n \xi_n \right) \tag{C29}$$

where in (C28) $\sum_{k=1}^{i} \varepsilon_k \neq 0$, while in (C29) $\sum_{k=1}^{i} \varepsilon_k = 0$.

$$\left\langle \left(\frac{\partial B^{(1)}}{\partial Q} + B^{(2)} \right)^2 \right\rangle = \left\langle \left(\frac{\partial B^{(1)}}{\partial Q} \right)^2 \right\rangle + \langle (B^{(2)})^2 \rangle + 2 \left\langle \frac{\partial B^{(1)}}{\partial Q} B^{(2)} \right\rangle \tag{C30}$$

and

$$2 \left\langle \frac{\partial B^{(1)}}{\partial Q} B^{(2)} \right\rangle = \sum_{l_1=1}^{i} \sum_{l_2=1}^{i} \sum_{\varepsilon_1} \cdots \sum_{\varepsilon_{2i}} \sum_{m_1} \cdots \sum_{m_{2i}} \sum_{j_1=1}^{n_{l_1}^{(1)}} \sum_{j_2=1}^{n_{l_2}^{(2)}} f_{j_1, l_1}^{(1)} f_{j_2, l_2}^{(2)}$$
$$\times \left\langle \cos\left(\sum_{n=1}^{2i} \varepsilon_n \xi_n \right) \right\rangle \tag{C31}$$

Now, in (C31) $\sum_{k=1}^{i} \varepsilon_k \neq 0$ and $\sum_{k=1+i}^{2i} \varepsilon_k = 0$, which implies that $\sum_{k=1}^{2i} \varepsilon_k \neq 0$ and therefore $\langle \partial B^{(1)}/\partial Q\, B^{(2)} \rangle = 0$. Hence, $\langle (\partial B^{(1)}/\partial Q + B^{(2)})^2 \rangle = \langle (\partial B^{(1)}/\partial Q)^2 \rangle + \langle (B^{(2)})^2 \rangle$, which implies that

$$\lambda_{\partial B^{(1)}/\partial Q + B^{(2)}} = \max\{ \lambda_{\partial B^{(1)}/\partial Q}, \lambda_{B^{(2)}} \} \tag{C32}$$

Similarly, one can show that

$$\lambda_B = \lambda_{B^{(1)} + B^{(2)}} = \max\{ \lambda_{B^{(1)}}, \lambda_{B^{(2)}} \} \tag{C33}$$

Moreover, when deriving Property 4 we showed that $\lambda_{\partial B^{(1)}/\partial Q} \geqslant \lambda_{B^{(1)}}$, so that finally (C31) and (C32) yield

$$\overline{\left\langle \left(\frac{\partial B^{(1)}}{\partial Q} \right)^2 \right\rangle_k} \leqslant \overline{\left\langle \left(\frac{\partial B^{(1)}}{\partial Q} + B^{(2)} \right)^2 \right\rangle_k} \tag{C34}$$

$$\overline{\langle (B)^2 \rangle_k} \leqslant \overline{\left\langle \left(\frac{\partial B^{(1)}}{\partial Q} + B^{(2)} \right)^2 \right\rangle_k} \tag{C35}$$

APPENDIX D: ESTIMATE OF $\overline{\langle X_i^2 \rangle_i}$

The function $X_i(Q, P, \tau)$ can be written as

$$X_i(Q, P, \tau) = Z_i(Q, P, \tau) - W_i(Q, P, \tau) \tag{D1}$$

where

$$W_i = \frac{1}{2} \sum_{j=1}^{i-1} \frac{\partial \Phi_j}{\partial Q} \frac{\partial \Phi_{i-j}}{\partial Q} \tag{D2}$$

$$Z_i = \sum_{l=1}^{i-1} \sum_{m=1}^{i-l} \frac{\partial^m h_l}{\partial Q^m} \frac{1}{m!} \sum_{i_1} \underset{\{1, i-l\}}{\cdots} \sum_{i_m} \prod_{j=1}^{m} \frac{\partial \Phi_{i_j}}{\partial P} \tag{D3}$$

the symbol $\{1, i-l\}$ meaning that the i_j's are such that

$$\begin{cases} i_1 \geqslant 1, \ldots, i_m \geqslant 1 \\ \sum_{j=1}^{m} i_j = i - l \end{cases} \tag{D4}$$

In order to find an estimate for $\overline{\langle X_i^2 \rangle_i}$, we begin to evaluate $\overline{\langle W_i^2 \rangle_i}$. From (59) one finds

$$\overline{\left\langle \left(\frac{\partial \Phi_j}{\partial Q} \frac{\partial \Phi_{i-j}}{\partial Q} \right)^2 \right\rangle_i} \leqslant \frac{i^2}{[\max(j, i-j)]^2} \frac{\binom{2i-1}{i} i!}{\binom{2(i-j)-1}{(i-j)} \binom{2j-1}{j} (i-j)!\, j!} \times \overline{\left\langle \left(\frac{\partial \Phi_{i-j}}{\partial Q} \right)^2 \right\rangle_{i-j}} \; \overline{\left\langle \left(\frac{\partial \Phi_j}{\partial Q} \right)^2 \right\rangle_j} \tag{D5}$$

Then, if for $j \leqslant i-1$, $\overline{\langle (\partial \Phi_j / \partial Q)^2 \rangle_j} \leqslant F \sigma^j (j!)^3 \binom{2j-1}{j}$, (D5) yields

$$\overline{\left\langle \left(\frac{\partial \Phi_j}{\partial q_1} \frac{\partial \Phi_{i-j}}{\partial q_1} \right)^2 \right\rangle_i} \leqslant 4 \binom{2i-1}{i} i!\, F^2 \sigma^i (j!)^2 ((i-j)!)^2 \tag{D6}$$

Using (57), it follows from (D6) that

$$\overline{\langle W_i^2 \rangle_i} \leqslant \binom{2i-1}{i} i!\, F^2 \sigma^i \left(\sum_{j=1}^{i-1} j!\, (i-j)! \right)^2 \tag{D7}$$

Now, it is easy to show by induction[7] that

$$S_{k,n} = \sum_{\substack{i_1,\ldots,i_k \geqslant 1 \\ i_1 + \cdots + i_k = n}} \prod_{j=1}^{k} i_k! \leqslant n! \tag{D8}$$

Therefore

$$\overline{\langle W_i^2 \rangle_i} \leqslant \binom{2i-1}{i} F^2 \sigma^i (i!)^3 \tag{D9}$$

Let us now evaluate $\overline{\langle Z_i^2 \rangle_i}$. Using (56) and (59), as well as (65), one finds

$$\begin{aligned} \overline{\left\langle \left(\frac{\partial \Phi_{i_1}}{\partial P} \frac{\partial \Phi_{i_2}}{\partial P} \right)^2 \right\rangle_i} \leqslant & \frac{(i_1+i_2)^2}{[\max(i_1, i_2)]^2} \binom{2(i_1+i_2)-1}{i_1+i_2} \\ & \times (i_1+i_2)!\, a^2 F^2 \sigma^{i_1+i_2} (i_1!)^2 (i_2!)^2 \end{aligned} \tag{D10}$$

Using several times (59), and making the product of the $(\partial \Phi_{i_j}/\partial \mathrm{P})$'s after sorting the i_j's in decreasing order, one finds

$$\begin{aligned} \overline{\left\langle \left(\prod_{j=1}^{m} \frac{\partial \Phi_{i_j}}{\partial P} \right)^2 \right\rangle_i} \leqslant & \frac{(i-l)^2}{(i_1+i_2+\cdots+i_{m-1})^2} \binom{2(i-l)-1}{i-l} (i-l)!\, a^m \\ & \times \prod_{j=1}^{m} (i_j!)^2 F^m \sigma^{i-l} \end{aligned} \tag{D11}$$

In (D11) $\sum_{j=1}^{m} i_j = i-l$, hence, when $m \geqslant 2$, $i_1 + \cdots + i_{m-1} \geqslant (m-1)(i-l)/m$. Therefore,

$$\overline{\left\langle \left(\prod_{j=1}^{m} \frac{\partial \Phi_{i_j}}{\partial P} \right)^2 \right\rangle_i} \leqslant m^2 \binom{2(i-l)-1}{i-l} (i-l)!\, (aF)^m \sigma^{i-l} \prod_{j=1}^{m} (i_j!)^2 \tag{D12}$$

Then, from (57) and (D8) it follows that

$$\begin{aligned} & \overline{\left\langle \left(\sum_{i_1} \underset{\{1, i-l\}}{\cdots} \sum_{i_m} \left(\prod_{j=1}^{m} \frac{\partial \Phi_{i_j}}{\partial P} \right) \right)^2 \right\rangle_i} \\ & \leqslant m^2 \binom{2(i-l)-l}{i-l} \sigma^{i-l} (aF)^m [(i-l)!]^3 \end{aligned} \tag{D13}$$

Using (59) and (60), one finds

$$\overline{\left\langle\left(\frac{\partial^m h_l}{\partial Q^m}\frac{1}{m!}\sum_{i_1}\underset{\{1,i-l\}}{\cdots}\sum_{i_m}\left(\prod_{j=1}^{m}\frac{\partial \Phi_{i_j}}{\partial P}\right)\right)^2\right\rangle_i}$$
$$\leqslant 4\binom{2i-1}{i} i!\, F\sigma^i (l!)^2 [(i-l)!]^2 \frac{(aF)^m}{[(m-1)!]^2} l^{2m} \tag{D14}$$

and then, using (57)

$$\overline{\left\langle\left(\sum_{m=1}^{i-l}\frac{\partial^m h_l}{\partial Q^m}\frac{1}{m!}\sum_{i_1}\underset{\{1,i-l\}}{\cdots}\sum_{i_m}\left(\prod_{j=1}^{m}\frac{\partial \Phi_{i_j}}{\partial P}\right)\right)^2\right\rangle_i}$$
$$\leqslant 4\binom{2(i-l)-1}{i-l} i!\, F\sigma^i (l!)^2 [(i-l)!]^2 aFl^2 \left[\sum_{m=1}^{i-l}\frac{(\sqrt{aFl^2})}{(m-1)!}\right]^2 \tag{D15}$$

F is then chosen so that

$$F \leqslant \frac{1}{an^2} \tag{D16}$$

which implies $\sum_{m=1}^{i-l}(\sqrt{aFl^2})^{m-1}/(m-1)! \leqslant \exp(\sqrt{aFl^2}) \leqslant e$. Therefore,

$$\overline{\left\langle\left(\sum_{m=1}^{i-l}\frac{\partial^m h_l}{\partial Q^m}\frac{1}{m!}\sum_{i_1}\underset{\{1,i-l\}}{\cdots}\sum_{i_m}\left(\prod_{j=1}^{m}\frac{\partial \Phi_{i_j}}{\partial P}\right)\right)^2\right\rangle_i}$$
$$\leqslant 4\binom{2(i-l)-1}{i-l} i!\, F\sigma^i aFe^2 (i-1)^2 (l!)^2 [(i-l)!]^2 \tag{D17}$$

Finally, using (57) and (D8), one finds

$$\overline{\langle Z_i^2\rangle_i} \leqslant 4aFe^2(i-1)^2 F\sigma^i \binom{2(i-l)-1}{i-l} (i!)^3 \tag{D18}$$

using definition (D1) of X_i, inequalities (D9) and (D18), together with (57), yield

$$\overline{\langle X_i^2\rangle_i} \leqslant F(1+2e\sqrt{a}\,(i-1))^2 F\sigma^i \binom{2i-1}{i} (i!)^3$$
$$\leqslant 4e^2 Fai^2 F\sigma^i \binom{2i-1}{i} (i!)^3 \tag{D19}$$

because $\sqrt{a} < n \geqslant 1$. Now, if the condition

$$F \leqslant \frac{1}{4e^2 a^2} \tag{D20}$$

is imposed, then

$$\overline{\langle X_i^2 \rangle_i} \leqslant \frac{i^2}{a} F\sigma^i \binom{2i-1}{i} (i!)^3 \tag{D21}$$

Note that (D20) implies (D16) because $a > n^2$. Moreover, (D16) is exactly condition (67). This proves that if (67) is satisfied, then inequality (68) is fulfilled.

APPENDIX E: EXPONENTIAL ESTIMATE OF THE REMAINDER

The exponential estimate for the remainder is obtained by using the following inequality, valid for $\varepsilon \leqslant 1$

$$\sqrt{\langle R^2(\varepsilon) \rangle} \leqslant \varepsilon^{n+1} \sup_{\varepsilon' \leqslant 1} \sqrt{\langle R^2(\varepsilon') \rangle} \tag{E1}$$

and by using value (83) for n. In order to bound $\sup_{\varepsilon' \leqslant 1} \sqrt{\langle R^2(\varepsilon') \rangle}$ from above, we use inequality (37)

$$\begin{aligned} \langle R^2(Q, P, \tau) \rangle \leqslant {} & (6n+3) \langle \mathfrak{R}_\alpha^2(Q, P, \tau) \rangle + 6n \langle \mathfrak{R}_\alpha'^2(Q, P, \tau) \rangle \\ & + 3 \langle \mathfrak{R}_\beta^2(Q, P, \tau) \rangle \end{aligned} \tag{E2}$$

derived in Appendix B, where $\mathfrak{R}_\alpha$, $\mathfrak{R}_\beta$ and $\mathfrak{R}'_\alpha$ are defined by (33), (34) and (38) respectively. The aim of this appendix is to find an upper bound for the right hand side of (E2). Writing $\mathfrak{R}_\alpha = \sum_i r_{\alpha, i}$, where $r_{\alpha, i}$ comes from the ith order of the perturbation analysis, we will actually derive an upper bound for $\langle \mathfrak{R}_\alpha^2 \rangle$ on the disk $D_{n, a}$ in terms of the $\overline{\langle r_{\alpha, i}^2 \rangle_n}$'s. Now, it is clear that if $r'_{\alpha, i}$ is the term of $\mathfrak{R}'_\alpha$ corresponding to $r_{\alpha, i}$, $\overline{\langle r'^2_{\alpha, n} \rangle_n} = \overline{\langle r^2_{\alpha, n} \rangle_n}$. Therefore, the upper bound we will derive for $\langle \mathfrak{R}_\alpha^2 \rangle$ will also hold for $\langle \mathfrak{R}'^2_\alpha \rangle$ on $D_{n, a}$. Moreover, because we only consider the case where $n \geqslant 1$ (otherwise $R = 0$), we can, in practice, replace (E2) by

$$\langle R^2(Q, P, \tau) \rangle \leqslant 15n \langle \mathfrak{R}_\alpha^2(Q, P, \tau) \rangle + 3 \langle \mathfrak{R}_\beta^2(Q, P, \tau) \rangle \tag{E3}$$

In order to estimate $\langle \mathfrak{R}_\alpha^2 \rangle$ and $\langle \mathfrak{R}_\beta^2 \rangle$ one cannot directly use inequality (57) because it was derived in the case where all the B_l's came from the same order of perturbation. Let us then consider here $B = B^{(i)} = \sum_{l=1}^{i} B_l$ and $C = C^{(j)} = \sum_{l=1}^{j} C_l$, coming respectively from the orders i and j and let us calculate $\langle B^{(i)}C^{(j)}\rangle$. If $(i+j)$ is odd, then it is clear that condition (A8), $2i$ being replaced by $(i+j)$, cannot be fulfilled, which implies that $\langle B^{(i)}C^{(j)}\rangle = 0$. If $(i+j)$ is even, we calculate

$$\langle B^{(i)}C^{(j)}\rangle = \sum_{l_1=1}^{i} \sum_{l_2}^{j} \langle B_{l_1} C_{l_2}\rangle \tag{E4}$$

Using a notation similar to (27), we denote

$$B_l = \frac{1}{\Delta v^{i/2}} \sum_{\varepsilon_1} \cdots \sum_{\varepsilon_i} \sum_{m_1} \cdots \sum_{m_i} b_l(P, \overline{m_{1,i}}) \cos\left(\sum_{k=1}^{i} \varepsilon_k \xi_k\right) \tag{E5a}$$

$$C_l = \frac{1}{\Delta v^{j/2}} \sum_{\varepsilon_1} \cdots \sum_{\varepsilon_j} \sum_{m_1} \cdots \sum_{m_j} c_l(P, \overline{m_{1,j}}) \cos\left(\sum_{k=1}^{j} \varepsilon_k \xi_k\right) \tag{E5b}$$

In (E4), the only non-zero terms are those such that $(l_1 + l_2)$ is even. The number of such terms is

$$\begin{cases} N = ij/2, & \text{if } i \text{ and } j \text{ are even} \\ N = (ij+1)/2, & \text{if } i \text{ and } j \text{ are odd} \end{cases} \tag{E6}$$

In any case, one can note from (40) that

$$N \leqslant \sqrt{n_i}\,\sqrt{n_j} \tag{E7}$$

As for $\langle B_{l_1} C_{l_2}\rangle$, we estimate, in a similar way as in Appendix A

$$\begin{aligned} |\langle B_{l_1} C_{l_2}\rangle| \leqslant & \sum_{\substack{\varepsilon_1, \dots, \varepsilon_{i+j} \\ \varepsilon_1 + \cdots + \varepsilon_{i+j} = 0}} \sum_{\sigma_1} \sum_{\sigma_2} \sum_{m_1, \dots, m_{(i+j)/2}} |b_{l_1}(P, \sigma(\overline{m_{1,i}}))| \\ & \times |c_{l_2}(P, \sigma(\overline{m_{1+i, i+j}}))| \end{aligned} \tag{E8}$$

Now, from the very definition of $\lambda_B^{(n)}$ and $\lambda_C^{(n)}$, (49), for any P of the disk $D_{n,a}$,

$$|b_{l_1}|\,|c_{l_2}| \leqslant \lambda_B^{(n)} \lambda_C^{(n)} \prod_{k=1}^{(l_1+l_2)/2} \left\| \frac{1}{(m_k/\Delta v - P)^2} \right\|_n \tag{E9}$$

We then denote

$$\begin{cases} \mu^{(n)}_{\sqrt{ij}} = \left(\displaystyle\sum_{|m/\Delta v - P^*| > 1} \left\| \frac{1}{(m/\Delta v - P)^2} \right\|_n \right)^{(i+j)/2}, \\ \qquad \text{if } \displaystyle\sum_{|m/\Delta v - P^*| > 1} \left\| \frac{1}{(m/\Delta v - P)^2} \right\|_n \geqslant \sum_{|m/\Delta v - P^*| \leqslant 1} 1 \\ \mu^{(n)}_{\sqrt{ij}} = \left(\displaystyle\sum_{|m/\Delta v - P^*| \leqslant 1} 1 \right)^{(i+j)/2 - 1} \left(\displaystyle\sum_{|m/\Delta v - P^*| > 1} \left\| \frac{1}{(m/\Delta v - P)^2} \right\|_n \right), \\ \qquad \text{in the opposite case} \end{cases}$$

and, in any case

$$\mu^{(n)}_{\sqrt{ij}} \leqslant \sqrt{\mu^{(n)}_i} \sqrt{\mu^{(n)}_j} \tag{E10}$$

then, using the fact that $\sum_{\substack{\varepsilon_1, \ldots, \varepsilon_{i+j} \\ \varepsilon_1 + \cdots + \varepsilon_{i+j=0}}} \sum_{\sigma_1} \sum_{\sigma_2} \leqslant \binom{(i+j)-1}{(i+j)/2} (i+j)/2)!$, one finds that for any P in $D_{n,a}$

$$|\langle B_{l_1} C_{l_2} \rangle| \leqslant \lambda^{(n)}_B \lambda^{(n)}_C \sqrt{\mu^{(n)}_i} \sqrt{\mu^{(n)}_j} \binom{(i+j)-1}{(i+j)/2} ((i+j)/2)! \tag{E11}$$

Then, from the definition (53) of $\overline{\langle (\ )^2 \rangle_n}$, and the inequality (E7), it follows that for any P in $D_{n,a}$

$$|\langle B^{(i)} C^{(j)} \rangle| \leqslant \frac{\binom{(i+j)-1}{(i+j)/2}}{\sqrt{\binom{2i-1}{1} i! \binom{2j-1}{j} j!}} \left(\frac{i+j}{2} \right)! \sqrt{\overline{\langle (B^{(i)})^2 \rangle_n}} \sqrt{\overline{\langle (C^{(j)})^2 \rangle_n}} \tag{E12}$$

Now, it can be easily shown[7] that

$$\frac{\binom{(i+j)-1}{(i+j)/2}}{\sqrt{\binom{2i-1}{i} i! \binom{2j-1}{j} j!}} \left(\frac{i+j}{2} \right)! \leqslant 1 \tag{E13}$$

Therefore, for any P in $D_{n,a}$

$$|\langle B^{(i)} C^{(j)} \rangle| \leqslant \sqrt{\overline{\langle (B^{(i)})^2 \rangle_n}} \sqrt{\overline{\langle (C^{(j)})^2 \rangle_n}} \tag{E14}$$

Then, from (E14), it is clear that

$$\left\langle\left(\sum_k B_k^{(i_k)}\right)^2\right\rangle \leqslant \left(\sum_k \sqrt{\langle (B_k^{(i_k)})^2\rangle_n}\right)^2 \tag{E15}$$

using the definitions of $\mathfrak{R}_\alpha$ and $\mathfrak{R}_\beta$, and the inequality (E15), one finds

$$\langle \mathfrak{R}_\alpha^2(Q, P, \tau)\rangle \leqslant \left[\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \varepsilon^{i+j}\sqrt{\left\langle\left(\frac{\partial\Phi_i}{\partial Q}\frac{\partial\Phi_j}{\partial Q}\right)^2\right\rangle_n} + \sum_{i=1}^{n}\varepsilon^i\sqrt{\langle h_i^2\rangle_n} + \sum_{i=1}^{n}\varepsilon^i\sqrt{\langle X_i^2\rangle_n}\right]^2 \tag{E16}$$

$$\langle \mathfrak{R}_\beta^2(Q, P, \tau)\rangle \leqslant \left[\sum_{i=1}^{n}\varepsilon^i\sqrt{\langle h_i^2(Q, P, \tau)\rangle_n}\right]^2 \tag{E17}$$

Now, using (59), (60) and (66), as well as inequality $\binom{2(i+j)-1}{i+j} \leqslant 4^{i+j-1}$ one finds

$$\varepsilon^{2i+2j}\left\langle\left(\frac{\partial\Phi_i}{\partial Q}\frac{\partial\Phi_j}{\partial Q}\right)^2\right\rangle_n \leqslant F(4\sigma\varepsilon^2)^i\,(i!)^2\,F(4\sigma\varepsilon^2)^j\,(j!)^2\,(i+j)! \tag{E18}$$

Using the fact that for any $i \geqslant 1$, $i! \leqslant i^{i-1}$, it follows form (E18) that

$$\begin{aligned}
&\varepsilon^{2i+2j}\left\langle\left(\frac{\partial\Phi_i}{\partial Q}\frac{\partial\Phi_j}{\partial Q}\right)^2\right\rangle_n \\
&\quad\leqslant (i+j)(4\sigma F\varepsilon^2)(4\sigma i^2(i+j)\,\varepsilon^2)^{i-1}\,(4\sigma F\varepsilon^2)(4\sigma F\varepsilon^2)(4\sigma j^2(i+j)\,\varepsilon^2)^{j-1} \\
&\quad\leqslant 2n(4\sigma F\varepsilon^2)(8\sigma n^3\varepsilon^2)^{i-1}\,(4\sigma F\varepsilon^2)(8\sigma n^3\varepsilon^2)^{j-1} \\
&\quad\leqslant 2n(4\sigma F\varepsilon^2)^2
\end{aligned} \tag{E19}$$

where, in order to derive the last line of (E19), we used the values (79) and (83) of σ and n respectively, which imply that $8\sigma n^3\varepsilon^2 \leqslant 1$. Finally, one finds

$$\begin{aligned}
\frac{1}{2}\sum_{j=1}^{n}\sum_{j=1}^{n}\varepsilon^{i+j}\sqrt{\left\langle\left(\frac{\partial\Phi_i}{\partial Q}\frac{\partial\Phi_j}{\partial Q}\right)^2\right\rangle_n} &\leqslant 2\sqrt{2}\,\sigma F\varepsilon^2 n^{5/2} \\
&\leqslant 2^{-23/16}\varepsilon^{9/8}
\end{aligned} \tag{E20}$$

where the values (70), (79) and (83) of F, σ and n respectively have been used. Using the same lines as the ones used to derive (E20), it is easily shown that

$$\sum_{i=1}^{n} \varepsilon^i \sqrt{\langle h_i^2 \rangle_n} \leqslant 2^{-19/16} \varepsilon^{5/8} \tag{E21}$$

$$\sum_{i=2}^{n} \varepsilon^i \sqrt{\langle X_i^2 \rangle_n} \leqslant \varepsilon \tag{E22}$$

Then, using (E3), (E16), (E17), (E20), (E21) and (E22), one finds that for any P in $D_{n,a}$

$$\begin{aligned}\langle R^2 \rangle \leqslant\ & 3 \times 2^{-19/8} \varepsilon^{5/4} + 15(2^{-4}\varepsilon^2 + 2^{-7/2}\varepsilon + 2^{-9/8}\varepsilon^{7/4} \\ & + 2^{-11/4}\varepsilon^{3/2} + 2^{-25/16}\varepsilon^{15/8} + 2^{-21/16}\varepsilon^{11/8})\end{aligned} \tag{E23}$$

If are only considered values of ε less than unity, the right hand side of (E23) is maximum when $\varepsilon = 1$, which implies that for any P in $D_{n,a}$

$$\sup_{\varepsilon' \leqslant 1} \sqrt{\langle R^2(\varepsilon') \rangle} \leqslant 5 \tag{E24}$$

APPENDIX F: ESTIMATE OF $\langle (p/\Delta v - P)^2 \rangle$

From (11), it follows that $p/\Delta v + P = \sum_{i=1}^{n} \varepsilon^i \, \partial \Phi_i / \partial q_1$. Φ_i being of the type (25)–(26), if the change of variables is performed about $P = P_j$, $\partial \Phi_i / \partial q_1$ can be written as

$$\partial \Phi_i / \partial q_1 = \sum_{l=1}^{i} \cos[i(q_1 - P_j \tau)] \, \Phi_i^{(c)} + \sum_{l=1}^{i} \sin[i(q_1 - P_j \tau)] \, \Phi_i^{(s)} \tag{F1}$$

The $\Phi_i^{(c)}$'s and the $\Phi_i^{(s)}$'s are of the type (25)–(26) where $(Q - m_k/\Delta v \tau)$ is replaced by the phase-independent variable $(m_k/\Delta v - P_j)\,\tau$. By using the inequality $xy < (x^2 + y^2)/2$, one can see that if $\langle (\Phi_i^{(c)})^2 \rangle$ and $\langle (\Phi_i^{(s)})^2 \rangle$ can be bounded from above independently of M and Δv, $\langle (p/\Delta v - P)^2 \rangle$ can be made arbitrarily small by decreasing ε. From (26) one can see that if $|P - P_j| < 1$, $\Phi_i^{(c)}$ can be expanded in Taylor series about $P = P_j$

$$\Phi_i^{(c)} = \sum_{m=0}^{+\infty} \phi_m (P - P_j)^m / m! \tag{F2}$$

where $\phi_m = (\partial^m \Phi_i^{c} / \partial P^m)_{P = P_j}$. Now, it is important to remark that if in (26) one replaces all the cosines by 1, and changes the m_k's so that in the denominator there are only terms such that $(m_k/\Delta v) > 1$, one then defines a function which can be expanded in a Taylor series whose coefficients are upper bounds of the $|\phi_m|$'s. Therefore, $\sum_{m=0}^{+\infty} |\phi_m|\ |P-P_j|^m/m!$ converges, which implies that $\sum_{m=0}^{+\infty} [\phi_m(P-P_j)^m/m!]^2$ also converges. Then, when $|P-P_j| < 1-\delta$, it follows from (F2) that

$$\langle (\Phi_i^{(c)})^2 \rangle \leqslant 2 \sum_{m=0}^{+\infty} \{\langle (\phi_m)^2 \rangle\ [(1-\delta)^m/m!]^2\} \tag{F3}$$

Let us now consider the function $F_i^{(c)}$ obtained from (26) by replacing the $v_{j,l}$'s by 0, the $s_{j,l}$'s by $|s_{j,l}|$ and by changing the m_k's such that $(m_k/\Delta v - P_j) < -1$ in $(2\ \Delta v P_j - m_k)$ so that $(m_k/\Delta v - P_j)$ as always larger than 1. The variance of this function, when calculated for a fixed P, can still be expressed in terms of Riemann sums which converge when $M \to +\infty$. Therefore, this variance can be bounded from above independently of M and Δv. When $|P-P_j| < 1$, $F_i^{(c)}$ can be expanded in Taylor series about $P = P_j$: $F_i^{(c)} = \sum_{m=0}^{+\infty} f_m(P-P_j)^m/m!$ and, by construction, $\langle f_{m_1} f_{m_2} \rangle \geqslant 0$ and $\langle (f_m)^2 \rangle \geqslant \langle (\phi_m)^2 \rangle$. Then, when evaluating $\langle (F_i^{(c)})^2 \rangle$ at $P = P_j + (1-\delta)$ one finds

$$\begin{aligned} \langle (F_i^{(c)})^2 \rangle &\geqslant \sum_{m=0}^{+\infty} \{\langle (f_m)^2 \rangle\ [(1-\delta)^m/m!]^2\} \\ &\geqslant 2 \sum_{m=0}^{+\infty} \{\langle (\phi_m)^2 \rangle\ [(1-\delta)^m/m!]^2\} \\ &\geqslant \langle (\Phi_i^{(c)})^2 \rangle \end{aligned} \tag{F4}$$

Therefore, $\langle (\Phi_i^{(c)})^2 \rangle$ can be bounded from above independently of M and Δv. The same property can be proven for $\langle (\Phi_i^{(s)})^2 \rangle$ in a similar way. Therefore, $\langle (P - p/\Delta v)^2 \rangle$ can be made arbitrarily small by decreasing ε.

ACKNOWLEDGMENTS

One of the authors (D.B) was supported by an Allocataire Moniteur Normalien grant from the French Ministère de l'Education Nationale et de la Recherche. The authors acknowledge the hospitality of the Consorzio RFX while writing this paper.

REFERENCES

1. A. N. Kolomogorov, *Dok. Akad. Nauk. SSSR* **98**:527 (1954), English translation in *Stochatistic Behavior in Classical and Quantum Hamiltonian Systems*, G. Casati and J. Ford, eds. (Springer-Verlag, New York, 1979).

2. G. Benettin, L. Galgani, A. Giorgilli, and J. M. Strelcyn, *Nuovo Cimento* **79**:201 (1984).
3. N. N. Nekhoroshev, *Russian Math. Surveys* **32**:1 (1977). N. N. Nekhoroshev, *Trud. Sem. Petroviski* **5**:5 (1979).
4. G. Benettin and G. Galavotti, *J. Stat. Phys.* **44**:293 (1986).
5. P. Lochak, *Nonlinearity* **6**:885 (1993).
6. D. Bénisti and D. F. Escande, *Phys. Plasmas* **4**:1576 (1997).
7. D. Bénisti, Ph.D. thesis, Marseille (1995).
8. D. Bénisti and D. F. Escande, *Phys. Rev. Lett.* **80**:4871 (1998).
9. T. H. Dupree, *Phys. Fluids* **9**:1773 (1966).
10. J. R. Cary, D. F. Escande, and A. D. Verga, *Phys. Rev. Lett.* **65**:3132 (1990).
11. P. Helander and L. Kjellberg, *Phys. Plasmas* **1**:210 (1994).
12. J. P. Ramis, *Séries divergentes et théories asymptotiques* (Société Mathématique de France, 1992).
13. B. A. Fuchs, *Introduction to the Theory of Analytic functions of Several Complex Variables* (Providence, RI, 1963).
14. L. Verlet, *Phys. Rev.* **159**:98 (1967).
15. L. Casetti, *Physica Scripta*, **51**:29 (1995).
16. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, p. 70–71 (Wiley International, 1968).
17. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*, p. 472–475 (Cambridge University Press, 1989).