Fluctuation theorem for quasi-integrable systems

T. Goldfriend and J. Kurchan

Laboratoire de Physique Statistique, Département de physique de l’ENS, École Normale Supérieure, PSL Research University; Université Paris Diderot, Sorbonne Paris-Cité; Sorbonne Universités, UPMC Univ. Paris 06, CNRS - 24 rue Lhomond, 75005 Paris, France

received on 25 September 2018; accepted by Ch. Beck on 5 October 2018
published online 30 October 2018

PACS 05.70.Ln – Nonequilibrium and irreversible thermodynamics
PACS 05.20.-y – Classical statistical mechanics
PACS 05.45.-a – Nonlinear dynamics and chaos

Abstract – A fluctuation theorem (FT), both classical and quantum, describes the large-deviations in the approach to equilibrium of an isolated quasi-integrable system. Two characteristics make it unusual: i) it concerns the internal dynamics of an isolated system without external drive, and ii) unlike the usual FT, the system size, or the time, need not be small for the relation to be relevant, provided the system is close to integrability. As an example, in the Fermi-Pasta-Ulam chain, the relation gives information on the ratio of probability of death to resurrection of solitons. For a coarse-grained system the FT describes how the system “skis” down the (minus) entropy landscape: always descending but generically not along a gradient line.

Introduction. – Quasi-integrable systems —those whose Hamiltonian slightly differs from an integrable one— are widespread in Nature, ranging from planetary systems, to weakly nonlinear waves, to some quantum chains. When the systems are macroscopic they often involve coherent stable structures like solitons [1]. A small breaking of integrability leads to slow dynamics toward equilibrium [2]. The fluctuations and the nature of irreversibility along this slow route is the subject of this letter. The discussion is very similar for classical and quantum systems.

The trajectories of a classical integrable system with \( N \) degrees of freedom correspond to a laminar flow along a \( N \)-dimensional torus, defined by \( N \) integrals of motion \( \{J_r(x)\}_{r=1}^{N} \). According to the Kolmogorov-Arnold-Moser (KAM) theorem, as the integrability-breaking interactions are switched on, some tori remain unbroken, or change exponentially in time according to the Nekhoroshev theorem. However, these intermediate regimes are expected to exist for a range of coupling parameters that is vanishingly small —and is often irrelevant— for macroscopic systems, which thus perform unimpeded diffusion towards equilibrium. A description of weakly conserved quantities is given for, e.g., energy cascade in weakly coupled wave modes (weak turbulence) [3,4], or slow relaxation in a gas with long-range interactions [5].

A generic and somewhat surprising feature is that the Lyapunov time, measuring the separation of nearby trajectories, is in such systems often much shorter than the characteristic diffusion time of the quasi-constants of motion. Such is the case of the Solar System [6] (Lyapunov time of 5 Myrs \( \ll \) 5 Gyrs for the stability time) and the Fermi-Pasta-Ulam-(Tsingou) (FPU) nonlinear chain [2,7,8] (Lyapunov time of \( 10^6 \) compared to thermalization time of \( 10^{10} \) for a chain of size \( N = 1024 \)). A simple way to rationalize this is to consider the solvable problem of an integrable system perturbed by weak stochastic noise [9]: chaos develops first tangent to the action (invariant) variables, allowing the system to have ergodic motion primarily within a torus, and diffusion at a longer scale; there is furthermore no KAM regime for any amplitude of the (white) noise.

We consider here systems with any number of quasi-conserved quantities, \( J_1,\ldots,J_p \), corresponding to any model with a finite number of constants of motion which is weakly perturbed. In addition, the systems may have strictly conserved quantities, such as the total energy or linear momentum, imposed by the symmetries of the complete perturbed Hamiltonian. We indicate these by \( J_{c1},\ldots,J_{cn} \) and assume that they are independent —no function of them vanishes identically— and that no function of the nonconserved \( J_1,\ldots,J_p \) is strictly conserved.\(^1\)

We shall only distinguish these two sets when necessary, otherwise we drop the superscript.

\(^1\)We may thus add a conserved \( J_{c}^* \) to any nonconserved \( J_{b} \), but not a nonconserved \( J_{a} \) to a conserved \( J_{c}^{**} \).
The statistics of a system with many conserved quantities may in some cases be described by a Generalized Gibbs Ensemble (GGE),

$$\rho_{\text{GGE}}(x) = e^{-Q(x)/Z}; \quad Z = \int dx e^{-Q(x)} , \quad (1)$$

where we define

$$Q(x) = \sum \beta_r J_r(x) \quad (2)$$

and \{\beta_r\} are Lagrange multipliers. The GGE construction can be understood as a form of maximal entropy principle \cite{10}, or, as a consequence of equivalence of ensembles (to be discussed below) in analogy to the Gibbs measure; see ref. \cite{11}. Recently, the GGE distribution has proved useful in describing quantum integrable systems \cite{12} (and references therein), and analogous classical systems \cite{13}, where a direct access to the \beta is granted by the response and correlation function \cite{14}. In the case in which the \beta are not true constants, the \beta would be slowly time-dependent \cite{15,16}.

Fluctuation theorems refer to a group of relations concerning a system which evolves under nonequilibrium conditions \cite{17-25}. Mainly, one considers a system driven away from an equilibrium state by an external noise or nonconservative forces, and studies the heat/entropy exchange or work extraction. Here, we study a generalized exchange fluctuation theorem (GXFT) for quasi-integrable systems. We consider a system which starts at a GGE state and study the fluctuations in the quasi-conserved quantities, \Delta J_r, as induced by the weak breaking of conservation. In particular, we define the quantity

$$\Delta E \equiv \sum \beta_r \Delta J_r(x),$$

and prove that it obeys the following relation:

$$\ln \frac{P(u)}{P(-u)} = u, \quad (3)$$

where \(P(u)\) is the probability that the system assumes a value \(u\).

FT for this quantity in strictly (quantum) integrable systems, where work is performed by an external agent, has been considered recently in refs. \cite{26,27}. Here the situation is different, we are interested in the endogenous entropy production due to integrability breaking.

Before we prove eq. (3) let us first discuss a simple example presented by Jarzynski and Wójcik \cite{28}, in fact the simplest case of a GGE system. We consider two isolated systems at temperatures \(T_1\) and \(T_2\), as depicted in fig. 1. The simplest statistics of the combined system can be written as a GGE, \(\rho \propto e^{-\epsilon_1 E_1 - \epsilon_2 E_2}\), where \(E_1, 2\) are the energy of each subsystem, or equivalently as \(\rho \propto e^{-\epsilon_{\pm} E_{\pm} - \epsilon_{\mp} E_{\mp}}\) with \(\beta_{\pm} = \beta_1 \pm \beta_2, E_{\pm} = (E_1 \pm E_2)/2\), and \(\Delta E = \Delta E_{\pm} = -\Delta E_{\mp} \) for a given time. Since \(E_{\pm}\) is conserved we get \(u = \beta_{\pm} \Delta E_{\mp}\). According to eq. (3), we must have \(P(\Delta E_{\pm})/P(-\Delta E_{\pm}) = \exp(\beta_{\pm} \Delta E_{\pm})\). This is exactly the result obtained in ref. \cite{28}. The generalized exchange fluctuation theorem, eq. (3), that we address in the current letter may be viewed as a situation in which a quasi-integrable system behaves as a set of weakly coupled subsystems, each of which stands for a conserved quantity of the integrable system. The coupling is intrinsic, being the interaction induced by the integrability breaking.

In the simple example of fig. 1 the parameter that decides whether fluctuations and flow reversals may or may not be observable is the thinness of the channel, rather than the size of the system: this, in our general setting, will be translated into the magnitude of the breaking of conservation.

**Fluctuation theorem.** – We now prove the generalized exchange fluctuation theorem given in eq. (3).

**Classical:** Let us denote the phase-space variables by \(x = (q, p)\), and consider a Hamiltonian \(H = H_0 + H_p\), where \(H_0\) has conserved quantities \{\(J_r\)\}, and \(H_p \ll H_0\) is a perturbation. We assume that the initial probability distribution is of the GGE form (eq. (1)).

Then, the system evolves from \(x\) to \(x'\) with \(H\) for some time \(t\), and we measure the probability of the quantity \(q(t) = \sum \beta_r (J_r(x') - J_r(x))\). Note that the \(\beta\) are the same for initial and final states. The proof is given under the conditions of (a) time-reversibility of the dynamics, and (b) the functions \(J_r(x)\) are invariant under \((q, p) \rightarrow (q, -p)\). We use the additional following notations: the operation \((q, p) \rightarrow (q, -p)\) is designated by \(x \rightarrow \tilde{x}\), and \(T(x', t; x)\) denotes the probability of finding the system at state \(x'\) at time \(t\) given that it was at state \(x\) at time \(t = 0\).

The quantity in question can be written explicitly as

$$P(u) = \int dxd{x'} \delta \left( \sum \beta_r (J_r(x') - J_r(x)) - u \right) \times T(x', t; x) e^{-Q(x')} / Z. \quad (4)$$

Exchanging the integration variables \((x \rightarrow x')\) and using the delta function to re-express \(Q(x') \rightarrow Q(x) + u\), this
becomes
\[ P(u) = e^u \int dx d\tilde{x} \delta \left( \sum_{i} \beta_i (J_i(\tilde{x}) - J_i(x)) + u \right) \times T(x, t; \tilde{x}^t)e^{-Q(x)/Z}. \] (5)

Time-reversibility allows us to preform the transform \( T(x', t; x) \rightarrow T(\tilde{x}, t; x) \), which, together with the assumed invariance \( J(x) = J(\tilde{x}) \), gives the desired result \( P(u(t) = u) = e^u P(u(t) = -u) \).

Quantum: We now consider a system where the Hamiltonian almost commutes with a series of operators \( \hat{J}_r \).
Consider the following experiment \[29]:

- Diagonalize the operator \( \hat{Q} = \sum \beta_i \hat{J}_r \), such that \( \hat{Q}|q \rangle = q|q \rangle \);
- Choose an eigenvector \( |q \rangle \) with probability \( e^{-q} \), and evolve it with the complete Hamiltonian for time \( t \);
- Add the amplitudes \( |\langle q'|e^{-i\frac{t}{2}}H|q \rangle| \) to the histogram of probabilities of \( u = (q' - q) \), and repeat.

We are in fact calculating
\[ P(q' - q = u) = \int dq dq' \delta(q' - q - u)|\langle q'|e^{-i\frac{t}{2}H}|q \rangle|^2 e^{-q}/Z. \] (6)

Exchanging as above \( (q \leftrightarrow q') \), using the delta function, and the requirement of time-reversal, \( H = H^* \) and \( |q \rangle = |q^* \rangle \), we easily obtain the fluctuation relation (3).

The fluctuation theorem implies no assumption but as it stands is of little use, for the following reason: it requires a GGE initial condition even for large deviations, something that does not follow from any obvious physical process, and if it does, it may hardly be expected to be preserved by the dynamics. Here is where an argument of equivalence of ensembles becomes necessary: we wish to argue that an initial condition may be considered as if it were GGE, with some \( \beta_r \). As is well known in these cases, ensemble equivalence may be expected to hold for “coarse-grained” constants. One further assumes that these describe correctly the physical situation—a hydrodynamic limit, the applicability of which does not only depend on the system but also on its initial conditions. (The issue of equivalence of ensembles in the case of other fluctuation theorems has been discussed previously, e.g., refs. \[30–32\], and in the context of GGE in ref. \[11\].) Formally, we may work as follows: putting \( \beta_1 = \ldots = \beta_m = \beta_1 \), then \( \beta_{m+1} = \ldots = \beta_{2m} = \beta_2 \), etc., we have that \( Q = \beta_1 (J_1 + \ldots + J_m) + \beta_2 (J_{m+1} + \ldots + J_{2m}) + \ldots \). In other words, considering a situation with \( \beta_r \) grouped into sets of \( m \) automatically yields a coarse-graining on the constants \( J_r \), which enter as sums of \( m \) terms. The FT clearly works for such an initial condition, as this is only one particular situation. For large \( m \) we may expect to get equivalence of ensembles, the question that remains is whether this grouping we made accurately reflects the original problem, i.e., if a coarse-grained description is faithful. Note however that we are dealing with large deviations rather than simple averages, so the kind of equivalence of ensembles we need is very demanding.

Let us consider a more general distribution \( \rho(x) \propto e^{-Nf(J^q(x))} \), with \( f = O(1) \) and the \( J^q(x) \) coarse-grained variables over blocks of size \( m = aN \) (\( \alpha \) small but \( O(1) \)) and normalized by \( m \) to be of \( O(1) \). The dependence of the distribution \( f \) on \( J^q \) reflects the assumption of a hydrodynamic limit; it is justified by the fact mentioned above that a classical quasi-integrable system typically visits the approximate torus ergodically in a time much shorter than the diffusion time.

Then, one may try to see whether eq. (3) is valid with the specific choice for the values of \( \{ \beta_r \} \),
\[ \beta_r = \frac{\partial Ns(J^q)}{\partial J^q} - J_r, \] (7)
where we define the entropy \( Ns(J^q) = \int \delta(J^q(x) - J)dz \) and \( J_r \) is a saddle point of \( Nf(J^q) = Ns(J^q) \), in analogy with standard thermodynamics. We now discuss the limitations of this statement, a full mathematical derivation is given in the Supplementary Material Supplementary material.pdf (SM).

Looking at the derivation following eqs. (4), (5) it is clear that a fluctuation theorem of the form \( \ln P(N\Delta f = u) - \ln P(N\Delta f = -u) = u \) can be readily proved. In itself it is not very useful, as we do not have a direct access to \( f \). The usual FT theorem, for \( \beta_r \) defined by (7) would follow if we could identify \( N\Delta f \approx \sum \beta_r \Delta J^q_r \). Such an approximation depends both on the model and on the time-interval considered. If we assume a large deviation principle for the transition probability for \( J^q \rightarrow J^{q'} \), we find that \( u \) is dominated by certain values of \( J^q_r \). We have then to assume that these values are bounded to be small, but still very large with respect to their fluctuations of order \( 1/\sqrt{N} \). Note that this still leaves room for large deviations, because the \( J^q \) are intensive quantities. See SM.

An example: FPU chain. – We now demonstrate our result by treating a specific quasi-integrable system: the Fermi-Pasta-Ulam-(Tsingou) (FPU) chain (see ref. \[33\] for reading on the significant contribution of Tsingou). This is a 1D nonlinear chain whose Hamiltonian reads
\[ H_{\text{FPU}} = \frac{1}{2} \sum_{n=1}^{N} p_n^2 + \sum_{n=0}^{N} V_{\text{FPU}}(q_{n+1} - q_n), \] (8)
where \( V_{\text{FPU}}(r) = r^2/2 + Ar^3/3 + Br^4/4 \), and we take a fixed-ends boundary condition \( q_0 = q_{N+1} = 0 \). The dynamics depends on the size of the chain \( N \), the energy density \( \epsilon = E/N \), and the parameter \( B \); The parameter \( A \) can be rescaled by the energy \( \epsilon \), and thus it is set to \( A = 1 \) hereafter \[2\]. The FPU potential can be written as a small

10002-p3
perturbation of the Toda potential, $V_{\text{Toda}}(r) = V_0(e^{\lambda r} - 1 - \lambda r)$. For the values $V_0 = (2A)^{-2}$ and $\lambda = 2A$ one finds $\mathcal{H}_{\text{FPU}} = \mathcal{H}_{\text{Toda}} + \mathcal{H}_p$, with $\mathcal{H}_p \sim (2A^2/3 - B)e^4/4 + (A^3/3)e^{3/2}/5$. The Toda chain is integrable [34,35], and a set of conserved quantities $\{J_r\}$ can be derived, having the properties that: i) they are exact constants of the Toda lattice; ii) for weak coupling they are very close to the Fourier modes, except for a small fraction, which may be associated with soliton numbers. See SM.

The fact that the Toda dynamics serves as an underlying integrable model for the FPU chain has been established in three main aspects: 1) the presence of solitons in FPU dynamics [1, 2] at short timescales, although being chaotic, the FPU dynamics completely explores Toda tori [2,36], 3) starting from a concentrated ensemble the FPU dynamics drifts between quasi-stable states, each of which can be characterized by Toda tori and a corresponding GGE ensemble [8].

Figure 2 demonstrates the GXFT for a system of size $N = 15$, with the parameters $A = 1$, $B = 2/3$. The initial GGE ensemble is generated with a Monte Carlo sampling. The values of $\{\beta_r\}$ are chosen to follow a step function profile, $\beta_{15} = \beta_0$ and $\beta_{15} < 15 = \beta_0$, with $\beta_0 > 1$ such that only the lower part of the set is excited and $\langle e \rangle \approx 0.01$. In fig. 2 the probability distribution of $u$ is shown for different times. At long times, negative values of $u$ are rare, indicating a net drift of the conserved quantities as the system approaches toward equipartition. For comparison, we present the distribution of $u$ in the case of the Toda dynamics. Here, any departure from a delta function is only due to a random numerical error. The inset of fig. 2 demonstrates the validity of eq. (3).

Discussion. – The FT stresses the fact that a coarse-grained system always goes down an (minus)-entropy landscape, but not necessarily as a gradient, just as a ski descent.

Let us distinguish as above the conserved $J_1, \ldots, J_n$, and non conserved $J_1, \ldots, J_p$ quantities. The inverse temperatures for the two groups are $\beta_1, \ldots, \beta_n$ and $\beta_1, \ldots, \beta_p$, respectively. Then it is easy to see that in the course of evolution the Lagrange multipliers $\beta_1, \ldots, \beta_p$ will tend to finite values (they impose a constraint for the truly conserved quantities), while $\beta_1, \ldots, \beta_p$ all go to zero, because as $t \to \infty$ they are no longer imposing any constraint. For instance, returning to the simple set-up in fig. 1, we know that at long times $T_1 = T_2$, which implies $\beta_+ \to 2/T_1 = 2/T_2$ and $\beta_- \to 0$.

Furthermore, the fluctuation theorem immediately implies in the usual way that, for $u = \beta_\Delta J_r$, 
\begin{equation}
\langle u \rangle = \int_{-\infty}^{+\infty} P(u) u \, du = \int_{0}^{+\infty} P(u) [1 - e^{-u}] \, du \geq 0
\end{equation}
This is a form of the Second Principle, since for a small change $\delta J_r$ we have that 
\begin{equation}
\delta S(J) = \sum \frac{\partial S}{\partial J_r} \delta J_r = \sum \beta_r \delta J_r,
\end{equation}
with $S(J) = N s(J)$ being the entropy.
Finally, we reflect the ideas discussed above in the slow thermalization of a large size FPU chain with low energy. This example will be discussed in detail in a future publication [8]. In the modern version of the original FPU numerical experiment one starts with an initial ensemble in which only the lowest Fourier modes are excited and studies its dynamics [2]. Figure 3 shows the time evolution of the profile $\langle J_k \rangle$ for a system with $N = 511$ and $\epsilon = 10^{-3}$. The time to fill the Toda tori in such system is of order $10^3$, shorter than the typical time for changes in $J_k$ induced by the breaking of integrability. The hydrodynamic limit, which is evident in the self-averaging profiles, suggests that at any time along the dynamics, the system can be described by a coarse-grained GGE. Indeed, we have verified that this system admits microcanonical ensemble averages with $m = 23$ coarse-grained quantities [8].

Moreover, we see from the inset of fig. 3 that after some time, the vast majority of the out-of-equilibrium (linear) variation of quasi-constants of motion is in the lower modes, that we know correspond to the soliton modes (see SM). Thus, the evolution of the $\beta_\tau$ (and $\Delta Q$) at late times describes the gradual death of the excessive solitons, and the fluctuation theorem describes the ratio of death and resurrection of those. Here we may note the remark made in the abstract: we could have a macroscopic system with only a few solitons, and the desired fluctuations scale with the number of solitons and not with the size, so they may be observable.

***

We thank Abhishek Dhar, Leticia Cugliandolo and Jacopo de Nardis for fruitful discussions. TG and JK are supported by the Simons Foundation Grant No. 454943.

REFERENCES

[1] Zabusky N. J. and Kruskal M. D., Phys. Rev. Lett., 15 (1965) 240.
[2] Benettin G., Christodoulidi H. and Ponno A., J. Stat. Phys., 152 (2013) 195.
[3] Zakharov V. E. and Filonenko N. N., J. Appl. Mech. Tech. Phys., 8 (1967) 37.
[4] Düring G., Josserand C. and Rica S., Phys. Rev. Lett., 97 (2006) 025503.
[5] Chavanis P.-H., Physica A, 377 (2007) 469.
[6] Laskar J., Icarus, 196 (2008) 1.
[7] Benettin G., Pasquali S. and Ponno A., J. Stat. Phys., 171 (2018) 521.
[8] Goldfriend T. and Kurchan J., eprint arXiv: 1810.06121 (2018).
[9] Lam K.-D. N. T. and Kurchan J., J. Stat. Phys., 156 (2014) 619.
[10] Jaynes E. T., Phys. Rev., 106 (1957) 620.
[11] Yuzbashyan E. A., Ann. Phys., 367 (2016) 288.
[12] Vidmar L. and Rigol M., J. Stat. Mech. Theory Exp., 2016 (2016) 064007.
[13] Cugliandolo L. F., Lozano G. S., Picco M. and Tartaglia A., J. Stat. Mech. Theory Exp., 2018 (2018) 063206.
[14] Foini L., Gambassi A., Konik R. and Cugliandolo L. F., Phys. Rev. E, 95 (2017) 052116.
[15] Essler F. H. L., Keihein S., Manmana S. R. and Robinson N. J., Phys. Rev. B, 89 (2014) 165104.
[16] Stark M. and Kollar M., eprint arXiv:1308.1610 (2013).
[17] Evans D. J., Cohen E. G. D. and Morriss G. P., Phys. Rev. Lett., 71 (1993) 2401.
[18] Evans D. J. and Searels D. J., Phys. Rev. E, 50 (1994) 1645.
[19] Gallavotti G. and Cohen E. G. D., J. Stat. Phys., 80 (1995) 931.
[20] Kurchan J., J. Phys. A, 31 (1998) 3719.
[21] Lebowitz J. L. and Spohn H., J. Stat. Phys., 95 (1999) 333.
[22] Evans D. J. and Searels D. J., Adv. Phys., 51 (2002) 1529.
[23] Crooks G. E., Phys. Rev. E, 60 (1999) 2721.
[24] Jarzynski C., Phys. Rev. Lett., 78 (1997) 2690.
[25] Hickey J. M. and Genway S., Phys. Rev. E, 90 (2014) 022107.
[26] Mur-Petit J., Relaño A., Molina R. and Jaksch D., Nat. Commun., 9 (2018) 2006.
[27] Jarzynski C. and Wójcik D. K., Phys. Rev. Lett., 92 (2004) 230602.
[28] Kurchan J., eprint arXiv:cond-mat/0007360 (2000).
[29] Cleuren B., Van den Broeck C. and Kawai R., Phys. Rev. Lett., 96 (2006) 056001.
[30] Talkner P., Morillo M., Yi J. and Hänggi P., New J. Phys., 15 (2013) 095001.
[31] Jeon E., Kim Y. W. and Yi J., J. Phys. A: Math. Theor., 48 (2015) 305002.
[32] Dauxois T., Phys. Today, 61, issue No. 1 (2008) 55.
[33] Hénon M., Phys. Rev. B, 9 (1974) 1921.
[34] Flaschka H., Phys. Rev. B, 9 (1974) 1924.
[35] Benettin G., Livr R. and Ponno A., J. Stat. Phys., 135 (2009) 873.

More precisely at times which are longer than the time to fill the initial Toda tori.