Journal of Statistical Physics, Vol. 29, No. 1, 1982

Rates of Convergence to Equilibrium in the Prigogine–Misra–Courbage Theory of Irreversibility[1]

R. de la Llave[2]

Received December 11, 1981

The rates of convergence to equilibrium in the Prigogine–Misra–Courbage theory of irreversibility, as developed by Goldstein, Misra, and Courbage, are examined. It is found that arbitrarily slow convergence to equilibrium should be present; in fact, in a certain precise sense, it should be the most abundant behavior. This is compared with the common beliefs in kinetic theory.

KEY WORDS: Irreversibility; kinetic theory; K flow; nonunitary conjugation; Markov process.

1. INTRODUCTION

Misra, Prigogine, and Courbage have proposed a general theory of irreversibility in (classical and quantum) dynamical systems[(1–3)] based on ideas of Prigogine *et al.*[(4)] and Misra.[(5)] The theory has been further developed by these and others[(6–8)] and constitutes a conceptual framework in which to discuss irreversibility of dynamical systems. It has been shown that, in this framework, very unstable classical dynamical systems are indeed irreversible and that their convergence to equilibrium is monotonic.[(6)]

Since irreversible behavior in nature has many other features besides monotonic irreversibility (these features constitute the lore of kinetic theory), it seems advisable to study which of them can be fitted into Prigogine–Misra–Courbage conceptual framework. This paper is concerned with rates of approach to equilibrium in unstable classical dynamical systems.

In Section 2 a brief summary of the basis of the theory is given. In Section 3 it is shown how this program of establishing monotone approach

---

[1] Research supported in part by NSF Grant No. PHY 78-23952.

[2] Department of Mathematics, Princeton University, Princeton, New Jersey 08544.

0022-4715/82/0900-0017$03.00/0 © 1982 Plenum Publishing Corporation

to equilibrium can be carried out for classical K systems. This section follows Ref. 6 very closely. In Section 4 the results on rates of approach to equilibrium are presented.

2. BRIEF SUMMARY OF PRIGOGINE–MISRA–COURBAGE THEORY

If the dynamics of a physical system is very unstable, we cannot make measurements and preparations accurate enough to follow a definite trajectory over an appreciable interval of time. So, since the concept of individual trajectories will not have any operational meaning, we should try to substitute for the usual dynamical description of the system, based on trajectories, another description somehow incorporating our inability to make accurate predictions.

It is hoped that this new description of dynamics will be free from recurrence (Poincaré–Zermelo) and time reversal paradoxes since both of them depend crucially on the validity of the concept of individual trajectories. In this new description all states should approach equilibrium monotonically.

Let us make this more precise. Let $(\Omega, \mathfrak{B}, \mu)$ be a probability space (which should be thought of as the phase space of a classical system, so we will assume that it contains sets of any measure less than or equal to one). Assume, moreover, that we have defined in this space a measure preserving time evolution, i.e., a group of transformations

$$T^t : \Omega \to \Omega, \qquad t \in \mathbb{R}$$

satisfying

$$\text{(i)} \qquad T^{t+s} = T^t T^s, \qquad \text{all } t, s \in \mathbb{R}$$

$$\text{(ii)} \qquad \mu\left[(T^t)^{-1} A\right] = \mu(A), \qquad \text{all } t \in \mathbb{R}, \text{ all } A \in \mathfrak{B}$$

(This should be thought of as Hamiltonian evolution.)

It is well known (Koopman's lemma[9]) that the family of operators

$$U_t : L^2(\Omega) \to L^2(\Omega)$$

defined by

$$(U_t \phi)(x) = \phi(T^{-t} x)$$

(we will sometimes consider U_t as defined in other spaces of integrable functions by this same formula) is a semigroup of unitary operators that also satisfy

$$\text{(i)} \qquad U_t \phi \geqslant 0, \qquad \text{whenever} \quad \phi \geqslant 0$$

$$\text{(ii)} \qquad \int U_t \phi \, d\mu = \int \phi \, d\mu$$

$$\text{(iii)} \qquad U_t 1 = 1$$

The first two conditions are necessary and sufficient for the operator U_t to map probability densities into probability densities. The third one means that U_t leaves 1, the microcanonical equilibrium state, invariant. Operators satisfying the first two properties are called stochastic and if they also satisfy the third one they are called doubly stochastic. There are many doubly stochastic operators that do not come from measure-preserving transformations. An example to keep in mind could be the heat kernel.

In accordance with the philosophy previously exposed, we will assume that, in the process of measurement, we do not observe ρ, but only a blurred version of it, $\tilde{\rho}$, related to ρ by

$$\tilde{\rho} = \Lambda \rho$$

where Λ is a doubly stochastic operator implementing this smearing out of states. If Λ is injective, we can also define a dynamics in the $\tilde{\rho}$ description to model the evolution of the physical system, namely,

$$\tilde{\rho}_t = \Lambda U_t \Lambda^{-1} \tilde{\rho}_0$$

Notice that Λ^{-1} will not be, in general, bounded. We will require, however, that it is densely defined.

In order that this dynamics can be interpreted as an evolution of states leaving equilibrium invariant, it is necessary that $\Lambda U_t \Lambda^{-1}$ be a doubly stochastic operator in its domain of definition. From that it follows Ref. 10 that $\Lambda U_t \Lambda^{-1}$ maps L^P into L^P (any $1 \leqslant P \leqslant \infty$) and that $\|\Lambda U_t \Lambda^{-1}\|_{L^P \to L^P} = 1$. So, it has a unique extension W_t to the whole L^P which is also doubly stochastic. It is to be remarked that we only make these requirements for $t \geqslant 0$. It is known that if we made them for all $t \in \mathbb{R}$ we would run into difficulties. If $W_t(\tilde{\rho} - 1)$ is to tend to zero, then W_t cannot be uniformly bounded. In particular, it cannot be doubly stochastic and therefore cannot be interpreted as (backwards) evolution of states. This causes no difficulty at all because it only emphasizes the time asymmetry of W_t evolution.

The goal is now to find a physically motivated Λ such that the approach to equilibrium in the $\tilde{\rho}$ description is monotonic for all states. Following standard practice in the theory, we will consider only states described by L^2 functions:

$$E = \left\{ \rho \,|\, \rho \in L^2,\ \rho \geqslant 0,\ \int \rho \, d\mu = 1 \right\}$$

This is quite reasonable for finite systems with Liouville measure, but for infinite systems endowed with Gibbs measure it appears physically unnatural to exclude states described by singular measures. So, our discussion will be interesting only for the first case. We will also take as a definition of monotonic approach to equilibrium

$$\|\tilde{\rho}_t - 1\|_{L^2} \underset{\text{monotonically}}{\rightarrow} 0$$

Notice that this requirement shows that $\|\tilde{\rho}_t - 1\|_{L^2}$ is a candidate for the entropy of the system. Other such candidates have been considered in Ref. 11. Though this point of view about irreversibility (existence of an ever-increasing entropy function) is related to Prigogine–Misra–Courbage theory, it is independent and will, therefore, not be considered in detail here. Notice also that there is no hope that this program can be carried out for very stable systems, having, e.g., a state evolving in a periodic fashion. It is known that if the program outlined above can be carried through, the system should satisfy certain properties stronger than ergodicity.

In the next section we are going to show, following Ref. 6 closely, that indeed we can find such Λ for K systems.

3. MONOTONIC APPROACH TO EQUILIBRIUM OF K SYSTEMS

We say that an abstract dynamical system is a K flow if there exists a σ algebra $\mathfrak{F}$ of measurable sets satisfying

(a) $\mathfrak{F} \subset T_t\mathfrak{F}, \qquad t \geqslant 0$

(b) $\bigvee_{t=-\infty}^{\infty} T_t\mathfrak{F} = \mathfrak{B}$, the algebra of all measurable sets

(c) $\bigcap_{t=-\infty}^{\infty} T_t\mathfrak{F} =$ algebra of all sets of measure zero or one

The physical interpretation of this is that $\mathfrak{F}$ is the σ algebra of all the sets that can be specified by measurements in the past of the trajectory. (We measure with finite accuracy.) So what these conditions mean is that specifying the trajectory up to a remote past does not specify very much, (c), but to predict the trajectory to a distant future, we will need to get a great deal of information, (b).

See Refs. 12 and 13 for more details.

It is known that K systems appear when the underlying physical process is very unstable.[12]

We will call P_λ the projections over the space of functions measurable with respect to $T_\lambda\mathfrak{F}$

$$P_\lambda\rho = E[\rho \mid T_\lambda\mathfrak{F}]$$

From properties (a), (b), (c) follows immediately

(a′) $P_\lambda < P_{\lambda'}$ whenever $\lambda < \lambda'$

(b′) $\lim_{\lambda\to+\infty} P_\lambda = Id$

$\lim_{\lambda\to-\infty} P_\lambda =$ projection over constant functions $\equiv P_{-\infty}$

(Limits are understood in strong operator sense.)

(c′) $U_t^* P_\lambda U_t = P_{\lambda-t}$

(d′) $U_t^* P_{-\infty} U_t = P_{-\infty}$

Calling

$$F_\lambda = P_\lambda - P_{-\infty}$$

we can define

$$\Lambda = \int h(\lambda)\, dF_\lambda + P_{-\infty}$$

We have to show that for some choices of h, this Λ has all the required properties.

(a) It is clear from the definition that Λ preserves integrals, and $\Lambda 1 = 1$.

(b) Λ is positivity preserving if h is positive, decreasing, $h(-\infty) \leqslant 1$, as can be seen integrating by parts in the definition of and observing that under the previous hypothesis it is a combination, with positive coefficients, of positivity-preserving operators:

$$\Lambda = -\int h'(\lambda) P_\lambda\, d\lambda + \left[1 - h(-\infty)\right] P_{-\infty}$$

(c) If $h(\lambda) \neq 0$ then Λ^{-1} exists and is densely defined.

Now, we remark that W_t is positivity preserving if and only if $U_t^* W_t$ is. A simple computation shows that

$$U_t^* W_t = \int \frac{h(\lambda + t)}{h(\lambda)}\, dF_\lambda + P_{-\infty}$$

so that W_t is positivity preserving provided that $h(\lambda + t)/h(\lambda)$ is decreasing in λ. Notice that by (b), we already have

(i) $$0 \leqslant \frac{h(\lambda + t)}{h(\lambda)} \leqslant 1$$

(ii) $$\lim_{\lambda \to -\infty} \frac{h(\lambda + t)}{h(\lambda)} = 1$$

$$\|(W_t - P_{-\infty})\rho\|_2^2 = \|(U_t^* W_t - P_{-\infty})\|_2^2$$

$$= \int \frac{h^2(\lambda + t)}{h^2(\lambda)}\, d\langle \rho, F_\lambda \rho\rangle$$

Since $d\langle \rho, F_\lambda \rho\rangle$ is a finite, positive measure, and $h^2(\lambda + \mathrm{t})/h^2(\lambda)$ decreases monotonically with t for any fixed λ, we will have, using Lebesgue monotone convergence

$$\|(W_t - P_{-\infty})\rho\|_2^2 \to 0 \text{ monotonically}$$

provided

$$\lim_{t \to \infty} \frac{h^2(\lambda + t)}{h^2(\lambda)} = 0 \qquad \text{for all } \lambda$$

This will happen if and only if $\lim_{\lambda \to \infty} h(\lambda) = 0$.

So, we will have finished the construction of Λ if we can find a function $h > 0$ which simultaneously satisfies the following:

(a) monotone decreasing

(b) $h(-\infty) \leqslant 1$

(c) $\dfrac{h(\lambda + t)}{h(\lambda)}$ decreasing

(d) $\lim_{\lambda \to +\infty} h(\lambda) = 0$

These conditions are satisfied by a function of the form

$$h(\lambda) = e^{-\Phi(\lambda)}$$

with Φ satisfying

(a) Φ convex, increasing

(b) $\Phi > 0$

(c) $\lim_{\lambda \to \infty} \Phi(\lambda) = +\infty$

which are obviously compatible.

4. RATES OF CONVERGENCE TO EQUILIBRIUM

The first result we will prove is independent of the Prigogine–Misra–Courbage theory. It is a constraint on rates of approach to equilibrium imposed by the structure of dynamics. In this sense, it could be said that it is reminiscent in spirit of the Poincaré–Zermelo recurrence paradox.

Theorem 1. Suppose we have a dynamics W_t, $\| W_t \| \leqslant 1$ that makes the system approach equilibrium. If all states decay at a certain fixed rate, this rate can be chosen to be exponential.

Proof. Suppose we had

$$\|(W_t - P_{-\infty})\rho\| \leqslant K(\rho) f(t), \qquad f(t) \to 0 \text{ for all } \rho \in E.$$

Then we would have a similar inequality valid for all $\rho \in L^2$. Using the uniform boundness principle, we can get

$$\|(W_t - P_{-\infty})\| \leqslant K f(t)$$

Choose τ such that $Kf(\tau) = \alpha < 1$. Then

$$\|(W_{n\tau} - P_{-\infty})\rho\| < \alpha^n$$

Remember that, by hypothesis, $P_{-\infty} = \lim W_t$. Therefore $W_\beta P_{-\infty} = P_{-\infty}$. So,

$$\begin{aligned}\|(W_{n\tau+\beta} - P_{-\infty})\rho\| &= \| W_\beta (W_{n\tau} - P_{-\infty})\rho\| \\ &< \|(W_{n\tau} - P_{-\infty})\rho\|\end{aligned}$$

We see that we can obtain

$$\|(W_t - P_{-\infty})\rho\| < \frac{1}{\alpha} e^{(ln\alpha)t/\tau}$$

which establishes the claim. ■

From now on we will specialize the discussion to Prigogine–Misra–Courbage theory, though it will be clear from the argument that part of the result will have a more general validity. It is possible that the results in this section also hold for some other methods proposed to drive Hamiltonian dynamics to equilibrium, as, for example, modifying boundary conditions. This question is still open, however.

We will need the following version of the uniform boundness principle.

Theorem 2. Call $E_p = E \cap L^p$, $2 \leqslant p \leqslant \infty$.

Let $A_\alpha : L^p \to L^2$ be a family of bounded operators such that $\|A_\alpha\|$ is unbounded as $\alpha \in I$. Then, there is a residual (in the L^p-induced topology) set in E_p all whose elements x have the property that $\|A_\alpha x\|_2$ is unbounded.

What we mean by a residual set is one that contains a countable intersection of open dense sets. It is known (Baire category theorem) that residual sets in complete metric spaces (closed subsets of Banach spaces, e.g.) are dense.[14] Moreover, it is clear from the definition that countable intersections of them are still residual and, therefore, also dense. This last property is what makes it reasonable to interpret these sets as very big. They are so big that not only are they close to every point in space but two of them cannot miss each other and have to overlap in a set, which is also big. This is reminiscent of what happens with sets of full measure and, in fact, the analogy can be carried further (see Ref. 15). So in sets which do not have a natural measure, but have a natural topology satisfying the conditions of the Baire theorem, it is customary to say that a property is true for most of the points or that it is "generic" when it holds for all points in a residual set. Notice that being generic depends on the topology as much as being true almost everywhere depends on the measure. Sets residual in a topology can be negligible (complement of residual) in another one. If we have both a topology and a measure, the two concepts of big associated to them can have similar disagreements even if both the topology and the measure are natural and are closely related (there are examples in the unit interval with the usual topology and Lebesgue measure of residual sets of zero measure).

Proof. Call

$$B_r = \{x \mid x \in L^p, \|A_\alpha x\|_{L^2} \leqslant r \text{ all } \alpha\}, \qquad r \in \mathbb{R}^+$$

Then, from the definition, $B_r \cap E_p$ are closed (in E_p). So, the only thing we have to prove is that they have empty interior because, in that case, $E_p - (B_r \cap E_p)$ would be open and dense. But $\bigcap_{r \in N} E_p - (B_r \cap E_p)$ is precisely the set of points x in E_p for which $\|A_\alpha x\|_{L^2}$ is unbounded.

The way to prove this is by contradiction. We will show that if any one of them had nonempty interior in E_p, $\|A_\alpha\|$ would be bounded. So, suppose there existed r, $\epsilon > 0$, $x_0 \in E_p$ such that $x \in E_p$, $\|x - x_0\|_{L^p}$ implies $x \in B_r$. We can, then, find $\delta > 0$ such that $\|y\|_{L^p} < \delta$ implies $\|(1 + \int y)^{-1}(x_0 + y) - x_0\|_{L^p} < \epsilon$. If $y \geqslant 0$, $(1 + \int y)^{-1}(x_0 + y) \in E_p$. So, by the assumption, should also belong to B_r. Therefore, $x_0 + y \in B_{r(1+\delta)}$ and $y \in B_{r(2+\delta)}$. In other words, any positive function of L^p norm less than ϵ belongs to $B_{r(2+\delta)}$. Writing any function of L^p norm less than ϵ as the difference of its positive and negative parts, we see it should be contained in $B_{r(4+2\delta)}$ and that implies a uniform bound for $\|A_\alpha\|$. ■

We are going to use this theorem to prove the two following ones.

Theorem 3. Let E_p be defined as before. Then, given any $f(t) \to 0$ and any Λ constructed as in Section 3, there exists an L^p-residual set R_p in E_p such that $[1/f(t)]\|(W_t - P_{-\infty})\rho\|_{L^2}$ is unbounded for all $\rho \in R_p$.

Since we are going to prove existence of states satisfying certain properties it could be that extra conditions in the definition of states removed them. One such restriction that comes naturally in statistical mechanics of many-particle systems is the symmetry of ρ under exchange of arguments. However, the conclusion still holds.

Theorem 4 (many-particle case). We have

$$\Omega = (\Delta)^n$$

Let

$$E_p = \Big\{\rho \mid \rho \in L^p(\Omega),\ \rho \geqslant 0,\ \int \rho = 1,$$

$$\rho \text{ symmetric under permutation of its arguments}\Big\}$$

Then, given any $f(t) \to 0$ and any Λ constructed as in Section 3, preserving symmetry under permutation of arguments, there exists an L^p-residual set R_p in E_p such that $(1/f(t))\|(W_t - P_{-\infty})\rho\|_{L^2}$ is unbounded for all ρ in R_p.

Notice that the extra condition that Λ preserves the symmetry under permutations can be justified on the same physical grounds (should be interpretable as a mapping from states into states) which were used to impose the others.

Proof. In view of Theorem 2, to prove Theorem 3 it will suffice to prove

(*) $\quad \|W_t - P_{-\infty}\|_{L^\infty \to L^2} = 1$

Since then, it would follow that $[1/f(t)]\|W_t - P_{-\infty}\|_{L^p \to L^2}$ is unbounded.

Proof of (*). If a function v in L^2 satisfies

$$F_\tau v = v$$

then

$$\|(W_t - P_{-\infty})v\|_2^2 = \|U_t^*(W_t - P_{-\infty})F_\tau v\|_2^2$$

$$= \int \frac{h^2(\lambda + t)}{h^2(\lambda)}\, d\langle F_\tau v, F_\lambda F_\tau v\rangle$$

Since

$$F_\lambda F_\tau = \begin{cases} F_\lambda, & \lambda \leqslant \tau \\ F_\tau, & \lambda \geqslant \tau \end{cases}$$

this is

$$= \int^\tau \frac{h^2(\lambda + t)}{h^2(\lambda)}\, d\langle v, F_\lambda v\rangle$$

$$\geqslant \int^\tau \frac{h^2(t + \tau)}{h^2(\tau)}\, d\langle v, F_\lambda v\rangle$$

$$= \frac{h^2(t + \tau)}{h^2(\tau)} \int^\tau d\langle v, F_\lambda v\rangle$$

$$= \frac{h^2(t + \tau)}{h^2(\tau)} \int d\langle F_\tau v, F_\lambda F_\tau v\rangle$$

$$= \frac{h^2(t + \tau)}{h^2(\tau)} \|v\|_2^2$$

Now take v of the form

$$v = \chi_A - \mu(A) \qquad (\chi_A = \text{characteristic function of } A)$$

By choosing $\mu(A)$ sufficiently close to $1/2$, we can make $\|v\|_2$ arbitrarily close to $\|v\|_\infty$. But $\bigvee_t \mathfrak{F}_t$ contains sets of measure $1/2$. Since all the σ-algebras $\mathfrak{F}_t$ can be obtained from each other by application of a measure-preserving transformation, it follows that $\mathfrak{F}_\tau$ contains sets of measure arbitrarily close to $1/2$. So we have

$$\|(W_t - P_{-\infty})\|_{L^\infty \to L^2} \geqslant \frac{h(t + \tau)}{h(\tau)}$$

Since τ is arbitrary and $h(t+\tau)/h(\tau)$ tends to 1 when $\tau \to -\infty$ so we get

$$\| W_t - P_{-\infty} \|_{L^\infty \to L^2} \geqslant 1$$

That the left-hand side is also $\leqslant 1$ is trivial. So (*) is proved. ■
(The inequality $\geqslant 1$ would suffice for the proof.) This finishes the proof of Theorem 3. ■

The proof of Theorem 4 is quite similar to the proof of Theorem 3. We only keep at every stage the requirement that all the functions we consider are symmetric under the permutation of arguments. The only nontrivial part is to check that (*) goes through, but this can also be done just by observing that U_t should preserve this symmetry (it should be interpretable as a mapping of states into states) and that $\bigvee \mathcal{F}_t$ contains symmetric sets of arbitrary measure. Except for this, the argument is the same as the earlier one.

Remark 2. The choice of L^2 as a setting for Theorems 2, 3, and 4 is unessential and was done only because it is for L^2 norm that we prove monotone convergence.

It is possible, however, to prove the following statement:

For any $f(t)$ 0 and any $1 \leqslant q \leqslant p \leqslant -\infty$, there is an L^p-residual set R_p in E_p such that $[1/f(t)] \|(W_t - P_{-\infty})\rho\|_{L^q}$ is unbounded for all $\rho \in R_p$.

The reasoning to prove the result is the same. We need only that

$$\| W_t - P_{-\infty} \|_{L^\infty \to L^1} \geqslant K > 0$$

This can be proved by observing that, for all functions x in L^∞, we have

$$\|x\|_{L^2}^2 \leqslant \|x\|_{L^1} \|x\|_{L^\infty}$$

Therefore,

$$\begin{aligned} \|(W_t - P_{-\infty})\rho\|_{L^2}^2 &\leqslant \|(W_t - P_{-\infty})\rho\|_{L^1} \|(W_t - P_{-\infty})\rho\|_{L^\infty} \\ &\leqslant \|(W_t - P_{-\infty})\rho\|_{L^1} \leqslant 2\|\rho\|_{L^\infty} \end{aligned}$$

Using the fact that $\|(W_t - P_{-\infty})\|_{L^\infty \to L^2} = 1$, we obtain $\| W_t - P_{-\infty} \|_{L^\infty \to L^1} \geqslant 1/2$.

We can therefore say that, in some precise sense, slow convergence to equilibrium is the most abundant behavior. This, however, does not imply that there is no exponential convergence. We have the following:

Theorem 5. There is a set $D \subset E$ (E can be taken as in Theorem 3 or Theorem 4) such that

$$\|(W_t - P_{-\infty})\rho\| \leqslant K(\rho) e^{-\alpha t}, \qquad \alpha > 0$$

where α is a constant that only depends on the choice of h.

This set D is nontrivial in the sense that it is L^2-dense in the set

$$\{\rho \in E : \|\rho - 1\|_\infty \leqslant 1/3\}$$

(Stronger statements about nontriviality of this set could be made, but they are not worthwhile for present purposes.)

Proof. If $\rho \in L^2$, and $\rho \in \mathrm{Ran}(I - F_\tau)$ for some τ

$$\|(W_t - P_{-\infty})\rho\|_{L^2}^2 = \int_\tau \frac{h^2(\lambda + 1)}{h^2(\lambda)}\, d\langle \rho, F_\lambda \rho\rangle$$

$$\leqslant \frac{h^2(\tau + t)}{h^2(\tau)} \|\rho - P_{-\infty}\rho\|_{L^2}^2$$

Since h is log-concave there exist $\alpha > 0$, $K_\tau > 0$ such that $h(\tau + t)/h(\tau)B \leqslant K_\tau e^{-\alpha t}$ for all $t \geqslant 0$.

Now if ρ is a state and $\|\rho - 1\|_{L^\infty} \leqslant 1/3$, then $(I - F_\tau)\rho$ is also a state since $\|I - F_\tau\|_{L^\infty \to L^\infty} \leqslant 3$ and $(I - F_\tau)1 = 1$, so that $(I - F_\tau)\rho \geqslant 0$. (That the integral is 1, and symmetry also holds, is trivial.) But

$$(I - F_\tau)\rho \xrightarrow{L^2} \rho$$

So all points in the set $\{\rho \in E \,\|\rho - 1\|_{L^\infty} \leqslant 1/3\}$ can be arbitrarily approximated (in L^2) by states decaying to equilibrium with an exponential rate. That establishes the theorem. ■

It should also be noted that the theorem as stated is not always optimal. Any $f(t)$ such that

$$h(\tau + t) \leqslant K(\tau) f(t), \qquad \text{all } t \geqslant 0, \tau \leqslant 0$$

could be used instead of $e^{-\alpha t}$. There are some h satisfying all the conditions imposed before such that all $f(t)$ verifying this condition are exponentials or slower. However, for other admissible h (like $e^{-e\lambda}$) we can take f faster than any exponential.

We have the following theorem to prove that this is essentially optimal.

Theorem 6. Let $f(t)$ satisfy $f(t)/h(\tau + t) \xrightarrow[t \to +\infty]{} 0$ for all τ. Then, there is no L^2 function ρ decaying to equilibrium with rate $f(t)$, i.e., there exists no $\rho \in L^2$ satisfying

$$\rho \neq 1, \qquad \int \rho = 1$$

$$\|W_t \rho - 1\|_2 \leqslant K(\rho) f(t), \qquad t \geqslant 0$$

Proof. Since $\int d\langle\rho, F_\lambda\rho\rangle = \|\rho - 1\|_{L^2}^2$ there should be a τ such that $\int_{-\infty}^{\tau} d\langle\rho, F_\lambda\rho\rangle > 0$. But $d\langle\rho, F_\lambda\rho\rangle$ is a positive measure and we have

$$\int \frac{h^2(\lambda + t)}{h^2(\lambda)} d\langle\rho, F_\lambda\rho\rangle \geqslant \int_{-\infty}^{\tau} \frac{h^2(\lambda + t)}{h^2(\lambda)} d\langle\rho, F_\lambda\rho\rangle \geqslant \frac{h^2(\tau + t)}{h^2(\tau)} \int_{-\infty}^{\tau} d\langle\rho, F_\lambda\rho\rangle$$

This establishes the theorem.

Remark 2. The proofs of Theorems 3, 4, and 5 are rather robust in the sense that we can modify the hypothesis in several directions without altering the conclusions. For example, the proofs of Theorem 3 and 4 carry through under the extra restriction on states of having continuous densities. We need only to make some extra assumptions to guarantee that we can approximate (in the L^2 sense) characteristic functions by continuous ones which still have unit norm.

It turns out that regularity of the measure and normality of Ω are enough.[16] These assumptions are rather harmless and are definitely satisfied if Ω is a manifold and μ is Liouville measure. Probably other conditions like differentiability of density functions could alter the picture. But, at the moment, lacking a more concrete physical interpretation of Λ, it is unclear which direction these investigations should take.

Remark 3. The proofs given of Theorems 3 and 4 appear at first glance to be very nonconstructive. However, going carefully over the proof of the uniform boundedness theorem we can see that, to construct these slowly decaying states, what we are doing is to put mass very close to $-\infty$ (in the F_λ spectral decomposition).

If we consider conditions like $\int|\lambda|\, d\langle\rho, F_\lambda\rho\rangle < \infty$, we obtain uniform rates of convergence in the set of states verifying them.

$$\int \frac{h^2(\lambda + t)}{h^2(\lambda)}\, d\langle\rho, F_\lambda\rho\rangle \leqslant \sup_{\lambda\in\mathbb{R}} \left[\frac{1}{|\lambda| + 1} \frac{h^2(\lambda + t)}{h^2(\lambda)} \right] \int (|\lambda| + 1)\, d\langle\rho, F_\lambda\rho\rangle$$

so that ρ decays to equilibrium faster than

$$\varphi(t) = \sup_{\lambda\in\mathbb{R}} \frac{1}{(|\lambda| + 1)^{1/2}} \frac{h(\lambda + t)}{h(\lambda)}$$

Using known information about h, it is not difficult to prove $\varphi(t) \approx t^{-1/2}$.

It should be clear how to obtain similar results for other growth conditions $\int g(\lambda)\, d\langle\rho, F_\lambda\rho\rangle < \infty$.

It is interesting that this condition has a physical interpretation in other slightly different versions of Prigogine–Misra–Courbage theory.[1] Namely, it is the requirement that the "age" (expected value of a time operator) of the state be finite.

Remark 4. Notice also that in passing from Theorem 3 to Theorem 4, the only property of symmetry under permutation of arguments that was essentially used is that it be an L^{∞}-closed condition. Other such conditions like spatial homogeneity can also be imposed without changing the conclusion. (The argument about existence of characteristic functions satisfying such symmetries is the same as the one given.) So we should expect slow convergence to take place, even starting with spatially homogeneous states, whenever Λ and U_t preserve this symmetry. (The contrary would be difficult to reconcile with the physical interpretation of smearing out.)

Remark 5. Notice that the construction of states in Theorem 3 and 4 has been done under the hypothesis that the construction of Λ is done in the same way as in Ref. 6. It seems that the general conceptual framework could admit other constructions and it might be that these new constructions altered the conclusions. The latter seems unlikely, but it is a question that needs further investigation.

Remark 6. It could also be thought that the states with the behavior exhibited in Theorems 3 and 4 might have been introduced when taking the closure of $\Lambda U_t \Lambda^{-1}$ or equivalently, that they could be suppressed by imposing the extra condition $\rho \in \operatorname{Ran} \Lambda$ for ρ to belong to E. This is not so, because the proofs of Theorems 3 and 4 can be reproduced verbatim for the operators

$$V_t = \Lambda U_t = \int h(\lambda + t)\, dP_\lambda + P_{-\infty}$$

It is clear that

$$V_t \tilde{\rho} = W_t \Lambda \tilde{\rho}$$

so that if $V_t \tilde{\rho}$ goes to equilibrium slowly, $W_t \Lambda \tilde{\rho}$ does also. Clearly $\Lambda \tilde{\rho} \in \operatorname{Ran} \Lambda$.

Remark 7. This prediction of "extra-long tails" may be related to the existence of dissipative structures. It is conceivable that almost any system could accommodate dissipative structures of arbitrarily long life. However, this suggestion has the difficulty that dissipative structures are generally thought of as associated with large distances from equilibrium while extra-long tails behavior takes place on dense sets and so, arbitrarily close to equilibrium.

Remark 8. The fact that the constant α in Theorem 5 depends only on the choice made in the construction of Λ and that, on the other hand, could possibly be calculated from kinetic theory shows even more clearly that a deeper understanding of the relationship of this theory to kinetic equations is needed. Progress in this question and on other crucial issues like the dependence of these sets or the number of particles seems to depend on a more detailed study of concrete realizations of Λ.

5. COMPARISON WITH KINETIC THEORY

Comparison of these results with those of kinetic theory is not quite straightforward. First of all, the formalism is different; Prigogine–Misra–Courbage theory is formulated in terms of distribution functions and aims to describe the approach to microcanonical equilibrium while kinetic theory is usually expressed in terms of reduced distribution functions and correlations and is supposed to describe the approach to canonical equilibrium. Even the equivalence of these two ensembles is a nontrivial problem from the rigorous point of view.[17] However, in this section, we will not worry about this. We are going to assume all the standard beliefs used in kinetic theory, so that the result of this section will be a constraint on how many of them it can be hoped will be justified by Prigogine–Misra–Courbage theory in the face of previous results.

If we consider our system as immersed in a much larger heat bath with which it is interacting weakly, the fact that the whole system (system plus bath) is tending to microcanonical equilibrium implies that the small system tends to a canonical one. If we assume that, at the beginning, the heat bath is in equilibrium, we can assume it remains so exactly. Therefore, the difference of the state of the whole system with microcanonical equilibrium is the difference of the state of the small system with the canonical one.

Now we assume that after a sufficiently long time "the description is contracted" and that we can describe the system not by a many-body density but by a one-particle density. The evolution equation that this one-particle density satisfies is called the kinetic equation. Several variations of this basic procedure lead to different ones: Boltzman, Landau, Vlasov, etc. It is believed, moreover, on physical grounds (existence of relaxation times) that this evolution sometimes leads to an exponential convergence to equilibrium. (See Ref. 18, Chapter 13 for more details.) This would imply that the whole system tends to equilibrium exponentially, which is in contradiction with the previous results.

Of course, this conventional wisdom is open to criticism. It could be suggested that the fluctuations are just transferred from the system to the bath so that the whole system tends to equilibrium slower than the small system. This explanation seems to go against the spirit of usual kinetic theory and notice that the argument for the heat bath remaining in equilibrium is stronger here since we have a built-in tendency to it.

The equivalence between the canonical and microcanonical examples is only rigorous in the thermodynamic limit, and it is argued by some authors that this taking of the thermodynamic limit (or modification thereof) is enough to derive kinetic equations without any reinterpretation of evolution. (See Ref. 19 for a review of progress in that direction.)

A proper discussion of these points is beyond the scope of this paper. The objective of this last section was to try to clarify how much of the folk wisdom of kinetic theory can be justified at the same time on the basis of Prigogine–Misra–Courbage theory alone.

ACKNOWLEDGMENTS

I want to express my gratitude to Professor Arthur Wightman for his encouragement and advice throughout the preparation of this paper, and to Sheldon Goldstein for very useful suggestions and comments. I also want to acknowledge support from the University of Texas and the Institute of International Education which enabled me to attend the "Workshop on Long-Time Prediction in Non-Linear Conservative Dynamical Systems," where I could have a first-hand account of P–M–C theory. I also want to thank Mrs. Anderson for her excellent transcription.

REFERENCES

1. B. Misra, I. Prigogine, and M. Courbage, *Physica* **98A**:1–26 (1979).
2. B. Misra, I. Prigogine, and M. Courbage, *Proc. Natl. Acad. Sci. USA* **76**:3607–3611 (1979).
3. B. Misra, I. Prigogine, and M. Courbage, *Proc. Natl. Acad. Sci. USA* **76**:4768–4772 (1979).
4. I. Prigogine, C. George, F. Henin, and L. Rosenfeld, *Chem. Scr.* **4**:5–32 (1973).
5. B. Misra, *Proc. Natl. Acad. Sci. USA* **75**:1627–1631 (1978).
6. S. Goldstein, B. Misra, and M. Courbage, *J. Stat. Phys.* **25**:111–126 (1981).
7. B. Misra and I. Prigogine, *Suppl. Prog. Theor. Phys.* **69**:101–110 (1980).
8. B. Misra and I. Prigogine, in Proceedings of Workshop in Long-Time Prediction in Non-Linear Conservative Systems, 1981 (to appear).
9. P. R. Halmos, *Ergodic Theory*, Chelsea, New York (1955).
10. J. R. Brown, *Ergodic Theory and Topological Dynamics*, Academic Press, New York (1976).
11. S. Goldstein and O. Penrose, *J. Stat. Phys.* **24**:325–343 (1981).
12. V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics*, Benjamin, New York (1968).
13. M. Smorodinsky, *Lecture Notes in Mathematics No. 214*, Springer, New York (1971).
14. M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Academic Press, New York (1972).
15. J. C. Oxtoby, *Measure and Category*, Springer New York (1971).
16. W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York (1970).
17. O. Lanford III, in *Lecture Notes in Physics No. 20*, Springer, New York (1973).
18. R. Balescu, *Equilibrium and Non-Equilibrium Statistical Mechanics*, Wiley, New York (1975).
19. H. Spohn, *Rev. Mod. Phys.* **52**:569–615 (1980).