International Journal of Theoretical Physics, Vol. 37, No. 4, 1998

Dynamics, Thermodynamics, and Time Asymmetry

Mario A. Castagnino[1] and Edgard Gunzig[2]

Received June 13, 1997

There are two schools, or lines, of thought which attempt to unify the apparently divergent laws of dynamics and thermodynamics and to explain the observed time asymmetry of the universe and most of its subsystems in spite of the fact that these systems are driven by time-symmetric evolution equations. They will be called the coarse-graining and the extended dynamics schools (even if these names only partially describe their philosophy). The coarse-graining school obtains time asymmetry via a projection of the state space onto a space of "relevant" states. The corresponding projection of the primitive reversible evolution laws yields effective irreversible evolution laws for the relevant states. Extended dynamics always uses the same primitive reversible evolution laws. But these laws (in adequate extensions of the usual spaces where they are formulated) have a set of solutions S that can be decomposed into two subsets S_+ and S_- of time-asymmetric solutions. Time asymmetry is established by choosing one of these two sets as the arena in which to formulate the theory. This paper explains in the simplest self-contained and unbiased way the main characteristics of both schools and points out the advantages and disadvantages of each, in such a way as to make explicit the debate between the schools. Some cosmological features of the theory are also considered, mainly the problem of the low-entropy initial state of the universe

1. INTRODUCTION

In this paper we will study and try to solve two long-standing problems of theoretical physics.

1.1. The Problem of Time Asymmetry

The puzzle of the existence of the arrow of time or, in other words, the time asymmetry of the universe can be stated by asking two questions: (*i*) *How can the universe be time-asymmetric if all the relevant physical laws*

[1] Instituto de Astronomía y Física del Espacio, 1428 Buenos Aires, Argentina.
[2] Instituts Internationaux de Physique et de Chimie, 1050 Bruxelles, Belgium.

0020-7748/98/0400-133 3$15.00/0

are time symmetric? (*ii*) *Why do all arrows of time point in the same direction?* In fact, the universe has several time asymmetries, corresponding to the various arrows of time (thermodynamic, electromagnetic, psychological, *etc.*), whereas the main laws of nature are time-symmetric; as usual, we will neglect the weak interaction laws, since it is very difficult to imagine a mechanism whereby they could explain the time asymmetry of the universe (Sachs, 1987). In this paper we would like to answer these questions by taking an appropriate mathematical formalism of the problem and using several, both old and new, well-established ideas (Tolman, 1987; Landau and Lifshitz, 1958; Davies, 1994). Before doing so, however, we must precisely define two important words: *conventional* and *substantial* (Sachs, 1987; Mackey, 1989). In mathematics we are used to working with *identical* objects, such as points, the two directions of an axis, the two semicones of a null cone, etc. In physics there are also identical objects: such as identical particles, spin directions, *etc.* Among identical objects there is always a mathematical transformation that exchanges these objects, but leaves the system unmodified. If we are forced to refer to identical objects independently, we will say that we are establishing a *conventional difference* between them, for example, when we call + and − the two directions of an axis, or "past" and "future" the two semicones of a null cone. If objects are different, we will say that there is a *substantial difference* between them. The problem of the arrow of time is that, in usual physical theories, past and future are usually only conventionally different, whereas we have the impression that they are in fact substantially different; events in the past have already happened, future events are yet to happen. In theories endowed with time-symmetric evolution equations, such as those with which we will be dealing, it is quite impossible to find time substantial asymmetry using rigorous mathematical manipulations. Normally, however, one can find within these theories, as we shall see, two identical mathematical structures, one related with the past and one with the future (*e.g.*, two subspaces of the space of solutions of the theory). Nevertheless, if we study a closed system as the universe, these structures are only conventionally different, because they are related to each other by a time inversion. Within these structures, however, the past is substantially different from the future. The particular choice of structure is physically immaterial, since time inversion exchanges them, leaving the universe unchanged. Therefore to create an arrow of time we just conventionally choose one of the structures. This choice is irrelevant, in the same way that it is irrelevant to choose one face of a die if all faces are marked with the same number. But when we have chosen one of the structures, a substantial difference is also created between past and future within it, and an arrow of time will have been established. This is the method we will use to create all the arrows of time, both in the coarse and the extended dynamics cases ("extended dynamics" is

also known as "intrinsic irreversibility," "restricted dynamics," or even "fine-graining") (see Section 6). To show that all the arrows of time point in the same direction, we will consider the master arrow of time to be the cosmological one. We will show that the expansion of the universe creates a thermodynamic instability within it, in such a way as to make the thermodynamic arrow of time necessarily point in the same direction as the cosmological one. We will refer to the literature for the problem of the coincidence of the other arrows of time with the cosmological master arrow.

1.2. The Problem of the Unification of Dynamics and Thermodynamics

A specific, but very important, example of the problem of Section 1.1 is the problem of the unification of the time-symmetric dynamical laws with the time-asymmetric thermodynamic laws. In fact, it is reasonable to expect that thermodynamic laws could be demonstrated by using either classical or quantum dynamical laws. However, it seems that this is not possible for the second law of thermodynamics, which says that entropy increases, in irreversible evolutions, leading the system to a state of thermodynamic equilibrium or maximal entropy. This problem can be stated as follows:

(i) Liouville's equation is the time-symmetric evolution equation for classical distribution functions or quantum density matrices ρ. (ii) This equation prevents the definition of any functional of ρ, $F(\rho)$ (if it is constructed only with ρ and mathematical elements of the Liouville-phase space), such that $F^{\cdot}(\rho) > 0$ or, in other words, it is impossible, as a consequence of Liouville's theorem, to define a Lyapunov variable, *i.e.*, a growing functional of ρ; for example, the volume or the support of a characteristic distribution function ρ is time constant, Gibbs and conditional entropies are time constant (Lasota and Mackey, 1985), and so on. (iii) Nevertheless, we actually see that the evolution leads the system to a thermodynamic equilibrium with a maximal-entropy, stationary state ρ^*.

Therefore the problem is how to combine Liouville's theorem with the obvious fact that everyday physical systems have a tendency to go to thermodynamic equilibrium. The solution of this problem is based on a theorem by Mackey and Lasota (1985) (Theorem 4.3.1 below):

Theorem. Let $S(t)$ be an ergodic transformation, with stationary equilibrium density ρ_* [of the associated Frobenius–Perron operator $P(t)$ in a phase space of finite ρ_*-measure]. Then, if $S(t)$ is ρ_*-mixing if and only if $P(t)\rho$ is weakly convergent to ρ_*, *i.e.*,

$$\lim_{t\to\infty}(P(t)\rho|g) = (\rho_*|g) \tag{1.2.1}$$

for all bounded measurable functions g, *i.e.*, if the time evolution in phase

space is $S(t)$, and the corresponding time evolution of the distribution functions is $\rho(t) = P(t)\rho(0)$, and this evolution is mixing, a chaotic property of evolutions that we will define below, and if there is an equilibrium density such that $P(t)\rho_* = \rho_*$, then (1.2.1) can be proved. However,

$$\lim_{t\to\infty} P(t)\rho \neq \rho_* \tag{1.2.2}$$

and in fact, as we will see, in many cases this limit does not even exist. Therefore we have a weak limit, but we have no strong limit (*i.e.*, a limit in the norm). Nevertheless we never see or measure ρ. What we see and measure are the mean values of physical quantities O such that

$$\langle O\rangle_\rho = (\rho|O) \tag{1.2.3}$$

Thus, what we actually see is that

$$\lim_{t\to\infty}\langle O\rangle_\rho = \langle O\rangle_{\rho*} \tag{1.2.4}$$

In fact, all the mean values of physical quantities go to their equilibrium values if the evolution of the system is ρ_*-mixing. So the solution of the problem is quite simple: (i) Liouville's theorem is embodied in (1.2.2): the system does not go (strongly) toward the equilibrium state. (ii) Tendency toward equilibrium is embodied in (1.2.4): the mean values of all the physical quantities goes to their equilibrium values. Clearly these facts are not contradictory. We will refer to this solution as the *nongraining* solution. As chaotic-mixing systems are very common in the universe, the problem is essentially solved. What remains to be studied are the different techniques to deal with the detailed calculations. These techniques attempt to find some logical modification of the theory in order to solve for the missing limit (1.2.2), which, even if unnecessary from the mathematical standpoint, is the way in which physicists are used to (or rather, very fond of) thinking, at least up to now. In fact there are two techniques, as follows.

1.2.1 Coarse-Graining

Let us define an arbitrary, but time-independent, projector:

$$P = |g)(g|, \qquad (g|g) = 1 \tag{1.2.1.1}$$

and let us define a coarse-graining density function as

$$\tilde{\rho} = P\rho = |g)(g|\rho) \tag{1.2.1.2}$$

From equation (1.2.1) we have

$$\lim_{t\to\infty} |g)(g|P(t)\rho) = |g)(g|\rho_*) \tag{1.2.1.3}$$

and therefore

$$\lim_{t\to\infty} \rho(\tilde{t}) = \tilde{\rho}_* \tag{1.2.1.4}$$

which is the coarse-graining version of equation (1.2.2) and the main equation of the first technique [of course, the same thing happens with the general projector $\Pi = \Sigma|g_i)(g_i|$, $(g_i|g_i) = \delta_{ij}$]. It is easy to demonstrate that (1.2.1.4) is a limit in norm. It is also evident that equation (1.2.1.4) may be obtained with a quite arbitrary state $|g)$, and that all the philosophy typical of the coarse-graining technique, namely the definition and consideration of macroscopic and microscopic states Misra *et al.*, 1979), is simply an intuitive justification to give physical meaning to the limit (1.2.1.4). But since this justification is really unnecessary [because the relevant and important limit is (1.2.4)], the physical explanation of the entire philosophy of the coarse-graining technique could be criticized on philosophical grounds Prigogine, 1980). This is the main problem with coarse-graining. It is an arbitrary method. It works perfectly well, but it is difficult to justify with physical-philosophical (metaphysical?) arguments. In fact, coarse-graining contains the misleading statement: *we cannot see microscopic states* (*i.e.*, ρ), *but we can see macroscopic states* (*i.e.*, $\tilde{\rho}$). This statement leads to the problem of finding a unique reasonable definition for these macrostates. This problem is unresolved and, in our opinion, it will remain so, since $|g)$ is essentially arbitrary. In addition, if we arbitrarily choose some definition of macrostates, we are introducing a physical element which is really alien to the system itself; this definition, therefore, even if natural in particular examples, will be suspicious from a general point of view. The correct "no-graining" statement is: *we cannot directly measure microscopic states* (*i.e.*, ρ), *we can only measure mean values of physical quantities or observables* [among them the projector $P = |g)(g|$, and therefore the arbitrarily defined macroscopic states]. This statement is completely true at the classical and quantum levels Ballentine, 1990) and refers to *all physical observables.* Then we can rigorously say that, for example, the two thermodynamic variables $\langle p\rangle$ and $\langle v\rangle$ define the thermodynamic macrostate of a perfect gas, and so on.

1.2.2. Extended Dynamics

Let $\mathcal{L}$ be the Hilbert–Liouville space of the physical states ρ, and let $\mathcal{L}^x = \mathcal{L}$ be the space of the (anti) linear functionals on $\mathcal{L}$. We may think that not all $O \in \mathcal{L} = \mathcal{L}^x$ are physical admissible observables. In fact, observables are measured by real physical devices that very likely are free of sophisticated mathematical behavior, *e.g.*, are related with continuous and

derivable functions and not with discontinuous, nondifferentiable functions, even if square-integrable. So it is reasonable to suppose that $O(x)$ should be, *e.g.*, a Schwarz function (we will make this point precise in Section 5). So, calling the space of physically admissible observables $\mathbf{\Phi}$, we have that

$$\mathbf{\Phi} \subset \mathscr{L} = \mathscr{L}^{\times} \tag{1.2.2.1}$$

If we consider the dual $\mathbf{\Phi}^{\times}$ of $\mathbf{\Phi}$ we have a Gel'fand triplet (cf. Section 4.6):

$$\mathbf{\Phi} \subset \mathscr{L} = \mathscr{L}^{\times} \subset \mathbf{\Phi}^{\times} \tag{1.2.2.2}$$

(as we will see in Section 5, if we give to the functions of Φ some analyticity properties, we can also consider the time-asymmetry problem within this framework). We will work with states that belong to $\Phi^{\times}$: for example, ρ_* normally belongs to this space. As is well known, we define the functional A, the sum of the functionals B and C, $A = B + C$, as the functional $A[g] = (A|g)$ defined by

$$A[g] = (A|g) = B[g] + C[g] = (B|g) + (C|g)$$

for all $g \in \Phi$. The same method is used to define the product of a functional by a number. Analogously, if we have a sequence of functionals $A_1, A_2, \ldots$ the limit $A = \lim_{t\to\infty} A_i$ is defined as the functional such that

$$A[g] = (A|g) = \lim_{t\to\infty} A_i[g] = \lim_{t\to\infty} (A_i|g)$$

for all $g \in \Phi$. Then, as ρ and ρ_* can be considered as functionals on Φ, equation (1.2.1) reads

$$\lim_{t\to\infty} P(t)\rho = \rho_* \tag{1.2.2.3}$$

and we have found a rigorous "strong" limit corresponding to equation (1.2.1) (Gel'fand and Shilov, 1968). Perhaps the main problem with the extended dynamics technique is that it is usual to consider the states of $\Phi^{\times}\backslash\mathscr{L}$ as unphysical states, or just effective states, where some characteristics of real physical states have been neglected (such as Zeno and Khalfin effects). Nevertheless we can also say that every state which can be used to measure the mean values of all observables of Φ is essentially a physical state, and this is the case with all the states of $\Phi^{\times}$. However, this point is not completely clear. So neither technique is completely without fault. Nevertheless, since the real physical problem is solved by the Mackey and Lasota theorem, we could say that the sins involved are merely venial ones. On the other hand, each technique has some advantages; (i) Coarse graining works just with one physical space, $\mathscr{L}$. Also, coarse-graining is unavoidable to calculate global thermodynamic variables such as temperature or pressure. But (ii) the time

evolution of $\rho(t)$ can be computed more easily using the extended dynamics technique, since we have the vectors of space Φ^x that may be used to find a new spectral expansion for the observables in the problem. Once we know $\rho(t)$, we can compute $\tilde{\rho} = P\rho(t)$, while the direct computation of $\tilde{\rho}(t)$ using coarse-graining techniques directly can be more difficult (Hu *et al.*, 1992a). Since the coarse-graining technique is well known (Gel'fand and Shilov, 1968; Hu *et al.*, 1992a; Zwanzig, 1961), this paper will be mostly devoted to studying the new ideas introduce by no-graining and extended dynamics. The arrangement of this paper is as follows: In Section 2 we describe the dynamics, both classical and quantum, and define the notions of time symmetry and reversibility. This section is based on Castagnino *et al.* (1996). In Section 3 we deal with thermodynamics, and give the different definitions of entropy. This and the two following sections are based on Mackey (1989) and Lasota and Mackey (1985), but we have added the new mathematical and physical structures which have been studied and which have recently appeared. In Section 4 we introduce the classical evolution equations and we study the ergodic, mixing, and, exact transformations. In Section 5 we find the quantum evolution equations. We study the no-graining and extended dynamics ideas in models with both discrete and continuous spectra, and we consider the Friedrichs model, for pure and mixed states. In Section 6 we study the coarse-graining and the extended dynamics projectors. We also study the problem of time asymmetry. In Section 7 we review the main equation of thermodynamics in curved space-time. In Section 8 we consider the alignments of the arrows of time. This section is mostly based on Tolman (1987), Davies (1994), and Castagnino *et al.* (1995). In Section 9 we draw our conclusions.

2. DYNAMICS

In this section we shall review the formalism that we will use throughout this work and we will see how the notions of reversibility and time asymmetry are introduced.

2.1. Classical Formalism

A classical system with N degrees of freedom is characterized by its Hamiltonian

$$H = H(x) = H(q_i, p_i) \tag{2.1.1}$$

which is a function of x, the generic point in the $2N$-dimensional phase space X, or a function of the configuration variables q_i and the momentum variables p_i ($i = 1, \ldots, N$). The system is solved if we compute the functions

$$q_i = q_i(t)$$
$$p_i = p_i(t)$$

or

$$x = x(t) \tag{2.1.2}$$

as solutions of Hamilton's equations

$$\frac{dq_i}{dt} = \partial_{p_i} H$$
$$\frac{dp_i}{dt} = -\partial_{q_i} H \tag{2.1.3}$$

satisfying, at time $t = 0$, the conditions

$$q_i(0) = q_i^0$$
$$p_i(0) = p_i^0 \tag{2.1.4}$$

The solution of the system of differential equations (2.1.3) is the map $S(t)$: $X \to X$, defined by

$$S(t)[x(0)] = x(t) \tag{2.1.4'}$$

We also call $S_t = S(t)$, and these S_t form a group. If $A \subset X$ is a subset of the phase space, we can compute the image of A, namely $S_t(A) = B$. Then, if μ_L is the Lebesgue measure on X, we can formulate the Liouville theorem:

Theorem 2.1.1. If $S(t)$ is the map obtained solving the classical dynamical evolution and A is a μ_L-measurable set of X, then

$$\mu_L(S(t)A) = \mu_L(A) \tag{2.1.4''}$$

i.e., classically the evolution preserves the "volume" in phase space.

Let us now define the notion of reversibility. Experimentally it is impossible to change the direction of time. The best we can do in order to simulate a time inversion is to film the motion under study and project the film backward. Then, if $q_i = q_i(t)$ and $p_i = p_i(t)$ give the real motion, the law of the fictitious motion obtained by playing the film backward will be $q_i = q_i(-t)$, $p_i = -p_i(-t)$, where to change t by $-t$ is just an easy way to avoid having to define new initial data (the final ones of the reversed motion). We can deduce that the time-reversal operator T acts on the configuration variables and the momentum variables as (Sachs, 1987; Messiah, 1962)

$$T(q_i, p_i) = (Tq_i, Tp_i) = (q_i, -p_i) \tag{2.1.5}$$

We can now consider the data (2.1.4) (which we have called "conditions at

zero time" and not "initial conditions" in order to avoid any reference to time, even though we will follow the common usage in other sections) and compute the reversed data

$$q_i^{\text{rev}}(0) = q_i(0)$$
$$p_i^{\text{rev}}(0) = -p_i(0) \tag{2.1.6}$$

With these conditions, "at zero time" we can calculate, using equations (2.1.3), a new real motion that we will call $q_i^{\text{rev}}(t)$, $p_i^{\text{rev}}(t)$. We will say that the motion is reversible if

$$q_i^{\text{rev}}(t) = q_i(-t)$$
$$p_i^{\text{rev}}(t) = -p_i(-t) \tag{2.1.7}$$

that is, if the motion in the backward film agrees with a real motion with reversed conditions at zero time (we see that the initial conditions of one motion are the final ones of the other). Usually H [cf. (2.1.1)] is quadratic in the p_i, so that

$$TH(q_i, p_i) = H(Tq_i, Tp_i) = H(q_i, -p_i) = H(q_i, p_i) \tag{2.1.8}$$

In this case we will say that the Hamiltonian is time-symmetric. Then, if we make a T transformation (2.1.5) on (2.1.3), we find

$$\frac{dq_i}{dt} = -\partial_{(-p_i)} H$$
$$-\frac{d(-p_i)}{dt} = -\partial_{q_{\text{rev}}} H \tag{2.1.9}$$

and if we now change t by $-t$, we find again (2.1.3) as

$$\frac{dq_i}{d(-t)} = \partial_{(-p_i)} H$$
$$\frac{d(-)p_i}{d(-t)} = -\partial_{q_i} H \tag{2.1.10}$$

From this equation and (2.1.3) a motion $(q_i^{\text{rev}}, p_i^{\text{rev}})$ with data (2.1.6) must satisfy (2.1.7). Therefore:

Theorem 2.1.2. A usual Hamiltonian, quadratic in the p_i, yields a reversible motion.

The only condition to obtain a reversible motion is (2.1.8), which implies that the Hamiltonian is time-symmetric. Then reversible motion forms a group. However, irreversible motion does not form a group, since their inverse

does not even exist. Moreover, we will say that the initial conditions are time-symmetric, with respect to $t = 0$, if $p_i(0) = 0$, or

$$q_i(0) = q_i(0)$$
$$p_i(0) = -p_i(0) \tag{2.1.11}$$

Then, if the motion is reversible, we will have

$$q_i(t) = q_i(-t)$$
$$p_i(t) = -p_i(-t) \tag{2.1.12}$$

We call this motion time-symmetric with respect to $t = 0$, since the curves $q_i(t)$ are symmetric with respect to the vertical axis and the curves $p_i(t)$ are symmetric with respect to the origin of the coordinate system, as in Fig. 1. Therefore:

Theorem 2.1.3. If the motion is reversible and the condition at $t = 0$ is time-symmetric, then the motion is time-symmetric with respect to $t = 0$.

If all the motion were time-symmetric with respect to $t = 0$, then it would be impossible to define any arrow of time at $t = 0$, since past and future would look exactly the same in this instant of time.

2.2. Quantum Formalism

The quantum wave function for the same system treated in Section 2.1 reads

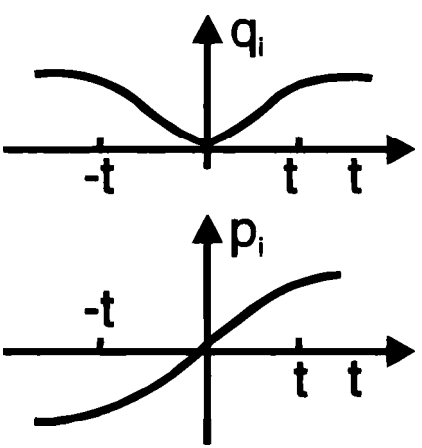

Fig. 1. The $q(t)$ and $p(t)$ functions for time-symmetric solutions, with respect to $t = 0$.

$$\Phi(q_i, t) = \langle q_i | \Phi(t) \rangle \tag{2.2.1}$$

This function belongs to a Hilbert space $\mathcal{H} = L^2$. This implies that if we introduce the inner product

$$(\Phi, \Psi) = \int \Phi^* \Psi \, d^N q \tag{2.2.1'}$$

with $(\Phi, \Phi) < \infty$, usually normalized as $(\Phi, \Phi) = 1$, it satisfies the Schrödinger equation

$$i \frac{\partial \Phi(q_i, t)}{\partial t} = H\Phi(q_i, t) \tag{2.2.2}$$

from which we can find the time evolution of the wave function $\Phi(q_i, t)$ by imposing conditions at zero time

$$\Phi(q_i, 0) = \Phi^0(q_i) \tag{2.2.3}$$

Then

$$\Phi(q_i, t) = e^{-iHt}\Phi(q_i, 0) = u(t)\Phi^0(q_i) \tag{2.3.3'}$$

Since we are now working in the configuration representation, in which the position and momentum operators are

$$\begin{aligned} \hat{q_i} &= q_i \\ \hat{p_i} &= -i\partial_{q_i}, \end{aligned} \tag{2.2.4}$$

the quantum version of (2.1.5) is

$$T\Phi(q_i, t) = \Phi^*(q_i, t) \tag{2.2.5}$$

For, if

$$\langle \hat{p_i} \rangle_\Phi = \int \Phi^*(q)(-i\partial_{q_i})\Phi(q) \, dq, \qquad \langle q_i \rangle_\Phi = \int \Phi^*(q) q_i \Phi(q) \, dq_i$$

then

$$\langle \hat{p_i} \rangle_{\Phi^*} = \int \Phi(q)(-i\partial_{q_i})\Phi^*(q) \, dq = -\langle \hat{p_i} \rangle_\Phi, \qquad \langle \hat{q_i} \rangle_{\Phi^*} = \langle \hat{q_i} \rangle_\Phi$$

(for more details see Messiah, 1962). In this case the wave function of the inverted motion will have as zero time data

$$\Phi_{\text{rev}}(q_i, 0) = \Phi^*(q_i, 0) = \Phi^{0*}(q_i) \tag{2.2.6}$$

and the motion will be reversible if

$$\Phi_{\text{rev}}(q_i, t) = \Phi^*(q_i, -t) \tag{2.2.7}$$

which is the quantum version of equations (2.1.7). If H is a Hamiltonian quadratic in p, it is easy to see that H is real (or time-symmetric), namely

$$H = H^* \tag{2.2.8}$$

We can thus formulate:

Theorem 2.2.1. If the Hamiltonian is real, the corresponding evolution is reversible.

Proof. From (2.2.6) and (2.2.8) we can obtain (2.2.7), since

$$\Phi_{\text{rev}}(t) = e^{-iHt}\Phi_{\text{rev}}(0) = e^{-iHt}\Phi^*(0) = (e^{iHt}\Phi(0))^* = \Phi^*(-t) \tag{2.2.9}$$

where we have omitted the variables q_i. Then, as in the classical case, a usual Hamiltonian yields a reversible motion.

We can also directly show that (2.2.2) is t-invariant, but the proof above is preferable because the role played by the condition at zero time can be seen explicitly. As in the classical case, reversible motion forms a group, since $u^{-1}(t) = u(-t)$ is a real motion, which is not the case for irreversible motion, in which the reversed motion is again not real. If $u(t_1)u(t_2) = u(t_1 + t_2)$ for $t_1, t_2, \geq 0$ only (and there is a unit), we will say that these motions form a semigroup. This is the case for irreversible motions. Let us now repeat all of this formalism, which so far we have introduced in the configuration representation, in an abstract way. The state of the system is defined by the ket $|\Phi(t)\rangle$ which belongs to the Hilbert space of states $\mathcal{H}$, and satisfies the Schrödinger equation

$$i\frac{d}{dt}|\Phi(t)\rangle = H|\Phi(t)\rangle \tag{2.2.10}$$

The inner product is symbolized as $\langle\Phi|\Psi\rangle = (\Phi, \Psi)$, and the normalization is $\langle\Phi|\Phi\rangle = 1$. We can find $|\Phi(t)\rangle$ by solving (2.2.10) with the condition at zero time

$$|\Phi(0)\rangle = |\Phi^0\rangle \tag{2.2.11}$$

This means that

$$|\Phi(t)\rangle = e^{-\mathrm{i}Ht}|\Phi(0)\rangle = u(t)|\Phi(0)\rangle \tag{2.2.11'}$$

Then, the T transformation can be defined as (Messiah, 1962; Roman, 1965)

$$T|\Phi(t)\rangle = K|\Phi(t)\rangle = |\Phi^*(t)\rangle \tag{2.2.12}$$

which means that we must conjugate the wave function in the configuration

representation and then go to the generic representation. K is known as the Wigner operator. More precisely, let $(\Phi(q_i, t)) = |\Phi(q_i, t))$ be the coordinates of the state vector in the configuration representation (wave function) and $|\Phi(t)\rangle$ the coordinates of the same vector in a generic representation; then

$$|\Phi(t)\rangle = U|\Phi(q_i, t)), \qquad UU^\dagger = 1 \tag{2.2.13}$$

Let K_0 be the conjugation operator in the configuration representation

$$K_0|\Phi(q_i, t)) = |\Phi^*(q_i, t)) \tag{2.2.14}$$

Then

$$K|\Phi(t)\rangle = KU|\Phi(q_i, t)) = UK_0|\Phi(q_i, t)) = UK_0U^\dagger|\Phi(t)\rangle \tag{2.2.15}$$

Thus, if K_0 is the conjugation in the configuration representation, the Wigner operator K in a generic representation reads

$$K = UK_0U^\dagger \tag{2.2.16}$$

It is easy to show that, in the configuration representation, K_0 has the following properties:

(a) K_0 is an antilinear, antiunitary operator, namely (Messia 1962), (a1) $K_0(\alpha|1\rangle + \beta|2\rangle) = \alpha^*K_0|1\rangle + \beta^*K_0|2\rangle$; (a2) if $|2\rangle = K_0|2\rangle$, $\langle 1| = \langle 1|K_0$, and $\hat{A} = K_0\, AK_0^\dagger$, then $\langle 1|\hat{A}\,|2\rangle = \langle 1|A|2\rangle^*$; (a3) $(\langle 1|K_0)|2\rangle = \langle 1|(K_0|2\rangle)^*$, *i.e.*, parentheses cannot be omitted.

(b) $K_0^2 = 1$ [at least for spin-zero fields (Messiah, 1962)].

(c)

$$K_0\hat{q}_iK_0^\dagger = \hat{q}_i$$

$$K_0\hat{p}_iK_0^\dagger = -\hat{p}_i$$

(d) We have

$$K_0cK_0^\dagger = c^{*\prime} \quad \text{if} \quad c \in \mathbb{C}$$

Therefore $K_0K_0^\dagger = 1$ and $K_0iK_0^\dagger = -i$. From (2.2.16), it is also easy to show that K has the same properties. As an exercise we can repeat formulas (2.2.6)–(2.2.9) in a generic representation. The time reversal is given by equation (2.2.12). The reverse initial condition is

$$|\Phi(0)_{\text{rev}}\rangle = K|\Phi(0)\rangle \tag{2.2.17}$$

and the condition of reversible motion reads

$$|\Phi(t)_{\text{rev}}\rangle = K|\Phi(-t)\rangle \tag{2.2.18}$$

We shall say that H is real if

$$H = KHK^{\dagger} \tag{2.2.19}$$

and usually H is endowed with this property, because equation (2.2.8) is satisfied in the configuration basis. Then, from (2.2.17) and (2.2.19) we can again deduce Theorem 2.2.1, but this time in a generic coordinate system:

$$\begin{aligned} |\Phi(t)_{\text{rev}}\rangle &= e^{-iHt}|\Phi(0)_{\text{rev}}\rangle = e^{-iHt}K|\Phi(0)\rangle \\ &= K(K^{\dagger}e^{-iHt}K|\Phi(0)\rangle) = Ke^{iHt}|\Phi(0)\rangle = K|\Phi(-t)\rangle \end{aligned} \tag{2.2.20}$$

Then, as in the classical case, a usual real Hamiltonian yields a reversible motion. In general, we will call a ket $|1\rangle$ (bra $\langle 1|$) real if

$$K|1\rangle = |1\rangle \qquad \text{or} \qquad \langle 1|K^{\dagger} = \langle 1|, \tag{2.2.21}$$

and an operator A real if

$$KAK^{\dagger} = A \tag{2.2.22}$$

From (2.2.19) we see that a usual Hamiltonian is a real operator. A basis $\{|i\rangle\}$ will be a real basis if all its kets are real:

$$K|i\rangle = |i\rangle \tag{2.2.23}$$

In a real basis, K is just the conjugation of the coordinates of the vectors, or of the coordinates of the operators:

$$K|\Phi\rangle = K\sum_i c_i|i\rangle = \sum_i c_i^*|i\rangle \tag{2.2.24}$$

$$KAK^{\dagger} = K\left(\sum_{ij} c_{ij}|i\rangle\langle j|\right)K^{\dagger} = \sum_{ij} c_{ij}^*|i\rangle\langle j|$$

Therefore, the configuration basis $\{|x\rangle\}$ is real. We will say that the conditions at $t = 0$ are time-symmetric if

$$|\Phi(0)\rangle = K|\Phi(0)\rangle \tag{2.2.25}$$

or in other words $|\Phi(0)\rangle$ is real. Then, if the evolution is reversible, we have

$$|\Phi(-t)\rangle = K|\Phi(-t)\rangle \tag{2.2.26}$$

and we say that the evolution is time-symmetric with respect to $t = 0$. So we have:

Theorem 2.2.2. If the evolution is reversible and the initial condition is time-symmetric, the evolution is time-symmetric. Then we can repeat what we said in the classical case. If all the quantum evolutions were time-symmetric with respect to $t = 0$, it would be impossible to define a quantum arrow of time at $t = 0$.

2.3. Statistical Formalism

We will deal simultaneously with the classical and quantum cases in order to establish an analogy or unified formalism that we will use below. Nevertheless, it should be stressed that there is a great difference between the classical and quantum cases. We will call the classical *distribution function or density* (resp., the quantum *density matrix*) a function (resp., matrix) endowed with the following properties:

$$\rho(q_i, p_i) \geq 0 \qquad \text{or} \qquad \rho(x) \geq 0 \tag{2.3.1}$$

$$\|\rho\| = \int_X \rho(q_i, p_i)\, dq_i\, dp_i = \int_X \rho(x)\, dx = 1$$

where X is the phase space. Distribution functions ρ belong to an L^1 Hilbert space called the classical Liouville space (resp., in the quantum mechanical formalism

$$\begin{aligned} \rho &= \rho^\dagger \\ tr(\rho) &= 1 \\ \rho_{\alpha\alpha} &\geq 0 \end{aligned} \tag{2.3.2}$$

density matrices ρ belong to a space $\mathscr{L} = \mathscr{H} \times \mathscr{H}$ called the quantum Liouville space). ρ satisfies the Liouville equation

$$i\partial_t \rho = L\rho \tag{2.3.3}$$

where

$$L = i\{H, ..\}_{PB} \tag{2.3.4}$$

[resp.,

$$L = [H, .,] = H \times 1 - 1 \times H \tag{2.3.5}$$

(cf. equation (2.A.24) for the definition of $\times$)]. Therefore the time evolution, in both classical and quantum cases, is

$$\rho(t) = e^{-iLt}\rho(0) = U(t)\rho(0) \tag{2.3.5$'$}$$

The T transformation of a density function is

$$T\rho(q_i, p_i) = \rho'(q_i, p_i) = \rho(q_i, -p_i) \tag{2.3.6}$$

(resp., the T transformation of a density matrix is

$$T\rho = \rho' = K\rho K^\dagger = \mathscr{K}\rho \tag{2.3.6$'$}$$

where $\mathscr{K} = K \times K^\dagger$). From equation (2.3.3), if the Hamiltonian is a usual time-symmetric one, we have classically

$$
\begin{aligned}
TL\rho(q_i, p_i) &= Ti\{H, \rho\}_{PB} \\
&= Ti\sum_i \partial_{q_i} H \partial_{p_i}\rho - \partial_{p_i} H \partial_{q_i}\rho \\
&= i\sum_i \partial_{q_i} H \partial_{-p_i} T\rho - \partial_{-p_i} H \partial_{q_i} T\rho \\
&= -i\{H, T\rho\}_{PB} = -LT\rho
\end{aligned} \tag{2.3.7}
$$

Therefore, if we T-transform classically equation (2.3.2), we obtain

$$
i\partial_t T\rho = -LT\rho \tag{2.3.8}
$$

[resp., if we T-transform the quantum Liouville equation (2.3.2), we obtain, if the Hamiltonian is a real usual one,

$$
KiK^\dagger\, \partial_t K\rho K^\dagger = (K \times K^\dagger L(K^\dagger \times K)K\rho K^\dagger \tag{2.3.8'}
$$

but $KiK^\dagger = -i$, and

$$
(K \times K^\dagger)L(K^\dagger \times K) = KHK^\dagger \times 1 - 1 \times KHK^\dagger = L
$$

so

$$
-i\partial_t K\rho K^\dagger = LK\rho K^\dagger \tag{2.3.8''}
$$

i.e., the same equation as the classical one (2.3.8)]. In both cases a minus sign appears. In the reverted solution we must change t by $-t$, namely

$$
T(t) = t' = -t \tag{2.3.9}
$$

So we have proved:

Theorem 2.3.1. The Liouville equation remains invariant under T transformations for a usual time-symmetric Hamiltonian.

Thus we have shown the complete isomorphism of the classical and quantum formalisms. From now on we will mainly use the quantum formalism, since it is the one which is better known to physicists. Let us therefore review the main properties of the usual Hamiltonian, in a real basis to simplify the treatment. From the equations of Section 2.2. we have

$$
\begin{aligned}
H &= H^\dagger \\
H &= H^* \\
H &= H^T
\end{aligned} \tag{2.3.10}
$$

which implies that the Hamiltonian is (I) self-adjoint, because it is an observable; (II) real, because for the usual Hamiltonian the motion is reversible; and (III) as a consequence, it is also symmetric. ρ belongs to a set which,

endowed with the inner product (2.A.1) below, becomes the Liouville–Hilbert space $\mathscr{L}$. From (2.A.29) and (2.3.10)) we can prove that the Liouvillian has the following properties in a real basis:

$$\begin{aligned} L &= L^{\dagger} \\ L &= L^{*} \\ L &= -L^{a} \\ L &= -L^{T} \end{aligned} \tag{2.3.11}$$

Then, from (2.3.11) and (2.A.18) we have

$$(iL) = (iL)^{a} \tag{2.3.12}$$

This property is important since from it we can deduce that the matrix ρ remains Hermitian under the time evolution always satisfying the Liouville equation (2.3.2). In fact, it follows from (2.A.18) that a product of self-associated commuting operators is also self-associated. Then, $(iL) = (iL)^{a}$ implies $(e^{-iLt})^{a} = e^{-iLt}$ and $\rho(0) = \rho(0)^{\dagger}$ implies $e^{-iLt}\rho(0) = [e^{-iLt}\rho(0)]^{\dagger}$, namely; $\rho(t) = \rho(t)^{\dagger}$. Finally, let us prove in another way that, if the Liouvillian is real, the evolution is reversible. Based on (2.1.7) and (2.2.7) we define a reversible motion, in a real basis, as

$$\rho_{\text{rev}}(t) = \rho^{*}(-t) \tag{2.3.13}$$

where $\rho_{\text{rev}}(t)$ is the motion with reversed condition at zero time:

$$\rho_{\text{rev}}(0) = \rho^{*}(0) \tag{2.3.14}$$

Now we can prove:

Theorem 2.3.2. If the Liouvillian is real, the evolution is reversible.

Proof. With the same reasoning as for (2.2.9) we have

$$\rho_{\text{rev}}(t) = e^{-iLt}\rho_{\text{rev}}(0) = e^{-iLt}\rho^{*}(0), = [e^{iLt}\rho(0)]^{*} = \rho^{*}(-t) \tag{2.3.15}$$

which shows that a motion with a real Liouvillian is reversible. QED

In a generic basis, (2.3.11)–to (2.3.15) read as follows: The Liouvillian is real or time symmetric if

$$\mathscr{K}L\mathscr{K}^{\dagger} = L \tag{2.3.16}$$

The evolution is time-symmetric if

$$\rho_{\text{rev}}(t) = K\rho(-t)K^{\dagger} = \mathscr{K}\rho(-t) \tag{2.3.17}$$

The conditions at time $t = 0$ are time-symmetric if

$$\rho_{rev}(0) = S\rho(0)K^{\dagger} = \mathcal{K}\rho(0) \tag{2.3.18}$$

A real Liouvillian and time-symmetric conditions at $t = 0$ yield a time-symmetric evolution since

$$\rho_{rev}(t) = e^{-iLt}\rho_{rev}(0) = e^{-iLt}K\rho(0)K^{\dagger} \tag{2.3.19}$$

$$K[e^{iLt}\rho(0)]K^{\dagger} = K\rho(-t)K^{\dagger}$$

The condition at $t = 0$ will be called time-symmetric if

$$\rho(0) = K\rho(0)K^{\dagger} = \mathcal{K}\rho(0) \tag{2.3.20}$$

Then, if the evolution is irreversible, we have

$$\rho(-t) = \mathcal{K}\rho(t) \tag{2.3.21}$$

and we find that the whole evolution is time-symmetric and we can repeat the procedure described for the previous cases. So we have:

Theorem 2.3.3. If the evolution is reversible and the condition at $t = 0$ is time-symmetric the evolution is time-symmetric with respect to $t = 0$.

Proof. If the Liouvillian satisfies (2.3.16) and the condition at $t = 0$ satisfies (2.3.20), all the evolution is time-symmetric, since

$$\mathcal{K}\rho(t) = \mathcal{K}(e^{-iLt}\rho(0)) = e^{i\mathcal{K}L\mathcal{K}^{\dagger}t}\mathcal{K}\rho(0) = e^{iLt}\rho(0) = \rho(-t) \tag{2.3.22}$$

Therefore the motion is time-symmetric if L is real and the condition at time $t = 0$ is time-symmetric. QED

2.4. Appendix 2A. Mathematical Theory of Superspace and Superoperators

We now have a short mathematical interlude, to define the notions of *superspace* and *superoperators* (Prigogine *et al.*, 1980).

2.4.1. The Quantum Case

Let us consider a Hilbert space $\mathcal{H}$ and the space $\mathcal{L} = \mathcal{H} \times \mathcal{H}$ of matrices on $\mathcal{H}$, *i.e.*, the Liouville–Hilbert space. Matrices will be symbolized by Greek lower case letters $\alpha, \beta, \ldots, \rho$ with coordinates $\alpha_{ij}, \beta_{ij}, \ldots, \rho_{ij}$. We will call the linear space of matrices the *superspace* $\mathcal{L}$ and the matrices *supervectors*. Let us define an inner product in the superspace $\mathcal{L}$:

$$\alpha \cdot \beta = (\alpha|\beta) = tr(\alpha^{\dagger}\beta) = \sum_{ij} \alpha_{ij}^{*}\beta_{ij} \tag{2.A.1}$$

Using this inner product, $\mathcal{L}$ becomes an L^2 Hilbert space. The norm of a supervector is thus

$$\|\alpha\| = \alpha \cdot \alpha = \sum_{ij} |\alpha_{ij}|^2 \geq 0 \tag{2.A.2}$$

We will consider the linear operators in superspace, which we will call *superoperators*, and which will be represented by capital Latin letters $A, B, \ldots, L$, with coordinates $A_{ij,kl}$, $B_{ij,kl}$. Superoperators act on matrices as

$$\begin{aligned} A\alpha &= \beta \\ \alpha A &= \beta \end{aligned} \tag{2.A.3}$$

We will use for these two equations the following rule for indices:

$$\begin{aligned} \sum_{kl} A_{ij,kl}\, \alpha_{kl} &= \beta_{ij}, \\ \sum_{kl} \alpha^T_{lk}\, A_{lk,ji} &= \beta^T_{ji} \end{aligned} \tag{2.A.4}$$

In the first equation we have used the usual multiplication "row by column" and α and β are considered as column vectors. In the second, we have transposed α and β since in these cases they are considered as row vectors. Since the superoperators have four indices, we can define more operations defining transposed and adjoints than for ordinary two-index matrices. So, for a superoperator A, we define the following:

(a) The *transpose* A^T is the superoperator such that

$$A\alpha = \alpha A^T \tag{2.A.5}$$

for all $\alpha \in \mathcal{L}$. Then,

$$A_{ij,kl}\alpha_{kl} = \alpha_{kl}\, A^T_{lk,ji} \tag{2.A.6}$$

so

$$A_{ij,kl} = A^T_{lk,ji} \tag{2.A.7}$$

Of course,

$$\begin{aligned} (A^T)^T &= A \\ (A_1 A_2)^T &= A_2^T\, A_1 \end{aligned} \tag{2.A.8}$$

and A is symmetric (antisymmetric) if

$$A = A^T \qquad (A = -A^T) \tag{2.A.9}$$

The *adjoint* $A^\dagger$ is the superoperator such that

$$A\alpha = (\alpha^\dagger A^\dagger)^\dagger \tag{2.A.10}$$

for all $\alpha \in \mathcal{L}$. Then,

$$A_{ij,kl}\alpha_{kl} = (\alpha^{*}_{lk}A^{\dagger}_{lk,ji})^{\dagger} = \alpha_{lk}(A^{\dagger}_{lk,ij})^{*} \tag{2.A.11}$$

so

$$A^{*}_{ij,kl} = A^{\dagger}_{kl,ij} \tag{2.A.12}$$

Of course,

$$(A^{\dagger})^{\dagger} = A$$

$$(A_1\ A_2)^{\dagger} = (A^{\dagger}_2 A^{\dagger}_1 \tag{2.A.13}$$

and A is Hermitian (anti-Hermitian) if

$$A = A^{\dagger} \qquad (A = -A^{\dagger} \tag{2.A.14}$$

The *associated* superoperator A^a is the superoperator such that

$$A\alpha = (A^a\alpha^{\dagger})^{\dagger} \tag{2.A.15}$$

for all $\alpha \in \mathcal{L}$. Then,

$$A_{ij,kl}\alpha_{kl} = (A^{a}_{ij,kl}\alpha^{\dagger}_{kl})^{\dagger} = (A^{a}_{ji,kl}\ \alpha^{*}_{lk})\dagger = (A^{a}\ *_{ji,kl}\ \alpha_{lk}) \tag{2.A.16}$$

so

$$A^{*}_{ij,kl} = A^{a}_{ji,lk} \tag{2.A.17}$$

Of course,

$$(A^a)^a = A$$

$$(A_1A_2)^a = A^a_1A^a_2 \tag{2.A.18}$$

and an operator is *adjoint-symmetric* (or *self-associated*) if

$$A = A^a \tag{2.A.19}$$

An adjoint-symmetric operator acting on a Hermitian matrix gives another Hermitian matrix. For, if

$$\alpha = \alpha^{\dagger}, \qquad A = A^a \tag{2.A.20}$$

then from equation (2.A.15) we have

$$A\alpha = (A\alpha)^{\dagger} \tag{2.A.21}$$

Putting everything together, we have

$$A_{ij,kl} = A^{T}_{lk,ji} = (A^{\dagger}_{kl,ij})^{*} = (A^{a}_{ji,lk})^{*} \tag{2.A.22}$$

and therefore

$$A^{aT} = A^{\dagger} \tag{2.A.23}$$

Let us now define a superoperator as a product of two operators, $A = \alpha \times \beta$, as

$$A\gamma = \alpha\gamma\beta, \qquad \forall\gamma \tag{2.A.24}$$

or, equivalently,

$$\sum_{kl} A_{ij,kl}\gamma_{kl} = \sum_{kl} \alpha_{ik}\gamma_{kl}\beta_{lj}$$

that is,

$$A_{ij,kl} = \alpha_{ik}\beta_{lj} \tag{2.A.25}$$

Then,

$$\sum_{kl} \gamma_{kl} A_{lk,ji} = \sum_{kl} \gamma_{kl}\alpha_{lj}\beta_{ik} = \sum_{kl} \beta_{ik}\gamma_{kl}\alpha_{lj} \tag{2.A.26}$$

and, from (2.A.4),

$$\gamma A = \beta\gamma\alpha \tag{2.A.27}$$

Therefore, we have from (2.A.24) and (2.A.27),

$$\begin{aligned} (\alpha \times \beta)\gamma &= \alpha\gamma\beta \\ \gamma(\alpha \times \beta) &= \beta\gamma\alpha \end{aligned} \tag{2.A.28}$$

The choice of the index position in equation (2.A.4) was made in order to obtain these simple multiplication rules. It is easy to prove that

$$\begin{aligned} (\alpha \times \beta)^T &= \beta \times \alpha \\ (\alpha \times \beta)^{\dagger} &= \alpha^{\dagger} \times \beta^{\dagger} \\ (\alpha \times \beta)^{\alpha} &= \beta^{\dagger} \times \alpha^{\dagger} \end{aligned} \tag{2.A.29}$$

The product $\times$ can be used to define the time inversion of matrices, since a time-inverted matrix is [equation (2.2.22)]

$$T\rho = K\rho K^{\dagger} = (K \times K^{\dagger})\rho = \mathcal{K}\rho \tag{2.A.30}$$

From this equation we can deduce the time inversion rule of superoperators, namely,

$$TA = (K \times K^{\dagger})A(K \times K^{\dagger})^{\dagger} = \mathcal{K}A\mathcal{K}^{\dagger} \tag{2.A.31}$$

Since [see equation (2.A.29)]

$$(K \times K^{\dagger})^{\dagger} = K^{\dagger} \times K$$

we have the alternative expression

$$TA = (K \times K^{\dagger})A(K^{\dagger} \times K) \tag{2.A.32}$$

We can also compute $(\alpha \times \beta)(\gamma \times \delta)$:

$$\begin{aligned}\sum_{kl} (\alpha \times \beta)_{ij,kl}(\gamma \times \delta)_{kl,nm} \\ = \sum_{kl} \alpha_{ik}\beta_{lj}\gamma_{kn}\delta_{ml} = \sum_{kl} \alpha_{ik}\gamma_{kn}\delta_{ml}\beta_{lj} \\ = (\alpha\gamma \times \delta\beta)_{ij,nm}\end{aligned} \tag{2.A.33}$$

giving

$$(\alpha \times \beta)(\gamma \times \delta) = (\alpha\gamma \times \delta\beta) \tag{2.A.34}$$

2.4.2. Classical Case

As we have seen, the quantum Liouville space is transformed in an L^2 Hilbert space by the inner product (2.A.1). In the same way it is convenient to define an inner product in the classical Liouville space $\mathscr{L}$, namely

$$(\rho|\sigma) = \int_X \rho^*(x)\sigma(x)\, dx \tag{2.A.35}$$

Using this inner product and Wigner functions (Section 6.7) the classical $\mathscr{L}$ becomes also an L^2 Hilbert space, and the classical equivalent of the quantum equation in the previous subsection can be found. Also, we can use the Wigner function integral of Section 6.7 to make this equivalence explicit.

3. THERMODYNAMICS

3.1. Classification of the Different Versions of the Second Law

The first law of thermodynamics is simply the conservation of energy. There is no conflict between dynamics and thermodynamics for this law. The problem is to derive the second law of thermodynamics based on dynamical considerations. The second law is expressed in many forms by different authors, and so we shall begin our research by making a classification of these forms. Let $S(t)$ denote the thermodynamic entropy of a closed system

(i) We denote a first-order second law by the statement

$$S(t) \geq S(t') \tag{3.1.1}$$

if $t \geq t'$ thus, according to this form, the entropy cannot decrease.

(ii) A stronger assertion would be a second-order second law: Equation (3.1.1) is satisfied and also

$$\lim_{t\to\infty} S(t) = S_* \tag{3.1.2}$$

In this case we assert that the system's entropy converges to a steady-state value S_*, which may not be unique; *e.g.*, it can be the entropy of a metastable state. Different preparations of the system could yield different final metastable states.

(iii) The final and strongest, or third-order, form of the second law is that (3.1.1) and (3.1.2) are satisfied, but also the limit (3.1.2) is unique. In this case the entropy of the system evolves to a unique maximum, irrespective of how it was prepared. We will find these different forms of the second law below.

3.2. Dynamics and Densities

In order to be as general as possible, we shall consider more generic systems than the ones of Section 2. This is not done just for the sake of mathematical generality, but also because we will later be forced, by the problems which we will be tackling, to consider such a system. So let us consider a system operating in a phase space X with an evolution law S_t more general than (2.1.3.), *i.e.*, a mapping S_t: $X \to X$ that changes the point x of X as t changes. X may have finite dimension d or infinite dimension, and t can be discrete or continuous. We will consider only "autonomous" processes, *i.e.*, such that $S_t(S_{t'}(x)) = S_{t+t'}(x)$, $S_0(x) = x$. Thus the mapping S can form either a group of transformations when $t, t' \in R$ (or Z) (*e.g.*, the evolutions with time-symmetric Hamiltonian or Liouvillian of Section 2) or a semigroup if $t, t' \in R^+$ (or N). In the two last cases (R^+, N) an equation such as (2.1.7) does not exist and the evolution is necessarily irreversible. For every point x_0 the successive points $S_t(x_0)$ form a system trajectory. To study an infinite number of initial points, or an infinite number of trajectories, we introduce the density functions $\rho(x) \in L^1(X)$, which obey

$$\int_X |\rho(x)|\, dx < \infty \tag{3.2.1}$$

such that

$$\rho(x) \geq 0, \qquad \|\rho(x)\| = 1 \tag{3.2.2}$$

where

$$\|\rho(x)\| = \int_X |\rho(x)|\, dx \tag{3.2.3}$$

is the L^1-norm of ρ. We postulate that a thermodynamic system is a system

which has, at any given time, states distributed throughout the phase space X, and the distribution of these states is characterized by the density function $\rho(x)$. We will define the ρ measure $\mu_\rho(A)$ of the set $A \subset X$ as

$$\mu_\rho(A) = \int_A \rho(x)\, dx \tag{3.2.4}$$

The Lebesgue (nonnormalized usual) measure of a set A will be denoted $\bar{\mu}_L(A)$. The uniform density will be

$$\rho_L(x) = \frac{1}{\bar{\rho}_L(X)} \tag{3.2.5}$$

and therefore the Lebesgue normalized measure is $\mu_{\rho_L}(X) = \mu_L(X) = 1$. We always write $\mu_L(dx) = dx$. Finally, X can be either Gibbs phase space Γ or Boltzmann phase space μ (Lasota and Mackey, 1985).

3.3. Gibbs Entropy

This entropy is defined as

$$H(\rho) = -\int_X \rho(x) \log \rho(x)\, dx \tag{3.3.1}$$

It is an additive quantity, by which we mean that the Gibbs entropy of a system formed by two subsystems is the sum of the two corresponding entropies. Then it is called an extensive quantity. The Gibbs entropy can be written as

$$H(f) = \int_X \eta(\rho(x))\, dx. \tag{3.3.2}$$

where the $\eta(\rho)$ function is defined as

$$\eta(\rho) = -\rho \log \rho \quad \text{for} \quad \rho > 0 \qquad \text{and} \qquad \eta(0) = 0 \tag{3.3.3}$$

and it is endowed with the property

$$\eta(\rho) \leq (\rho - \sigma)\eta'(\sigma) + \eta(\sigma) \tag{3.3.4}$$

Combining these last two formulas, we can prove the Gibbs inequality:

$$\rho - \rho \log \rho \leq \sigma - \rho \log \sigma \qquad \text{for} \quad \rho, \sigma > 0 \tag{3.3.5}$$

If ρ and σ are two normalized density functions, integrating the last equation, we have

$$-\int_X \rho(x) \log \rho(x)\, dx \leq -\int_X \rho(x) \log \sigma(x)\, dx \tag{3.3.6}$$

Only when $\rho = \sigma$ does the equality hold in (3.3.4.)–(3.3.6)

3.4. Microcanonical and Canonical Ensembles

Let us consider a space X with a finite Lebesgue measure: $\mu_L(X) < \infty$. Then the only density that will make the Gibbs entropy maximal is the uniform density of equation (3.2.5). Precisely we have:

Theorem 3.4.1. When $\mu_L(X) < \infty$, the density that maximizes the Gibbs entropy is the uniform density, $\rho_L(x)$ [cf. equation (3.2.5)]. For any other density $\rho \neq \rho_L, H(\rho) < H(\rho_L)$.

Proof. Choosing an arbitrary density ρ, from (3.3.6) we have

$$H(\rho) \leq -\int_X \rho(x) \log \sigma\,(x)\, dx \tag{3.4.1}$$

However, if $\sigma(x) = 1/\mu_L(X)$, the integrated Gibbs inequality (3.3.6.) gives

$$H(\rho) \leq -\log\left[\frac{1}{\mu_L(X)}\right] \tag{3.4.2}$$

since ρ is normalized to one. The equality holds if $\rho = \rho_L$, but the entropy corresponding to ρ_L is

$$H(\rho_L) = -\log\left[\frac{1}{\rho_L(X)}\right] \tag{3.4.3}$$

therefore $H(\rho) \leq H(\rho_L)$ for any density ρ and $H(\rho) < H(\rho_L)$ for $\rho \neq \rho_L$.

Clearly, if X is normalized so that $\mu_L(X) = 1$, then $H(\rho) \leq 0$. QED

The uniform density is also called the density of a microcanonical ensemble, and, as we can show, to define it we do not need to use any particular property of the thermodynamic system under consideration. Another, even more interesting, theorem is the following:

Theorem 3.4.2. Assume that there exists a nonnegative measurable function $\alpha(x)$ as well as an average or expectation mean value $\langle\alpha\rangle_\rho$ of that function over the entire X, weighted by the density ρ:

$$\langle\alpha\rangle_\rho = \int_X \alpha(x)\rho(x)dx \tag{3.4.4}$$

Then the maximum of the Gibbs entropy $H(\rho)$ subject to the constraint $\langle\alpha\rangle_\rho = \text{const}$ occurs for the density

$$\rho_*(x) = Z^{-1} e^{-\nu\alpha(x)} \tag{3.4.5}$$

where

$$Z = \int_X e^{-\nu\alpha(x)}\, dx \tag{3.4.6}$$

and ν is implicitly defined by the normalization condition

$$\langle\alpha\rangle_\rho = Z^{-1} \int_X \alpha(x) e^{-\nu\alpha(x)}\, dx \tag{3.4.7}$$

Proof. The proof again uses the integrated Gibbs equality (3.3.6) so

$$H(\rho) \leq -\int_X \rho(x) \log \rho_*(x)\, dx = -\int_X \rho(x)[-\log Z - \nu\alpha(x)]\, dx$$

$$= \log Z + \nu \int_X \rho(x)\alpha(x)\, dx = \log Z + \nu \langle\alpha\rangle\rho \tag{3.4.8}$$

However, it is equally easy to show that

$$H(\rho_*) = \log Z + \nu\langle\alpha\rangle\rho \tag{3.4.9}$$

and therefore $H(\rho) \leq H(\rho_*)$, with the equality holding if and only if $\rho = \rho_*$. QED

If $\alpha(x)$ is the energy of the system, ρ_* is the density of the Gibbs canonical ensemble at a temperature $T = \nu^{-1}$. (With many constraints $\langle\alpha_i\rangle_\rho$ we could define the density of a grand canonical ensemble.) We postulate also that there is a one-to-one correspondence between thermodynamic equilibrium states and the states of maximum entropy. Then, from the preceding theorems it would be natural also to postulate that the thermodynamic entropy S coincides with Gibbs entropy $H(\rho)$. In fact with this postulate we can obtain the usual equilibrium thermodynamics. But, as we will see below, this identification is not what we need to build a nonequilibrium thermodynamics, since the Gibbs entropy does not have the right properties in this case.

3.5. Reversible and Irreversible Systems

In Section 2 the properties of the Hamiltonian force the motion to be either reversible or irreversible. But in this section we study more general motions, so we are forced to repeat these definitions for the more general cases. Nevertheless, in order to prove some theorems, the motions cannot be completely general, so we will restrict ourselves to motion produced by Markov operators. Any linear operator P_t: $L^1 \to L^1$ such that

$$(a) \quad P_t\rho \geq 0, \qquad (b) \quad \|P_t\rho\| = \|\rho\| \tag{3.5.1}$$

for all $t \in R$, $\rho \geq 0$, $\rho \in L^1$, is a Markov operator, *i.e.*, an operator, which acting on a density, gives a another density. Markov operators have a number of useful properties. The most important is that if $\rho \in L^1$ and it is not restricted to $\rho \geq 0$, then

$$\|P_t\rho\| \leq \|\rho\| \tag{3.5.2}$$

which is known as the contractive property. A Markov operator is reversible (or time-symmetric) if

$$(a) \quad P_0\rho = \rho, \qquad (b) \quad P_t(P_{t'}\rho) = P_{t+t'}\rho \tag{3.5.3}$$

for all t, $t' \in R$ (or Z), or, in other words, reversible Markov operators form a group. The evolution operator $U(t) = e^{iLt}$ of (2.3.5′) is an example of a reversible Markov operator. This is so because it is generated by a time-symmetric or real Liouvillian L. However, if in the last definition we substitute R and Z by R^+ and N, we have the definition of an irreversible Markov operator. Irreversible Markov operators form a semigroup. The Gibbs entropy cannot be used in nonequilibrium theory since it may decrease under the action of some Markov operators (Lasota and Mackey, 1985), and therefore we cannot use this entropy to formulate a second law of thermodynamics, even in the first-order form. Nevertheless the Gibbs entropy is completely successful in equilibrium situations, so the entropy we will choose for non-equilibrium situations must coincide with the Gibbs entropy at equilibrium.

3.6. Conditional Entropy

If ρ and σ are two densities such that supp $\rho \subset$ supp σ, then the conditional entropy of density ρ with respect to density σ is

$$H_C(\rho|\sigma) = -\int_X \rho(x) \log \frac{\rho(x)}{\sigma(x)}\, dx \tag{3.6.1}$$

The conditional entropy is always definite; *i.e.*, it is finite or equal to $-\infty$. Evidently the conditional entropy measures the deviation of the density ρ from the density σ. Conditional entropy has two very important properties: (i) Since ρ and σ are both densities, the integrated Gibbs inequality (3.3.6) implies that $H_C(\rho|\sigma) \leq 0$. It is only when $\rho = \sigma$ that the equality holds. (ii) If ρ_L is the constant density of the microcanonical ensemble throughout the phase space X, then $H_C(\rho|\rho_L) = H(\rho) - \log \overline{\mu}_L(X)$. Therefore, in this case, conditional entropy is a generalization of Gibbs entropy. As $H_C(\rho|\rho_*) = 0$ when $\rho = \rho_*$ it is reasonable to postulate that

$$S - S_* = H_C(\rho|\rho_*) \tag{3.6.2}$$

e.g., when ρ_* is the density of the canonical ensemble. We will see that this definition is completely satisfactory, and that, using equation (3.6.2), we can formulate the second law of thermodynamics in its second- and third-order forms. The first result along these lines is a weak first-order form of the law of thermodynamics, namely that the conditional entropy is never decreasing, as is proved by:

Theorem 3.6.1. Let P_t be a Markov operator. Then

$$H_C(P_t\rho|P_t\sigma) \geq H_C(\rho|\sigma) \tag{3.6.3}$$

for all densities ρ and σ. An additional result is the following: if $\sigma = \rho_*$ is stationary, namely $P_t\rho_* = \rho_*$, then

$$H_C(P_t\rho|\rho_*) \geq H_C(\rho|\rho_*) \tag{3.6.4}$$

Thus this conditional entropy is always a nondecreasing function bounded above and $H_{\max} = H_C(\rho_*|\rho_*) = 0$. Therefore this conditional entropy converges as $t \to \infty$, although more information about the evolution is required in order to find the limiting value. Furthermore, if the stationary density is uniform, namely that of the microcanonical ensemble, we have

$$H(P_t\rho) \geq H(\rho) \tag{3.6.5}$$

for all nonnegative ρ. Now $H_{\max} = -\log[1/\overline{\mu}_L(X)]$ and, as in the general case, we have convergence when $t \to \infty$. Therefore (3.6.2) seems a reasonable assumption. But when the Markov operator is reversible all these nice inequalities become equalities and the problem of the thermodynamic entropy reappears. In fact:

Theorem 3.6.2. If P_t is a reversible Markov operator, then the conditional entropy is absolutely constant for all times t and equal to the value determinated by the choice of the initial densities ρ and σ. That is,

$$H_C(P_t\rho|P_t\sigma) = H_C(\rho|\sigma) \tag{3.6.6}$$

for all t.

Proof. Since P_t is reversible, by the previous theorem it follows that

$$\begin{aligned} H_C(P_{t+t'}\rho|P_{t+t}\sigma) &= H_C(P_tP_{t'}\rho|P_tP_{t'}\sigma) \\ &\geq H_C(P_t\rho|P_t\sigma) \geq H_C(\rho|\sigma) \end{aligned} \tag{3.6.7}$$

for all t, t' since P_t is reversible. So let us choose $t' = -t$; then for all times we have

$$H_C(\rho|\sigma) \geq H_C(P_t\rho|P_t\sigma) \geq H_C(\rho|\sigma) \tag{3.6.8}$$

and therefore

$$H_C(P_t\rho|P_t\sigma) = H_C(\rho|\sigma) \tag{3.6.9}$$

for all t. QED

So in this case the conditional entropy is for ever fixed and determined by the method of preparation of the system. So we have gained nothing if the Markov operator is reversible.

4. THE CLASSICAL EVOLUTION

In this section we will study "classical evolutions" in the sense that these evolutions are not quantum ones. Nevertheless the evolutions will be just as general as that of the preceding section, *i.e.*, not necessarily those of Section 2.

4.1. The Frobenius–Perron Operator

A transformation S_t is called a measurable transformation if $\mu_*(S_t^{-1}(A))$ is well defined for all subsets $A \subset X$, where $S_t^{-1}(A) = B$ is the counterimage of A, or in other words, $S_t(B) = A$. It should be remarked that even if a unique $S_t^{-1}(x)$ may not exist (as in the case of irreversible evolutions) the counterimage does exist, since it is the set of all the points $x \in B$ that will go to A under the action of S_t. The transformation is nonsingular if $\mu_*(S_t^{-1}(A) = 0 \Leftrightarrow \mu_*(A) = 0$. If S_t is a nonsingular transformation, then the unique operator P_t: $L^1 \to L^1$ defined by

$$\int_A P_t\rho(x)\, dx = \int_{S_t^{-1}(A)} \rho(x)\, dx \tag{4.1.1}$$

is called the Frobenius–Perron operator corresponding to S_t. For each S_t the Frobenius–Perron operator is unique. If $\rho \geq 0$, then $P_t\rho \geq 0$. As $S_t^{-1}(X) = X$ then $\|P_t\rho\| = \|\rho\|$ and these operators are Markov operators. Operator $U(t)$ of (2.3.5′) is a Frobenius–Perron operator. S_t is ρ-measure preserving if

$$\mu_\rho(S_t^{-1}(A)) = \mu_\rho(A)$$

for all sets A. Measure-preserving transformations are necessarily nonsingular; we will also say that the measure μ_ρ is invariant under the transformation. The Liouville theorem shows that transformation $U(t)$ of (2.3.5′) is Lebesgue-measure preserving. We will call a state ρ steady if $P_t\rho = \rho$, for all t and it will be symbolized by ρ_*. In due time it will be also call a state of "thermodynamic

equilibrium." The relation between invariant measures and the Frobenius–Perron operator is stated by:

Theorem 4.1.1. Let S_t be a nonsingular transformation and P_t its Frobenius–Perron operator. Then there exist a state of thermodynamic equilibrium whose density ρ_* is a stationary state of P_t if and only if the measure μ_*

$$\mu_*(A) = \int_A \rho_*(x)\, dx \tag{4.1.2}$$

is invariant.

Therefore the transformation $U(t)$ which preserves Lebesgue measure necessarily has an equilibrium steady state, *e.g.*, the uniform state of the microcanonical ensemble. But Theorem 4.1.1 says nothing about the uniqueness of the equilibrium state. We will discuss this problem in the next section. A point $x \in A \subset X$ is called a recurrent point if there is some time $t > 0$ such that $S_t(x) \in A$. An important result is Poincaré recurrence:

Theorem 4.1.2. Let S_t be a transformation with an invariant measure μ_* operating in a finite space X, $\mu_*(X) < \infty$, and let A be a subset of X with positive ρ_*-measure. Then there exists a point x in A that is recurrent.

Proof. Assume the contrary, *i.e.*, that there are no recurrent points in A. This then implies that $S_t^{-1}(A) \cap A = \emptyset$ for all times $t > 0$, and thus that $S_t^{-1}(A) \cap S_{t'}^{-1}(A) = \emptyset$ for all positive times $t \neq t'$. However, since S_t is measure preserving, this implies that $\mu_*(S_t^{-1}(A)) = \mu_*(S_{t'}^{-1}(A))$ and this, coupled with the pairwise disjoint nature of the sets $S_t^{-1}(A)$ and $S_{t'}^{-1}(A)$, leads to

$$\sum_{t=0}^{\infty} \mu_*(A) = \sum_{t=0}^{\infty} \mu_*(S_t^{-1}(A)) = \mu_*\left[\bigcup_{t=0}^{\infty} S_t^{-1}(A)\right] \leq \mu_*(X) < \infty \tag{4.1.3}$$

The only way this inequality can be satisfied is for $\mu_*(A)$ to be zero, which is a contradiction. Thus, we conclude that A contains recurrent points. QED.

Therefore, in the ordinary mechanical motion of finite systems, almost any point is recurrent, since the sets of nonrecurrent points have zero measure. This fact seems to prevent the existence of irreversible evolutions; it is impossible to reach a final equilibrium state, since the system will come back as close to its initial condition as we wish, if we wait long enough. The period we have to wait is called the Poincaré recurrence time. There are two ways to avoid this problem: (i) The practical way is to compute the recurrence time. It turns out that in a usual system (say with a number of molecules of

the order of Avogadro's number) the time is much bigger than the age of the universe, so the return to the initial conditions is practically unobservable. (ii) The theoretical way is to consider that irreversibility is not a notion of classical mechanics, but a notion which may only be defined in statistical mechanics, where we deal with statistical ensembles of infinite identical systems. Then the recurrent time of the ensemble, namely the time such that we reobtain the initial condition in *each one of the infinite identical systems* is of course infinite and the problem is theoretically solved. In the following subsections we will study some properties of dynamical systems ordered by their increasing chaotic behavior.

4.2. Ergodicity

It would be interesting to know whether the equilibrium state of Theorem 4.1.1 is unique or not. To answer this question we must introduce some new concepts. (i) A set A such that $S_t^{-1}(A) = A$ is called an invariant set. (ii) Any invariant set A such that $\mu_*(A) = 0$ or $\mu_*(X \backslash A) = 0$ is called trivial. (iii) A nonsingular transformation S_t is called ρ_*-ergodic if every invariant set A is a trivial subset of the phase space X, *i.e.*, either $\mu_*(A) = 0$ or $\mu_*(X \backslash A) = 0$. This means that if we consider a generic nonsingular subset A, the time-evolved counterimage of this subset, $S_t^{-1}(A)$, will wander around all X since A cannot be invariant. (iv) If ρ_* is the uniform density of the microcanonical ensemble, we will say that S_t is uniformly ergodic. The motion within almost all tori of integrable classical mechanical systems is ergodic (Tabor, 1980; Arnold and Avez, 1968). Ergodicity is therefore a very usual property of the mechanical systems of Section 2. The connection of the uniqueness of the equilibrium state to the properties of the operators is stated in the following:

Theorem 4.2.1 (Lasota and Mackey, 1985). Let S_t be a nonsingular transformation and P_t the corresponding Frobenius–Perron operator. S_t is ρ_*-ergodic if and only if P_t has a unique state of thermodynamic equilibrium with associated stationary density ρ_*, namely a density such that $P_t\rho_* = \rho_*$. Hence, ergodicity is the necessary and sufficient condition for the uniqueness of thermodynamic equilibrium, and allows us to formulate a third-order second law. But this of course is only half the picture, because we must also understand why the system evolves to this equilibrium state. Let us state an important result:

Theorem 4.2.2 (Lasota and Mackey, 1985). Let S_t be a nonsingular transformation and P_t the corresponding Frobenius–Perron operator with stationary density $\rho_* > 0$ for all points in phase space X. Then S_t is ρ_*-ergodic if and only if $\{P_t\rho\}$ is Césaro convergent to ρ_* for all densities ρ, *i.e.*, if and only if

$$\lim_{t\to\infty} \frac{1}{t} \sum_{k=0}^{t-1} (P_k \rho|\sigma) = (\rho_*|\sigma) \tag{4.2.1}$$

in the discrete-time case, or if and only if

$$\lim_{t\to\infty} \frac{1}{t} (P_t \rho|\sigma)\, dt = (\rho_*|\sigma) \tag{4.2.2}$$

in the continuous-time case, for all bounded measurable functions σ and where $(\rho|\sigma) = \int_X \rho(x)\sigma(x)\mu(x)\, dx$ [in this case $\mu(x)$ is an arbitrary measure] is a generalization of inner product (2.A.2.1).

4.3. Mixing

This will be the main property of the dynamical systems we will study, and it serves to guarantee the approach of the system to an equilibrium state. Let S_t be a ρ_*-measure-preserving transformation operating in a normalized space X [$\mu_*(X) = 1$]. Then S_t is called ρ_*-mixing if

$$\lim_{t\to\infty} \mu_*(A \cap S_t^{-1}(B)) = \mu_*(A)\mu_*(B) \tag{4.3.1}$$

for all sets A and B. If ρ_* is the uniform density of the microcanonical ensemble, then we will say that S_t is uniformly mixing. Some tori of mechanical nonintegrable system are broken, and then a chaotic motion in phase space takes place. Chaos, most likely with mixing properties, occurs very frequently in mechanical systems (Tabor, 1980; Arnold and Avez, 1968). A very important and popular example of a uniformly mixing transformation is the so-called baker's transformation, which operates in the phase space $X = [0, 1] \times [0, 1]$ and is defined by the following procedure: (i) Squeeze the 1×1 square to a $2 \times \frac{1}{2}$ rectangle, and (ii) cut the rectangle vertically into two rectangles and pile them up to form another 1×1 rectangle. In doing so, the point of the square will move as

$$(x, y) \to S(x, y) = \begin{cases} \left(2x, \frac{1}{2} y\right) & \text{if} \quad 0 \le x \le \frac{1}{2} \\ \left(2x - 1, \frac{1}{2} + \frac{1}{2} y\right) & \text{if} \quad \frac{1}{2} \le x \le 1 \end{cases} \tag{4.3.2}$$

The transformation is shown in Fig. 2, where in the first square, the one corresponding to $t = 0$, the lower half is shaded and corresponds to a subset B. It is easy to see that this transformation is reversible. The fate of this area B evolving to the future is shown on the right-hand side of the figure (it is transformed into a great number of horizontal strips with area $\frac{1}{2}$), and evolving

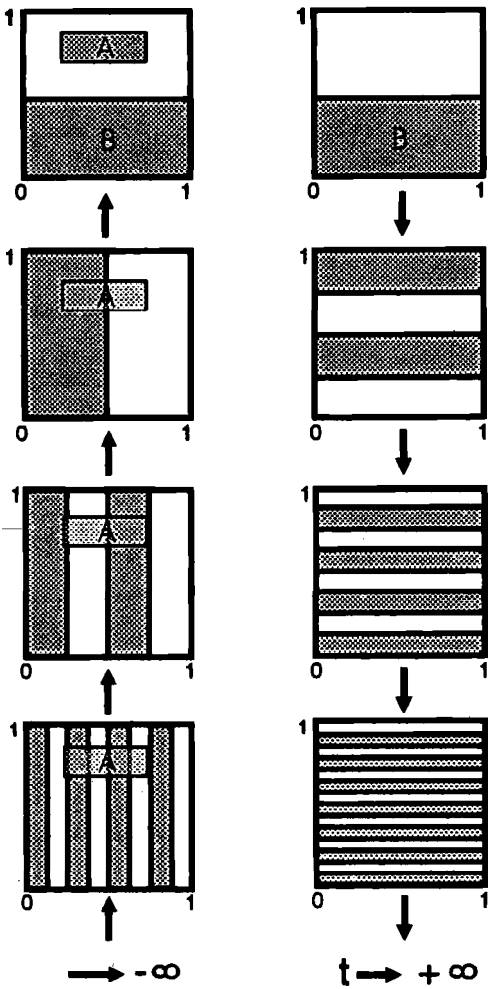

Fig. 2. The baker's transformation.

to the past in the left-hand side (the strips are now vertical). A smaller subset A is also shown. It is then easy to verify that condition (4.3.1) is fulfilled [the final measure of $S_t(B) \cap A$ will be $\frac{1}{2}\mu_L(A)$, since the initial measure of B is just $\frac{1}{2}$, so (4.3.1) is satisfied]. We will study this transformation in detail in Section 4.5 using the extended dynamics technique. Much more complicated mixing evolutions than (4.3.2) can be invented. In fact, the baker's transformation is the simplest of all; it is the simplest model of the famous Gibbs ink drop. Gibbs attempted to explain the essence of irreversibility with the ink drop model. If a drop of blue ink is introduced in a glass of water,

even if the volume of the ink drop remains constant (as the volume of any subset of mechanical phase space, according to Liouville's theorem) we will have, after a while, a homogeneous mixture of bluish water. What happens is that the motion of the water is mixing and therefore the ink drop is deformed (even if its volume is constant) in such a way that it is transformed into a set of very thin filaments that are present in every portion of the water, giving the result that the water has become bluish. The growth of this filament-like structure gives an arrow of time and it was for Gibbs the essence of irreversibility. This phenomenon is modeled by the baker's transformation. In fact, let us consider a small rectangle $a \times b$ within the square 1×1, and let us consider a small stain of lower quality flour within the bread dough. The height of the stain will successively become $\frac{1}{2}b, \frac{1}{4}b, \ldots, (1/t)b, \ldots$, while the base of the stain will become $2a, 4a, \ldots, ta, \ldots$, in such a way that the area is conserved. Eventually a time is reached when $ta > 1$, and then the stain will be cut in two, and then in four, eight, etc., in such a way that it will become a set of horizontal filaments of decreasing height, namely a "cubist" picture of the ink drop, so that the baker's transformation is just a model of the ink drop phenomenon. If now we consider the much more complicated evolution of the ink drop, and if the volume of the ink drop is 1% of the volume of the water, it is clear that the motion of usual water is mixing according to definition (4.3.1), as the baker's transformation. In fact, if the motion is mixing, when $t \to \infty$ every subset $A \subset X$ will have 1% of ink and therefore the distribution of ink will become homogeneous. As this is the case with the real ink drop, we can conclude that the real motion is mixing. It is a straightforward consequence of the definition that ρ_*-mixing implies ρ_*-ergodicity. In fact, if B is an invariant set, (4.3.1) reads

$$\mu_*(A \cap B) = \mu_*(A)\, \mu_*(B)$$

for all sets A. Now if we take $B = A$, we obtain $\mu_*(B) = [\mu_*(B)]^2$ and therefore either $\mu_*(B) = 0$, or $\mu_*(B) = 1$ [so in this case $\mu_*(X - B) = 0$]. So the evolution is ρ_*-ergodic. Now we have arrived at our most important theorem.

Theorem 4.3.1 (Lasota and Mackey, 1985). Let S_t be an ergodic transformation, with stationary density ρ_* of the associated Frobenius–Perron operator, operating in a phase space of finite ρ_*-measure. Then S_t is mixing if and only if $\{P_t\rho\}$ is weakly convergent to ρ_*, *i.e.*, if

$$\lim_{t\to\infty} (P_t\rho, \sigma) = (\rho_*, \sigma) \tag{4.3.3}$$

where σ is a bounded measurable function.

If a sequence is weakly convergent as well as Césaro convergent, then we can see again that mixing evolutions are ergodic. So the mixing property

ensures a weak convergence of $\{P_t\rho\}$ to ρ_*. But in the example of the baker's transformation the strong limits toward the far past or the far future do not exist. In fact, the support of any distribution function (if it has a measure $<$ 1) will be a set of infinite horizontal, toward the future or vertical lines toward the past. These sets cannot be the support of any regular distribution function. Nevertheless the weak limit (4.3.3) does exist with $\rho_* = 1$. The physical meaning of Theorem 4.3.1 is very clear: Let us consider a (nonviscous) fluid in motion in a cubic box. As energy is conserved, the motion will never stop, and therefore, according to the laws of mechanics, equilibrium will never be attained, and $P_t\rho$ will have no limit. This will be the case if the motion is oscillatory, namely a pressure wave which oscillates back and forth between two parallel walls of the box. But if the motion is mixing, it is so complicated that there are portions of the fluid moving in every direction in the vicinity of every point of the box. In this case, if we take the inner product $(P_t\rho|\sigma)$, we are making an average that goes to an equilibrium average $(\rho_*|\sigma)$ when $t \to \infty$. Therefore, even if there is always motion, the motion average gives an image of equilibrium. This is the profound meaning of Theorem 4.3.1 and the way to obtain a synthesis of the apparent contradiction of dynamics and thermodynamics: even if the dynamics says that the energy is conserved and the motion will never stop, there is a thermodynamic equilibrium in average, because the motion is mixing. From this point, the extended dynamics and coarse graining follow different paths, as we have explained in the introduction and we will discuss below. But let us remember that the problem is not completely solved, since all the nice inequalities of Section 3.6, which are necessary to explain the second law are equalities for reversible systems, and all systems in nature are considered to be, at least microscopically, reversible. There are systems endowed with properties more chaotic than mixing. They are (i) Kolmogorov systems (Lasota and Mackey 1985) that are necessarily mixing (Walter, 1982), (ii) Anosov systems (Lasota and Mackey, 1985; Arnold and Avez, 1968; Anosov, 1963), and (iii) Bernoulli systems, the most chaotic of all (Schild, 1979). The baker's transformation is, in fact, a Bernoulli system (Antoniou and Tasaki, 1991).

4.4. Exactness

We will now introduce a property which will (apparently) solve all our problems. If S_t is a ρ_*-measure-preserving transformation operating in a phase space X, then S_t is said to be ρ_*-exact if

$$\lim_{t\to\infty} \mu_*(S_t(A)) = 1 \tag{4.4.1}$$

for all sets A of nonzero measure. This is possible even if S_t is ρ_*-measure preserving, since an evolution is measure preserving if (4.1.2) is satisfied

and this equation is not equivalent to $\mu_*(S_t(A)) = \mu_*(A)$ if the evolution is not reversible. The Renyi map is a good example. Let us consider a dyadic Renyi map:

$$R: \quad [0, 1), \to [0, 1), \qquad x \to Rx = 2x(\text{mod } 1) \tag{4.4.2}$$

As the length of any subset A is multiplied by two in each transformation, this map is exact since it satisfies (4.4.1). In any case it is also measure preserving. In fact, let us consider, *e.g.*, the subset $A = [0,1/2)$; $R^{-1}(A)$ is $[0, 1/4) \cup [1/2, 3/4)$, and therefore both subsets have measure 1/2. If ρ_* is the uniform density of the microcanonical ensemble, we say that S_t is uniformly exact. The essential point to understand is that reversible system cannot be exact. In fact, for a reversible ρ_*-measure-preserving transformation, we have

$$\mu_*(S_t(A)) = \mu_*[S_t^{-1}(S_t(A))] = \mu_*(A) \tag{4.4.3}$$

and thus the definition of exactness is violated. Since, usually, classical dynamical systems are (by the Liouville theorem) measure preserving, and also reversible, they are not exact. Nevertheless, as we will see, exactness is really the property we are looking for. Precisely:

Theorem 4.4.1 (Lasota and Mackey, 1985). If S_t is a ρ_*-measure-preserving transformation operating on a finite normalizable phase space X and P_t is the associated Frobenius–Perron operator corresponding to S_t, then S_t is ρ_*-exact if and only if

$$\lim_{t\to\infty} \|P_t\rho - \rho_*\| = 0 \tag{4.4.4}$$

Therefore ergodicity corresponds to Césaro convergence; mixing corresponds to weak convergence; exactness corresponds to strong convergence (*i.e.*, convergence in the norm). A strongly convergent sequence is also weakly convergent. Thus we can deduce that exact evolutions are also mixing evolutions and therefore ergodic evolutions. Moreover, since we are looking for a strong limit, we see that working with ordinary distribution functions, we will find this limit only if the transformation is exact. But ordinary classical (microscopical) systems are not exact, since they are reversible and measure preserving. As an example, we have shown that the reversible baker's transformation does not have strong limits toward the past and the future. In fact the baker's transformation, being reversible, cannot be exact. Thus our problem can now be clearly stated: if we want a strong limit, our evolutions must be exact. But exact evolutions are not reversible and all microscopic transformations are reversible, and therefore we cannot have a strong limit. Furthermore we also have the following result:

Theorem 4.4.2 (Lasota and Mackey, 1985). Let P_t be a Markov operator operating in phase space X. Then the conditional entropy of $P_t\rho$ with respect to density ρ_* goes to a maximum value of zero as $t \to \infty$, *i.e.*,

$$\lim_{t\to\infty} H_C(P_t\rho|\rho_*) = 0 \tag{4.4.5}$$

if and only if P_t is ρ_*-exact. This theorem tells us the necessary and sufficient criteria to be able to state the second law of thermodynamics in third-order form, namely, for the entropy of the system to converge to its maximum value regardless of the way in which the system was prepared. This condition is that the system must evolve according to an exact transformation. But such systems do not exist in nature. So dynamics cannot be related, at least trivially, with thermodynamics. Therefore our theory must be modified one way or the other.

4.5. Mixing Studied by the Extended Dynamics Technique

Mixing evolutions are studied by the extended dynamics technique in Antoniou and Tasaki (1991, 1993a, b), using a perturbative method, which can be implemented in any example. For technical reasons we will present the most important mixing evolutions and we refer to the above papers for the general perturbative method.

4.5.1. The Renyi Maps

The β-adic Renyi map R on the interval [0, 1) is the multiplication, modulo 1, by the integer $\beta \geq 2$:

$$R: \quad [0.1) \to [0, 1); \qquad x \to Rx = \beta x \,(\text{mod } 1) \tag{4.5.1.1}$$

The forward iteration of the Renyi map n times defines a "cascade" or time evolution with time $t = n \in Z$. This evolution only preserves the Lebesgue measure, as we have shown in (4.4.3). The density functions $\rho(x)$ evolves according the Frobenius–Perron operator U:

$$U\rho(x) = \frac{1}{\beta} \sum_{r=0}^{\beta-1} \rho\left(\frac{x + r}{\beta}\right) \tag{4.5.1.2}$$

The Gel'fand-Maurin theorem (Theorem 4.A.1; see below) tells us that we can find a spectral expansion in the eigenvectors of this operator in an adequately rigged Hilbert space. In fact, using the perturbative methods of Antoniou and Tasaki (1991, 1993a, b), the spectral decomposition of U can be found, and reads

$$U = \sum_{n=0}^{\infty} \frac{1}{\beta^n} |B_n)(\tilde{B}_n| \tag{4.5.1.3}$$

where

$$|B_n(x)) = |x^n + \sum_{m=0}^{n-1} x^m \frac{n!}{m!\,(n-m)!} B_{n-m}) \tag{4.5.1.4}$$

where $B_n(x)$ is the n-degree Bernoulli polynomial defined by the generating function

$$\frac{ze^{zx}}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n(x)}{n!} z_n \tag{4.5.1.5a}$$

and

$$(\tilde{B}_n| = \begin{cases} (1| & n = 0 \\ \dfrac{(-1)^{(n-1)}}{n!} \{\delta^{(n-1)}(x-1) - \delta^{(n-1)}(x)\}|, & n = 1, 2, \end{cases} \tag{4.5.1.5b}$$

where $(1|$ is the constant distribution function. From (4.5.1.5b) we can see that the elements of the spectral decomposition (4.5.1.3) do not belong to $\mathscr{L}$, but to a larger space where the Dirac δ must have a precise mathematical meaning. This space is, in fact, a rigged Hilbert space, which we will define below, in full agreement with the Gel'fand–Maurin theorem. The system $\{|B_n), (\tilde{B}_n|\}$ is biorthonormal and complete, obeying

$$(\tilde{B}_n|B_m) = \delta_{nm} \tag{4.5.1.6}$$

$$\sum_{n=0}^{\infty} |B_n)(\tilde{B}_n| = 1 \tag{4.5.1.7}$$

The spectral decomposition (4.5.1.3) acquires a precise mathematical meaning if we define as test space Φ the space of polynomials $\mathscr{P}$. This space is dense in $\mathscr{L} = L^2$, nuclear (in fact, it is the union of an infinite and discrete set of finite-dimensional spaces), complete, and stable under U, and U is continuous in the topology of $\mathscr{P}$. It is therefore an appropriate test space to give meaning to the spectral decomposition whose elements belong to $\Phi^{\times}$. But other kinds of test function spaces can be defined and we will obtain different spectra. For example, a continuous set of eigenfunctions can be found, with adequate rigging, showing that the Renyi maps have continuous spectrum, namely the set of complex numbers z such that $|z| < 1$ (all the mixing operators have a spectral decomposition with a continuous spectrum in Hilbert space). If $t \in Z$, from (4.5.1.3 and 4.5.1.6) we can see that the evolution operator is

$$U^t = \sum_{n=0}^{\infty} \frac{1}{\beta^{nt}} |B_n)(\dot{B}_n| = |1)(1| + \sum_{n=1}^{\infty} \frac{1}{\beta^{nt}} |B_n)(\tilde{B}_n| \qquad (4.5.1.8)$$

If we want to work only in space $\mathscr{L}$, we must remember that all the formulas above are just weak equations; for example, the last is just

$$(\sigma|U^t\rho) = (\sigma|1)(1|\rho) + (\sigma| \sum_{n=1}^{\infty} \frac{1}{\beta^{nt}} |B_n)(B_n|\rho) \qquad (4.5.1.9)$$

for all $\rho, \sigma \in \mathscr{P}$, and ρ is a density. Then as $\beta > 1$ and ρ is normalized, we have

$$\lim_{t\to\infty} (\sigma, U^t\rho) = (\sigma|1) \qquad (4.5.1.10)$$

in perfect agreement with Theorem 4.3.1. Equation (4.5.1.9) is simply the weak version (or coarse graining) of (4.5.1.8), which allows us to work within space $\mathscr{L}$, but always using weak limits as (4.5.1.10). However, if we work with functionals directly, namely in space $\Phi^{\times}$, then from (4.5.1.8) we can say that

$$\lim_{t\to\infty} U^t|\rho) = |1) \qquad (4.5.1.11)$$

which is a strong limit, namely the extended dynamics version of (4.5.1.10). If we call $\beta = e^{-\gamma} < 1$, then from (4.5.1.9) we can also say

$$\rho(t) = U^t\rho = \rho_* + \rho_1(t)e^{-\gamma t}, \qquad \rho_* = |1) \qquad (4.5.1.12)$$

where $\rho_* = |1)$ is the equilibrium distribution function and $\rho_1(t)e^{-\gamma t}$ is something like a "fluctuation" around the equilibrium state. We include this last equation because we will find a similar one in the quantum case.

4.5.2. The Baker's Transformation

The β-adic, $\beta = 2, 3, \ldots$, baker's transformation in the unit square $Y = [0, 1) \times [0, 1)$ is a two-step operation: (i) Squeeze the 1×1 square to a $\beta \times 1/\beta$ rectangle, and (ii) cut the rectangle vertically into β rectangles and pile them up to form another 1×1 square. Then

$$(x, y) \to B(x, y) = \left(\beta x - r, \frac{y + r}{\beta}\right) \qquad (4.5.2.1)$$

$$\text{for} \qquad \frac{r}{\beta} \le x < \frac{r + 1}{\beta}, \qquad r = 0, \ldots, \beta - 1$$

This equation is an obvious generalization of (4.3.2), which is the particular case of (4.5.2.1) for $\beta = 2$. As we can see, we have sort of two Renyi maps, one in each coordinate. The baker's transformation is a Bernoulli shift and

has Kolmogorov–Sinai entropy $\log_2\beta$ (Schild, 1963). The invariant measure is Lebesgue measure. The density function $\rho(x, y)$ evolves according to the Frobenius–Perron operator U:

$$U\,\rho(x, y) = \rho(B^{-1}(x, y)) = \rho\left(\frac{x + r}{\beta},\ \beta y - r\right)$$

$$\text{for} \qquad \frac{r}{\beta} \le y < \frac{r + 1}{\beta}, \qquad r = 0, \ldots, \beta - 1 \tag{4.5.2.2}$$

This operator is unitary in the Hilbert space $\mathcal{L} = L^2$, the equilibrium distribution function is the constant function $\rho_* = 1$, and the Lebesgue spectrum is the unit circle plus the simple eigenvalue 1. As the baker's transformation B is the natural extension of the Renyi map R, the conclusion we can obtain is the same and we refer the reader to Bohm and Gadella (1989) for details. B acts on the Liouville–Hilbert space $\mathcal{L} = L^2 = L_x^2 \otimes L_y^2$ and a suitable initial biorthonormal system can be constructed from the tensor products of the eigenfunctions of the β-adic Renyi map [cf. (4.5.1.4) and (4.5.1.5)]

$$|\varphi_{nm}\rangle = B_n(x)B_m(y) \qquad \text{and} \qquad \langle\tilde{\varphi}_{nm}| = \tilde{B}_n(x)B_m(y) \tag{4.5.2.3}$$

Using these bases and the same perturbative method as before, the following spectral decomposition can be obtained:

$$U = |f_{00}\rangle\langle\tilde{f}_{00}| + \sum_{\nu=1}^{\infty}\left\{\sum_{r=0}^{\nu}\frac{1}{\beta_\nu}|f_{\nu,r}\rangle\langle\tilde{f}_{\nu,r}| + \sum_{r=0}^{\nu-1}|f_{\nu,r+1}\rangle\langle\tilde{f}_{\nu,r}|\right\} \tag{4.5.2.4}$$

where the vectors $|f_{\nu,r}\rangle$ and $\langle\tilde{f}_{\nu,r}|$ can be obtained from the vectors of (4.5.2.3). As we have said, the Liouville spectrum is the unit circle plus the eigenvalue 1, so in the new spectral decomposition we have found new eigenvalues $1/\beta^\nu < 1$. The vectors φ_{nm} and $\tilde{\varphi}_{nm}$ are linear functionals over the spaces $\Phi_- = L_x^2 \otimes \mathcal{P}_y$ and $\Phi_+ = \mathcal{P}_x \otimes L_y^2$. Furthermore, it can be shown that the vectors $f_{nm} \in \Phi_-^\times$ and $f_{nm} \in \Phi_+^\times$ are also functionals over the same spaces, so the spectral decomposition (4.5.2.4) can be implemented if we use these functional vector spaces. We have enlarged the state space with densities which can be distributions in the y coordinate, in the case of $\Phi^\times$: e.g., if the y distributions are Dirac deltas, we will have a distribution whose support is a set of horizontal straight lines, which we will call a "horizontal Dirac comb." In the case of $\Phi_+^\times$ we must change the y by the x, and we would have, e.g., "vertical Dirac's combs." Now, *mutatis mutandis*, we can repeat what we have said in (4.5.1.9) to (4.5.1.12), and we will find similar equations for the baker's transformation. The equilibrium distribution in this case will be

$$\rho_* = |f_{00}\rangle \tag{4.5.2.4}$$

where $|f_{00}\rangle \in \Phi_-^\times$.

4.6. Appendix 4A. Rigged Hilbert Spaces (Ballentine, 1990; Bohm and Gadella, 198

As is well known, all linear spaces of the same dimension are isomorphic if the dimension is finite. This is not the case if the dimension is infinite. In fact, let us consider the infinite sequence of the vectors of a basis of an infinite-dimension vector space

$$\{\Phi_n: n = 1, 2, \ldots\} \tag{4.A.1}$$

Let V be the vector space of all finite linear combinations of the vectors of the basis above, namely $\Psi \in V$ if

$$\Psi = \sum_{n=1}^{i} c_n \Phi_n \tag{4.A.2}$$

V is a linear space of infinite dimension, but we will see that we can build other spaces using the basis $\{\Phi_n\}$. For instance, we can add to V the limit points of all the convergent infinite sequences of vectors of V. But, defining different criteria of convergence, we will have different sets of limit points and therefore different vector spaces. The most useful convergence is the convergence in the norm. The sequence $\{\Psi_i\}$ converges in the norm to a limit point χ if

$$\lim_{i \to \infty} \|\chi - \Psi_i\| = 0 \tag{4.A.3}$$

If the sequences $\{\Psi_i\}$ are sequences of vectors of V and they converge in the norm and we add the limit points χ of these sequences to V, then we obtain a larger space $\mathcal{H}$ where we have finite sequences like (4.A.2), but also limit points of infinite sequences. We will say that $\mathcal{H}$ is the closure of V and also that V is dense in $\mathcal{H}$. However, we can use other kinds of convergences, namely other topologies, and we will obtain different spaces. Let us suppose that we choose the sequences such that the coefficients c_n satisfy the condition

$$\sum_{n=1}^{\infty} |c_n|^2 < \infty \tag{4.A.4}$$

Then, adding the corresponding limit points, we obtain again a Hilbert space $\mathcal{H}$ which contains all the sequences which converge in norm; it is also called the completion of V with respect to the topology of the norm. However, we can also consider an infinite-dimensional linear space Ξ of all, either finite

or infinite, linear combinations of the basis $\{\Phi_n\}$, namely all the linear combinations $\xi = \Sigma_n c_n\Phi_n$ with no limitations imposed on the coefficients c_n. Of course we cannot define a norm in such a space, but we now have three infinite-dimensional linear spaces such that

$$V \subset \mathcal{H} \subset \Xi \tag{4.A.5}$$

Let us define the inner product

$$(f, h) = \sum_n b_n^* c_n \tag{4.A.6}$$

Then $\mathcal{H}$ is the space of the vectors $h = \Sigma_n c_n\Phi_n$ such that $(h, h) = \Sigma |c_n|^2 < \infty$. Let us now define the conjugated space of $\mathcal{H}$: $\mathcal{H}^\times \subset \Xi$ of all linear functionals over $\mathcal{H}$, namely vectors $f = \Sigma_n b_n\Phi_n$ such that the inner product

$$f[h] = (f, h) = \sum_n b_n^* c_n \tag{4.A.7}$$

is convergent for all $h \in \mathcal{H}$. The convergence of this inner product is a consequence of the Schwarz inequality

$$|(f, h)|^2 \leq (h, h)(f, f) \tag{4.A.8}$$

so (4.A.7) converges if $(f, f) = \Sigma_n |b_n|^2$ converges and, therefore, $f \in \mathcal{H}$, so $\mathcal{H} = \mathcal{H}^\times$. (We can in addition define the antilinear space as $f[h] = (h, f)$ *i.e.*, bras are linear functionals and kets can be considered as antilinear functionals.) Let us now define a new space Ω as the space of all vectors $\omega = \Sigma_n u_n\Phi_n$, endowed with coefficients c_n such that they satisfy the following set of infinite conditions:

$$\sum_n |u_n|^2 n^m < \infty, \qquad m = 1, 2, 3, \ldots \tag{4.A.9}$$

Obviously $\Omega \subset \mathcal{H}$. Let us now find the conjugate space of Ω, $\Omega^\times \subset \Xi$, namely the space of convergent linear continuous functionals over Ω. These functionals read $\sigma = \Sigma_n v_n^*\Phi_n$ and they are such that

$$\sigma[\omega] = (\sigma, \omega) = \sum_n v_n^* u_n \tag{4.A.10}$$

is convergent for all $\omega \in \Omega$. Therefore

$$\sum_n |v_n|^2 n^{-m} < \infty, \qquad m = 1, 2, 3, \ldots \tag{4.A.11}$$

In fact, according to Schwarz's lemma, we have

$$\left|\sum_n v_n^* n^{-m/2} u_n n^{m/2}\right|^2 < \left(\sum_n |v_n|^2 n^{-m}\right)\left(\sum_n |u_n|^2 n^m\right) < +\infty \quad (4.A.12)$$

and the rhs is convergent if (4.A.9) and (4.A.11) are fulfilled. As it is obvious that $V^\times = \Xi$, we now have the following set of infinite-dimensional spaces:

$$V \subset \Omega \subset \mathcal{H} = \mathcal{H}^\times \subset \Omega^\times \subset V^\times = \Xi \quad (4.A.13)$$

Any triplet

$$\Omega \subset \mathcal{H} \subset \Omega^\times \quad (4.A.14)$$

such as those of (4.A.12) and others which may be obtained [*e.g.*, if we change the condition $m = 1, 2, \ldots$ of (4.A.9) by $1 \leq m \leq M$, for some $M \in N$] are called Gel'fand triplets. Ω is known as the test space and $\Omega^\times$ as the rigged space. Mathematically, it is convenient that the test space should be a nuclear space. Heuristically speaking, nuclear (barreled) spaces are the infinite-dimensional spaces endowed with the largest number of properties of finite-dimensional spaces, one of them being discrete spectral decomposition. Precisely, nuclear spaces are spaces obtained, so to speak, as the union of an infinite sequence of spaces of finite dimension. Since space Ω is, from our point of view, the space of operators corresponding to real measurement apparatus, and since these devices make only a finite (so less than discrete) number of measurements, logically Ω must be a nuclear space. In fact, even if physical devices make a finite number of measurements, we can conceive that these numbers grow with the progress of technology. Then an infinite, but discrete, number of measurements would correspond to the limit of an infinitely long period of technological progress. A finite number of measurements will correspond to a test space of a finite number of dimensions. Then the test space corresponding to the limit of technological progress will be a nuclear space, since this space is the limit of a sequence of finite-dimensional spaces.

For example, a measurement device makes n measurements, which can define n points on a curve, which can be interpolated by a polynomial of degree n. The space of polynomials of degree n will be the test space corresponding to this device. In the limit of technological progress, the test space will be the space $\mathcal{P}$ of polynomials of any degree, in fact a nuclear space. Generally speaking, by choosing different nuclear test function spaces, we can also choose the physical properties of our measuring devices. In finite-dimensional vector spaces the eigenvalue problem, for every self-adjoint linear operator A, can be solved in a unique way. Namely we can find a unique spectrum $\{a_i\}$ and an orthonormal basis $\{\Psi_n\}$ such that $A\Psi_n = a_n \Psi_n$. This is not so for infinite-dimensional

linear spaces, since the spectrum depends on the rigging we use; nevertheless one can demonstrate:

Gel'fand–Maurin Theorem 4.A.1. If A is a self adjoint operator in $\mathcal{H}$, there is always a complete set of eigenvectors of A in some rigged Hilbert space $\Omega^{\times}$.

Let us give some very important examples:

(i) Let Ξ be the space of functions $f(x)$ of one real variable x and let $A = P = -i\, d/dx$ be the self-adjoint momentum operator in $\mathcal{H} = L^2$. The eigenvectors of P are the plane waves e^{ikx} which do not belong to L^2, since they do not have finite norm. Nevertheless they can be considered as functionals over a convenient space test function Φ, since

$$e^{ikx}[f] = (e^{ikx}, f(x)) = \int_{-\infty}^{+\infty} e^{-ikx} f(x)\, dx \approx f(k) \tag{4.A.15}$$

where $\widehat{f(k)}$ is the Fourier transform of $f(x)$ and Φ is any subspace of $\mathcal{H}$ such that (4.A.15) is convergent. Then we have the Gel'fand triplet $\Phi \subset \mathcal{H} \subset \Phi^{\times}$ and $e^{ikx} \in \Phi^{\times}$.

(ii) Let Ξ be as in the example above and $A = Q = x$ be the position self-adjoint operator in $\mathcal{H} = L^2$. The eigenvectors of Q are the Dirac deltas $\delta(x - y)$, since $Q\delta(x - y) = y\delta(x - y)$; these distributions do not belong to L^2 since they are not even functions. Nevertheless they can be considered as functionals over a convenient space of test functions Φ, since we can rigorously define these deltas as the functionals

$$\delta_y[f(x)] = f(y) \tag{4.A.16}$$

where $f(x)$ is any function of Φ. Usually physicists write this last equation as

$$\delta_y[f(x)] = (\delta(x - y), f(x)) = \int_{-\infty}^{+\infty} \delta(x - y) f(x)\, dx = f(y) \tag{4.A.17}$$

even if the integral in this last equation does not have a rigorous definition Usually Φ is the set of functions with nice properties, *e.g.*, continuous, derivable, with compact support, etc. Then $\delta_y \in \Phi^{\times}$. These examples show that usual operators do not have their eigenvalues in $\mathcal{H}$, but in properly chosen rigged Hilbert spaces.

5. THE QUANTUM EVOLUTION

As the laws of quantum evolution are well known (Ballentine, 1990; Messiah, 1962; Roman, 1965), in this section we will see the use of the no-graining and coarse-graining techniques in quantum mechanics.

5.1. The Case of Discrete Spectra

Let us begin by simply making a heuristic calculation. Let $\mathcal{H}$ be the quantum Hilbert space. Let $\{|i\rangle\}$ be an energy eigenbasis of this Hilbert space, where i is a discrete index. The quantum Liouville space is $\mathcal{L} = \mathcal{H} \times \mathcal{H}$, and a generic density matrix reads

$$\rho = \sum_{i,j} \rho_{ij} |i\rangle\langle j| \tag{5.1.1}$$

where since $\rho = \rho^\dagger$, one has $\rho_{ij} = \rho_{ji}^*$. Also, $\rho_{ii} \geq 0$. Let O be a self-adjoint operator; it reads

$$O = \sum_{i,j} O_{ij} |i\rangle\langle j| \tag{5.1.2}$$

where $O_{ij} = O_{ji}^*$. The mean value of operator O in the quantum state ρ is

$$\langle O \rangle_\rho = (\rho|O) = \mathrm{tr}(\rho O) = \sum_{ij} \rho_{ij} O_{ji} \tag{5.1.3}$$

As $|i\rangle$ is an energy eigenstate, we have

$$H|i\rangle = \omega_i |i\rangle \tag{5.1.4}$$

where ω_i is the energy of state $|i\rangle$. The time evolution of this eigenstate reads

$$|i(t)\rangle = e^{-i\omega_i t} |i\rangle \tag{5.1.5}$$

Therefore the time evolution of ρ is

$$\begin{aligned} \rho(t) &= \sum_{ij} \rho_{ij} |i(t)\rangle\langle j(t)| = \sum_{ij} \rho_{ij} e^{i(\omega_i - \omega_j)t} |i\rangle\langle j| \\ &= \sum_i \rho_{ii} |i\rangle\langle i| + \sum_{i \neq j} \rho_{ij} e^{i(\omega_i - \omega_j)t} |i\rangle\langle j| \end{aligned} \tag{5.1.6}$$

Then the time evolution of the mean value of equation (5.1.3) is

$$\langle O \rangle_{\rho(t)} = (\rho(t)|O) = \sum_i \rho_{ii} O_{ii} + \sum_{i \neq j} \rho_{ij} e^{i(\omega_i - \omega_j)t} O_{ji} \tag{5.1.7}$$

Now let us suppose that the steps of the spectrum are so small and the function under the second summation on the rhs of the last equation is so nice (precisely L^1) that this summation can be approximated by an integral. Therefore, if the function is nice enough, from the Riemann–Lebesgue theorem we have

$$\lim_{t \to \infty} (\rho(t)|O) = \sum_i \rho_{ii} O_{ii} = (\rho_*|O) \tag{5.1.8}$$

where we have defined an equilibrium density matrix $\rho_{*ij} = \rho_{ii} \delta_{ij}$. This

equation would be the quantum equivalent of the classical equation (4.3.3) for mixing systems (even if in the quantum case the vacuum is not unique); it shows that both systems have a similar behavior and it opens the possibility of using classical theorems in the quantum case also. Of course this demonstration is not rigorous, but it motivates the study of continuous spectra in the next section. Using continuous spectra, we will find a rigorous theorem. The role played by continuous spectra in this case is not strange, since evolution operators of mixing systems have this kind of spectrum (Halmos, 1956). In any case, we can also say that the most we can get is a weak limit. Furthermore, the decomposition of the rhs of equation (5.1.6) is not a decomposition within space $\mathcal{H}$, since its second term has a null trace [cf. (2.3.2)].

5.2. The Case of a Continuous Spectrum (Castagnino and Laura, 1997)

In the next subsection we will consider the Friedrichs model, which can be defined in Hilbert space H with a energy eigenbasis $\{|1\rangle, |\omega\rangle\}$, $0 \leq \omega < \infty$, with Hamiltonian operator

$$H = \omega_1 |1\rangle\langle 1| + \int_0^\infty d\omega\, \omega |\omega\rangle\langle\omega| + \lambda \int_0^\infty d\omega\, g(\omega)[|\omega\rangle\langle 1\omega + |1\rangle\langle\omega|] \tag{5.2.1}$$

In this section this formula will only be used as an example of an operator expanded in a continuous spectrum basis, to conclude that the expansion of a generic self-adjoint operator reads

$$O = \int_0^\infty d\omega\, O_\omega |\omega\rangle\langle\omega| + \int\int_0^\infty d\omega\, d\omega'\, O_{\omega\omega'} |\omega\rangle\langle\omega'| \tag{5.2.2}$$

where O_ω, $O_{\omega\omega'}$ are regular functions such that $O_\omega \in R$, $O^*_{\omega'\omega} = O_{\omega\omega'}$, and they belong to the Schwarz class S. Below we will say that $O \in \Phi$, a space with some other properties which we will choose for convenience. Thus functions O_ω, $O_{\omega\omega'}$ will be restricted by this choice. The first term on the rhs of (5.2.2) will be called the singular component of O, since it could be written as the second term, but with a singular coefficient $O_{\omega\omega'} = O_\omega \delta(\omega - \omega')$. The second term will be called the regular term. Let us consider the density matrix at time $t = 0$:

$$\rho(0) = \int\int_0^\infty d\omega d\omega'\, \rho_{\omega\omega'} |\omega\rangle\langle\omega'| \tag{5.2.3}$$

At time t this state reads

$$\rho(t) = \int\int_0^\infty d\omega\, d\omega'\, \rho_{\omega\omega'} |\omega\rangle\langle\omega'| e^{-i(\omega-\omega')t} \tag{5.2.4}$$

If we consider that O can be written as the ρ of equation (5.2.3) but with coefficients $O_\omega \delta(\omega - \omega') + O_{\omega\omega'}$ in (5.2.2), the mean value of operator O in the state $\rho(t)$ is

$$\langle O\rangle_{\rho(t)} = (\rho(t)|O) = \mathrm{tr}(\rho(t)O) \int_0^\infty d\omega\, \rho_{\omega\omega} O_\omega + \int\int_0^\infty d\omega\, d\omega'\, \rho_{\omega\omega'} O_{\omega'\omega} e^{-i(\omega-\omega')t} \tag{5.2.5}$$

Now, if the space $\Phi\rho_{\omega\omega'}$ belongs also to the Schwarz class S, the Riemann–Lebesgue theorem can be used, because $\rho_{\omega\omega'} O_{\omega'\omega} \in L^1$, and we have

$$\lim_{t\to\infty} \langle O\rangle_{\rho(t)} = \int_0^\infty d\omega\, O_\omega \rho_{\omega\omega} \tag{5.2.6}$$

As this equation is valid for any operator $O \in \Phi$, we can try to find a density matrix ρ_* such that

$$\lim_{t\to\infty} \langle O\rangle_{\rho(t)} = \lim_{t\to\infty} (\rho t)|O) = (\rho_*|\, O) \tag{5.2.7}$$

It is easy to see that the density matrix ρ_* cannot be found if $\rho_{\omega\omega'}$ is a regular function of variable ω, ω'; *i.e.*, from (5.2.5) we see that to obtain this result it is necessary that $\rho_{\omega\omega'} = 0$, $\omega \neq \omega'$, and $\rho_{\omega\omega} \neq 0$, but we cannot write $\rho_{\omega\omega'} = \rho_\omega \delta(\omega - \omega')$, because in this case the ρ is not regular (Antoniou *et al.*, 1995). Then we are forced to consider states with diagonal singularities, that is, with the same operator pathology. So we are forced to introduce singular components in the density matrix, which means the $\rho_{\omega,\omega'}$ of (5.2.3) cannot be regular and it must read something like $\rho_\omega \delta(\omega - \omega') + \rho_{\omega,\omega'}$. But now if we try to find the mean value (5.2.5), the $O_\omega \delta(\omega - \omega')$ term and the $\rho_\omega \delta(\omega - \omega')$ term produce the result

$$\int\int_0^\infty O_\omega \delta(\omega - \omega') \rho_{\omega'} \delta(\omega - \omega')\, d\omega\, d\omega' = \int_0^\infty O_\omega \rho_\omega \delta(0)\, d\omega \to \infty$$

which is divergent. Therefore to have a formalism free from these problems we are forced to start afresh, and to consider the operators O to be defined by the regular functions O_ω, $O_{\omega\omega'}$ and the state functions, ρ to be the matrices of rigged space $\Phi^\times$ defined as the linear operators on space Φ but with some extra properties, precisely defined by two regular Schwarz functions ρ_ω, $\rho_{\omega,\omega'}$). Then we have

$$(\rho|0) = \int_0^\infty d\omega\ \rho_\omega O_\omega + \int\int_0^\infty d\omega\ d\omega' \rho\omega'\omega O_{\omega\omega'} \tag{5.2.8}$$

where $\rho_\omega \in R^+$, $\rho^*_{\omega'\omega} = \rho_{g\nu\omega'}$. So we are forced to introduce a singular component ρ_ω also in the density matrices. Now ρ_* may be found; it is the functional of space Φ^x with $\rho_\omega \neq 0$, $\rho_{\omega\omega'} = 0$. The consistency of this method is proved by the logical physical results of Castagnino and Laura (1997). Equation(5.2.7) can now be considered as the rigorous quantum equivalent of the classical equation (4.3.3). We can call the weak limit of this equation the "quantum mixing" property and state the following.

Theorem 5.2.1. Quantum systems with continuous spectra are endowed with the quantum mixing property (provided we use the formalism based on (5.2.8). Equation (5.2.7) can also be considered a proof of a weak decoherence in quantum systems. This would be the no-graining conclusion. However, we would like to have a strong decoherence. Then we can follow two methods. We can either use coarse-graining, a well-known technique discussed in Hu *et al.* (1993) and Caldeira and Leggett (1995), or we can use extended dynamics. In this case, in order to obtain a strong limit from the weak limit of (5.2.8), we must give a precise sense to all terms on the rhs of equation (5.2.5), rigging the Hilbert space $\mathcal{H}$ in such a way that all the mathematical characters are well defined. So, working with the functional of space $\Phi^\times$, we can write the strong limit

$$\lim_{t\to\infty} \rho(t) = \rho_* \tag{5.2.9}$$

which corresponds to the classical strong limit (4.4.3). Now we would like to obtain not just a limit, but the time-irreversible evolution of $\rho(t)$, which yields the limit (5.8.9). Unfortunately our ability to work with continuous spectra is very limited (Gadella and Rudin, 1996), so we are forced to mix various techniques, as we will see in the next subsections.

5.3. The Friedrichs model

5.3.1. The General Formalism

We believe that the well-known Friedrichs model is the best quantum example to fix the ideas. In this example we have a free (naked) stable quantum state $|1\rangle$ (which is postulated to be real: $K|1\rangle = |1\rangle$) which becomes unstable when coupled to a continuous field $|\omega\rangle$. The stable state may be considered as a simplified model of an atom in an excited stable state, becoming unstable if coupled to an electromagnetic field, which, in the model, is represented by a continuous field. Thus, let us consider a Hilbert space $\mathcal{H}$ with a basis $\{|1\rangle, |\omega\rangle\}$, $0 \leq \omega < \infty$, such that

$$\langle 1|1\rangle = 1, \quad \langle 1|\omega\rangle = 0 \quad \langle\omega|\omega\rangle = \delta(\omega - \omega') \tag{5.3.1}$$

$$1 = |1\rangle\langle 1| + \int_0^\infty |\omega\rangle\langle\omega|\, d\omega \tag{5.3.2}$$

and a system with free Hamiltonian

$$H_0 = \omega_1 |1\rangle\langle 1| + \int_0^\infty |\omega\rangle\langle\omega|\, d\omega \tag{5.3.3}$$

and $\omega_1 > 0$. Therefore the spectrum of H_0 is R_+ with a degeneracy at ω_1. Let the interaction Hamiltonian be

$$H_1 = \lambda \int_0^\infty g(\omega)(|1\rangle\langle\omega| + |\omega\rangle\langle 1|)\, d\omega \tag{5.3.4}$$

where $g(\omega)$ is an interaction function endowed with all sorts of nice properties; it is analytic, well behaved at $\omega \to +\infty$, and so on. The total Hamiltonian is

$$H = H_0 + H_I \tag{5.3.5}$$

This Hamiltonian can be diagonalized using standard techniques. We then obtain

$$H = \int_0^\infty \omega |\omega, \overset{\text{adv}}{\text{ret}}\rangle\langle\omega, \overset{\text{adv}}{\text{ret}}|\, d\omega \tag{5.3.6}$$

where $\{|\omega, \overset{\text{adv}}{\text{ret}}\rangle\}$ are the usual retarded or advanced bases (Bohm, 1979). We can see, comparing (5.3.3) and (5.3.6), that the interaction has erased the discrete component of the spectrum. In fact, state $|1\rangle$ has became unstable and now it is simply a pole in the corresponding S-matrix. In any case, by using (5.3.6), we can compute the time evolution of any state, for example the state $|1\rangle$ at $t = 0$. As we have just said, the state $|1\rangle$ of the free system (5.3.2) is transformed into an unstable state by the interaction (5.3.4) in such a way that the survival probability $P(t) = |\langle 1|1(t)\rangle|^2$ vanishes when $t \to +\infty$. It is also known that $P(t)$ has a vanishing derivative when $t = 0$ (Zeno effect), then a decreasing exponential behavior, and finally oscillates for large t (the Khalfin effect) (Fig. 3) (Sudarsham *et al.*, 1978).

5.3.2. Hilbert and Rigged Spaces

Let us put aside for the moment the problem of the unification of dynamics and thermodynamics, and let us introduce some equations related to the problem of time asymmetry, as stated in Section 1.1. As we will see, the previous equations are all we need to define the quantum arrow of time

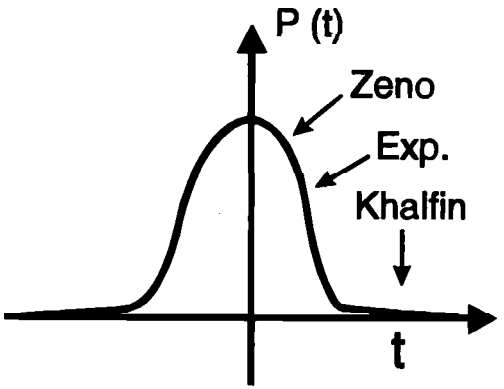

Fig. 3. The $P(t)$ graph, showing the Zeno effect, the exponential behavior, and the Khalfin effect.

according to the coarse-graining school. As this school always works within Hilbert space $\mathcal{H}$, the following property holds:

$$K: \quad \mathcal{H} \to \mathcal{H} \tag{5.3.7}$$

On the other hand, for the extended dynamics school we need a richer structure. In fact, we need to define two convenient subspaces $\phi_\pm \subset \mathcal{H}$. To do so, let us consider a vector $|\varphi\rangle \in \mathcal{H}$ and its components $\langle\omega|\varphi\rangle$, and let us promote the real energy ω to a complex variable z. Then

$$|\varphi\rangle \in \phi_\pm \qquad \text{iif} \qquad \langle z|\varphi\rangle \in \theta(H^2_\pm \cap S) \tag{5.3.8}$$

where $H^2_\pm$ are the Hardy classes from above and below, respectively (cf. Section 5.5) and S is the Schwarz class of functions. It can be proved that $\phi_\pm$ are nuclear spaces. Then we can define two Gel'fand triplets:

$$\phi_- \subset \mathcal{H} \subset \phi_-^\times \tag{5.3.9}$$

$$\phi_+ \subset \mathcal{H} \subset \phi_+^\times \tag{5.3.10}$$

5.3.3. The Rigged Hilbert Space Formalism

Using analytical continuation techniques Castagnino *et al.*, 1996; Sudarsham *et al.*, 1978; Antoniou and Prigogine, 1979; Bohm, 1979), essentially just Cauchy's theorem, it is possible to obtain a new spectral decomposition of the identity operator 1 and the Hamiltonian operator H as

$$1 = |z_1, -\rangle\langle z_1, +| + \int_\Gamma |z, -\rangle\langle z, +|\, dz \tag{5.3.11}$$

$$H = z_1|z_1, -\rangle\langle z_1, +| + \int_\Gamma z|z, -\rangle\langle z, +|\, dz \tag{5.3.12}$$

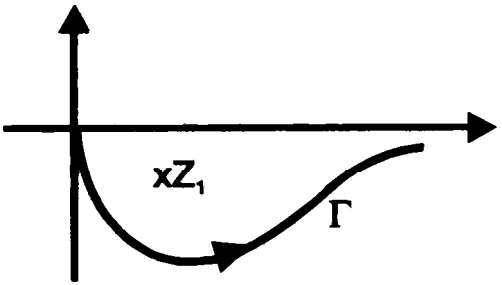

Fig. 4. The Γ curve.

where $|z_1, -\rangle \in \phi_+^\times$, $|z_1, +\rangle \in \phi_-^\times$, z_1 is a complex rot of the equation $\alpha(z) = 0$, where

$$\alpha(z) = z - \omega_1 + \lambda \int_\Gamma \frac{g^*(z^*)g(z)}{z - \omega_1}\, dz \tag{5.3.13}$$

and Γ is any curve going from the origin of the complex plane to positive infinity on the real axis and passing below z_1 (Fig. 4). The first terms of equations (5.3.11) and (5.3.12) are produced by the residues of the poles corresponding to the roots located at the zeros of equation $\alpha(z) = 0$, or, in other words, the poles of the *S*-matrix. We can see that the discrete component of the spectrum, which we have lost in (5.3.6), reappears in (5.3.12) in the form of a matrix of the rigged Hilbert space. There are several possibilities for choosing the curve Γ, which vary from author to author: (i) A generic curve Γ. (ii) The curve Γ' of Fig. 5, used in such a way that, as the vertical paths of the curve are mutually canceled, we are mostly integrating on the real positive axis. (iii) To use the negative real axis as integration path. (iv) (Defining a tilde operation as

$$\int_\Gamma f(z)g(z)\, dz = \int_0^\infty \tilde{f}(x)g(x)\, dx$$

for all $g(x)$ in the test function space. In this case the complex integral

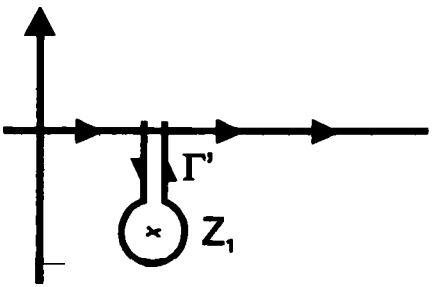

Fig. 5. The Γ' curve.

formally becomes a real one. If we use this last method and forget the tilde, (5.3.12) reads

$$H = z_1|z_1, -\rangle\langle z_1, +| + \int_0^\infty \omega|\omega, -\rangle\langle\omega, +|\, d\omega \qquad (5.3.12')$$

So we have built a basis $\{|z_1, -\rangle, |\omega, -\rangle\}$ for the space ϕ_- [see Antoniou and Prigogine (1979) for details]. These vectors read

$$|z_1 -\rangle = \langle 1|z_1, -\rangle\,(|1\rangle + \lambda \int_0^\infty d\omega \frac{g(\omega)}{[z_1 - \omega]_-}|\omega\rangle) \qquad (5.3.12'')$$

$$|\omega, -\rangle = |\omega\rangle + \frac{\lambda g(\omega)}{\alpha(\omega)}(|1\rangle + \lambda \int_0^\infty d\omega' \frac{g(\omega)}{\omega - \omega' + i\varepsilon}|\omega'\rangle) \qquad (5.3.12''')$$

where the subscript minus sign in the denominator of the integral in (5.3.12″) means that the curve Γ' must be used for the integration. Now we have two spectra to compare: (5.3.5) and (5.3.12″). The main difference is that (5.3.5) is structurally unstable when $\lambda \to 0$, while (5.3.12″) is stable. In fact, an algorithm is called structurally stable if it changes only slightly under small changes of the coefficients. When $\lambda = 0$, the spectral decomposition of H is (5.3.3). If λ is small, a small change of λ that makes $\lambda = 0$ produces a big change in the usual decomposition, which goes from (5.3.5), with no discrete term, to (5.3.3) with the discrete term $\omega_1|1\rangle\langle 1|$. The sudden vanishing of this term when $\lambda \to 0$ is a catastrophe (specifically a Poincaré catastrophe) which leads to many problems if we try to perform an expansion around $\lambda = 0$ in Hilbert (or Liouville) space. On the contrary, (5.3.12″) is stable, since it has the term $z_1|z_1, -\rangle\langle z_1, +|$ which goes to $\omega_1|\rangle\langle 1|$ when $\lambda \to 0$, as we shall see. From (5.3.12) it can be seen that $|z_1, -\rangle$ and $\langle z_1, +|$ are respectively the left and right eigenvectors of H corresponding to the eigenvalue z_1. It can be proved that

$$\begin{aligned} \langle z_1, +|z_1, -\rangle &= 1 \\ \langle z, +|z, -\rangle &= 0 \\ \langle z, +|z_1, -\rangle &= 0 \\ \langle z, +|z, -'\rangle &= \delta(z - z') \end{aligned} \qquad (5.3.14)$$

It can also be proved that

$$\begin{aligned} \langle z_1, -|z_1, -\rangle &= 0 \\ \langle z_1, +|z_1, +\rangle &= 0 \end{aligned} \qquad (5.3.14')$$

namely there are nonnull vectors of zero "norm" in spaces $\phi_+^\times$ and $\phi_-^\times$

(Castagnino *et al.*, 1997). Let us call $z_1 = \beta_1 - (i/2)\gamma_1$, where $\gamma_1 > 0$. Then from (5.3.12′) we can obtain the time evolution of $|z_1(t), -\rangle$ and $|z_1(t), +\rangle$:

$$|z_1(t), -\rangle = e^{-ix_1 t}|z_1(0), -\rangle = e^{-i\beta_1 t}e^{-(\tau_1/2)t}|z_1(0), -\rangle$$

$$|z(t), +\rangle = e^{-iz_1^* t}|z_1(0), +\rangle = e^{-i\beta_1 t}\, e^{(\tau 1/2)t}|z_1(0), +\rangle \qquad (5.3.15)$$

These equations show that $|z_1(t), -\rangle$ is a decaying state whereas $|z_1(t), +\rangle$ is a growing state. It can be proved that (Castagnino *et al.*, 1996)

$$K|z_1, -\rangle = |z_1, +\rangle \qquad (5.3.16)$$

$$K|z^1, +\rangle = |z_1, -\rangle$$

which is natural, since growing states must be transformed into decaying states by the time-inversion operator and vice versa. It can also be proved that

$$K: \quad \phi_-^{\times} \to \phi_+^{\times} \qquad (5.3.17)$$

$$K: \quad \phi^{\times} \to \phi_-^{\times}$$

The following limits are valid [cf. (5.3.12″)]:

$$\lim_{\lambda\to\infty} |z_1, -\rangle = \lim_{\lambda\to\infty} |z_1, +\rangle = |1\rangle \qquad (5.3.18)$$

Therefore $|z_1, -\rangle$ and $|z_1, +\rangle$ can be considered as versions of the unstable state $|1\rangle$ in spaces $\phi_+^{\times}$ and $\phi_-^{\times}$. In fact, the difference between these vectors and $|1\rangle$ is $O(\lambda)$, since when $\lambda = 0$, the interaction disappears. Let us remember that the survival probability of state $|1(t)\rangle$ was

$$P(t) = |\langle 1|1(t)\rangle|^2 = \langle 1|1(t)\rangle\langle 1(t)|1\rangle \qquad (5.3.19)$$

$P(t)$ shows the initial Zeno effect behavior, then an exponential behavior, and finally the oscillatory Khalfin effect behavior. If we make the substitution $|1(t)\rangle \to |z_1(t), -\rangle$, we obtain

$$P(t) \to P'(t) = \langle 1|z_1(t), -\rangle\langle z_1(t), -|\rangle = e^{-\gamma_1 t} \qquad (5.3.20)$$

and only the exponential behavior remains. Thus the physical nature of the state $|z_1, -\rangle$ would be that of a decaying unstable ideal state, where we have eliminated the Zeno and Khalfin effects, since these effects are contained in the last term on the rhs of (5.3.12″) (also called "the background"). Specifically, the three effects are mixed if we use the time evolution based on (5.3.6), but Zeno and Khalfin effects can be detached from the exponential behavior if we use the evolution based on (5.3.12). Equation (5.3.20) also shows that γ_1^{-1} is the mean lifetime of the unstable states. Extended dynamics can be thought of as an approximation to real states which eliminates the unimportant

Zeno and Khalfin effects. The Zeno effect is unimportant because it takes place at $t = 0$ whereas we are usually interested in the phenomena at $t \to 0$. However, as we said in the introduction, it is not completely clear if these ideal exponential states are merely mathematically useful, effective states or real physical states. Furthermore, the extended dynamics school needs to work with rigged Hilbert spaces $\phi^{\times}_{-}$ and $\phi^{\times}_{+}$ to solve the problem of the arrow of time, as we shall see in the next section. The Friedrichs model is only an example, but its rigged Hilbert space structure can be found in every scattering process (Bohm, 1979). Therefore, even if we base our reasoning only on this model, what we explain below is rather general.

5.3.4. Mixed States

Let us now introduce the arguments of the next subsection by writing the evolution equations of mixed states in our model. A mixed arbitrary state at time $t = 0$ can be expanded in a basis $\{|\omega, \text{ret}\rangle\}$, as

$$\rho = \int \int_0^\infty \rho_{\omega\omega'} |\omega, \text{ret}\rangle\langle\omega', \text{ret}\, d\omega \tag{5.3.21}$$

and its time evolution reads

$$\rho(t) = \int\int_0^\infty \rho_{\omega\omega'}\, e^{-i(\omega-\omega')t} |\omega, \text{ret}\rangle\langle\omega', \text{ret}\, d\omega \tag{5.3.22}$$

We could just as well use the advanced basis, but this is all we can say in space $\mathcal{H}$. In space $\phi^{\times}_{+}$, however, we can use the basis $\{|z_1, -\rangle, |\omega, -\rangle\}$ [introduced in (5.3.12′)] and expand ρ as

$$\begin{aligned} \rho = \rho_{11} |z_1, -\rangle\langle z_1, -| &+ \int_0^\infty (\rho_{1\omega} |z_1, -\rangle\langle\omega, -| \\ &+ \rho_{\omega 1} |\omega, -\rangle\langle z_1, -|)\, d\omega \\ &+ \int\int_0^\infty \tilde{\rho}_{\omega\omega'} |\omega, -\rangle\langle\omega, -|\, d\omega \end{aligned} \tag{5.3.23}$$

and its time evolution reads

$$\rho(t) = \rho_*(t) + e^{-\gamma_1 t/2} \rho_1(t) + e^{-\gamma_1 t}\, \rho_2(t) \tag{5.3.24}$$

where

$$\rho_*(t) = \int\int_0^\infty \tilde{\rho}_{\omega\omega'}\, e^{-i(\omega-\omega')t} |\omega, -\rangle\langle\omega', -|\, d\omega \tag{5.3.25}$$

and

$$\rho_1(t) = \int_0^\infty (\rho_{1\omega} e^{-i(\theta_1-\omega)t} |z_1, -\rangle\langle\omega, -| + \rho_{\omega 1} e^{-i(\omega-\beta_1)t} |\omega, -\rangle\langle z_1, -|)\, d\omega \tag{5.3.26}$$

$$\rho_2(t) = \rho_{11} |z_1, -\rangle\langle z_1, -| \tag{5.3.27}$$

Now since $\gamma_1 > 0$, $\rho_1(t)$ oscillates, and $\rho_2(t)$ is time invariant, we have

$$\lim_{t\to\infty} (\rho(t) - \rho_*(t)) = 0 \tag{5.3.28}$$

which seems very close to the strong limit we are looking for. The only problem is that $\rho_*(t)$ is not a dynamical equilibrium state, since it oscillates. Nevertheless from the thermodynamic point of view $\rho_*(t)$ is a thermodynamic equilibrium state, since it has a constant (and maximum) Gibbs entropy. In fact, it is evident (from the quantum version of Theorem 3.6.2) that the Gibbs entropy is constant for the time evolution (5.3.22), and therefore it is constant for the time evolution (5.3.25), which is similar. It is clear that that is all we can ask of the model, since the field cannot go to dynamical equilibrium because the modes of the field are decoupled. Therefore, from the thermodynamic point of view, (5.3.28) reads

$$\lim_{t\to\infty} \rho(t) = \rho_* \tag{5.3.29}$$

and it is the strong limit we are looking for. As in the case of equation (4.5.1.11), this limit belongs to the corresponding rigged Hilbert space. What was the miracle which allowed us to pass from the oscillatory evolution (5.3.22) with no limit to the partially damped evolution (5.3.24) with a thermodynamic limit? It was the fact that (5.3.22) is valid in space $\mathscr{L} = \mathscr{H} \times \mathscr{H}$, while (5.3.24) is valid in space $\Phi_+^\times = \phi_+^\times \times \phi_+^\times$, so that (5.3.28) is actually a functional equation which may be interpreted as

$$\lim_{t\to\infty} (\rho(t)|O_-) = (\rho_*|O_- \tag{5.3.30}$$

where O_- is an operator of the test operator space $\Phi_+ = \phi_+ \times \phi_+$, the space of chosen measurement operators. Therefore this miracle happens simply because we have chosen a convenient test space for our physical measurement apparatus. Equation (5.3.30) is a weak limit, similar to that of mixing classical states, and, since our model has a continuous spectrum, a consequence of Theorem 5.2.1. We will continue this line of reasoning once the more complete example in the next subsection has been introduced. In the practical case which will be studied in Section 8, the field $|\omega\rangle$ will be the thermic radiation field within the universe, which we can consider as being thermalized from

its beginning by interactions other than those of equation (5.3.4); it can then be classically chosen as a Boltzmann thermic distribution function:

$$\rho_* = ZT^{-1/2}e^{-\omega/T} \tag{5.3.31}$$

where T is the temperature and Z a normalization function, and the damping terms are produced by nuclear reaction phenomena within the stars, and γ_1^{-1} is the characteristic time of these nuclear reactions. We will use this model in Section 8.

5.4. The Friedrichs Model for Many Oscillators

We will now introduce a not very realistic physical model, which, however, is the simplest one for our purpose. Let us consider an infinite set (or a great number) of uncoupled harmonic oscillators, labeled by ω, with Hamiltonian

$$H_\omega = \omega(a_\omega^\dagger a_\omega + \tfrac{1}{2}) \tag{5.4.1}$$

where $a_\omega^\dagger$ and a_ω are respectively the creation and annihilation operators of the harmonic oscillator. The total Hamiltonian reads

$$H = \int_0^{\text{in}} H_\omega \, d\omega \tag{5.4.2}$$

H_ω can also be written

$$H_\omega = \sum_n H_\omega^{(n)} \tag{5.4.3}$$

where

$$H_\omega^{(n)} = \omega(n|n, \omega\rangle\langle n, \omega| + \tfrac{1}{2}) \tag{5.4.4}$$

$|n, \omega\rangle$ is the ω-oscillator in the n excited state ($n = 0$ corresponds to the ground state). Let us suppose that each of these states is coupled with a field represented by a set of infinite states $|n,\omega,w\rangle$ in such a way that now the coupled $H_\omega^{(n)}$ reads

$$H_\omega^{(n)} = \omega(n|n, \omega\rangle\langle n, \omega| + \tfrac{1}{2}) + \int_0^\infty dw \; |n, \omega, w\rangle\langle n, \omega, w|$$
$$+ \lambda \int_0^\infty dw \; g_{n,\omega}(w)(|n, \omega\rangle\langle n, \omega, w| + |n, \omega, w\rangle\langle n, \omega|) \tag{5.4.5}$$

where λ is a coupling constant and $g_{n,\omega}(w)$ is a coupling function which necessarily has the property $g_{0,\,\omega}(w) = 0$, since the ground state of each

oscillator is stable; it is therefore not coupled with the corresponding field that would produce its instability. Thus we have constructed a model which can be considered as an infinite repetition of the Friedrichs model of the last subsection. In this nonrealistic model the instability of all the states, with the exception of the ground states, is obtained by coupling a field to each oscillation mode. It is, in fact, not a particularly economical method, but serves our purpose, which here is simply to find the laws of unstable evolutions. Now, using the procedure of the previous subsection in each Friedrichs model of each mode, we can diagonalize each operator $H_\omega^{(n)}$ to obtain

$$H_\omega^{(n)} = z_\omega^{(n)}(|n, \omega, +\rangle\langle n, \omega, -| + \tfrac{1}{2}) \tag{5.4.6}$$

where, for simplicity, we have omitted the field term, where $z_\omega^{(0)} = \frac{1}{2}\omega$, since the ground states are not perturbed, and we have put the factor n inside $z_\omega^{(n)}$, and Im $z_\omega^{(n)} < 0$ for $n \neq 0$. If we renormalize and eliminate the 1/2-terms, we obtain

$$H = \int_0^\infty \sum_n z_\omega^{(n)}(|n, \omega, +\rangle\langle n, \omega, -| \, d\omega \tag{5.4.7}$$

Let us now consider a density matrix $\rho = \rho(0) \in \mathscr{L} = \mathscr{H} \times \mathscr{H}$, which can be expanded in the basis $\{|n, \omega, -\rangle\}$ as

$$\rho = \int_0^\infty \sum_n \rho_{n,n';\omega}|n, \omega, -\rangle\langle n, \omega, -| \, d\omega \tag{5.4.8}$$

We will always work with these density matrices below. And herein lies the essential fact. Since $|n, \omega, -\rangle \in \phi_{\text{pl}}^\times$ and, therefore, $\rho \in \phi_+^\times \times \phi_+^\times$, what we have done in choosing the expansion (5.4.8) is to assume that our operator space is $\mathbf{\Phi}_+ = \phi_+ \times \phi_+$ in such a way that now we have the Gel'fand triplet

$$\mathbf{\Phi}_+ \subset \mathscr{L} \subset \mathbf{\Phi}_+^\times \tag{5.4.9}$$

and therefore $\rho \in \mathbf{\Phi}_+^\times = \phi_+^\times \times \phi_+^\times$. In a physical sense, what we are doing is to postulate that our measurement apparatus correspond to operators in $\mathbf{\Phi}_+$. We will discuss this postulate below, but we can immediately see that this is the price we have to pay to get the strong limit we are looking for and the corresponding unstable time evolution. In fact, from (5.3.12) we can obtain the time evolution

$$\begin{aligned}\rho(t) &= e^{-iLt}\rho(0) \\ &= \int_0^\infty \sum_n \rho_{mn,n'\omega} \exp[-i(z_\omega^{(n)} - z^{(\omega')*})t]|n, \omega, -\rangle\langle n', \omega, -| \, d\omega\end{aligned} \tag{5.4.10}$$

but since

$$-i(z_{\omega}^{(n)} - z_{\omega}^{(n')*}) = -i(\beta_{\omega}^{(n)} - \beta_{\omega}^{(n')}) - \frac{i}{2}(\gamma_{\omega}^{(n)} - \gamma_{\omega}^{(n')}) \quad (5.4.11)$$

and $\gamma_{\omega}^{(n)}, \gamma_{\omega}^{(n')} \geq 0$ (only $\gamma_{\omega}^{(0)} = 0$), thus

$$\lim_{t\to\infty} \rho(t) = \int_0^{\infty} \sum_n \rho_{0,0;\omega} |0, \omega, -\rangle\langle 0, \omega, -| = \rho_* \quad (5.4.12)$$

and we have obtained our strong limit, equivalent to (5.2.9). Furthermore, now we have the time evolution to this limit, (5.4.10). To obtain this result, we have used an infinite set of continuous fields, which has been neglected in all the formulas above. Somehow we have "traced away" these fields. But the result will not change from the physical point of view if we include all these fields. The result we have obtained regarding the states of the harmonic oscillator will be the same; these oscillators reach equilibrium, evident in the last equation, but the fields will continue to oscillate and they always will be far from equilibrium (as in the last part of the last subsection). This is not surprising, since these fields have no self-interaction or mutual interaction, and therefore they cannot reach equilibrium. Thus to neglect these fields was merely a useful shorthand with no physical consequences (bearing in mind the caveats of the last part of the last subsection). We must also remember that the quantities of (5.4.12) are just functionals over the space Φ_+. Thus, if we contract this equation with any vector of this space, we will find the weak version of the limit (5.4.12), showing that in this example Theorem 5.2.1 is fulfilled. If we collectively denote by 2γ all the γ's, or 2γ is the inverse of the characteristic lifetime of the system, or if we call 2γ the smaller of them to maintain the leading term only, (5.4.11) reads

$$\rho(t) = \rho_* + \rho_1 e^{-\gamma t} \quad (5.4.13)$$

As usual we have

$$\operatorname{tr} \rho = \operatorname{tr} \rho_* = 1 \quad (5.4.14)$$

since the matrix ρ is the usual one, the matrix ρ_* is an expansion of stable states (5.4.12), and the norm must be conserved (cf. Appendix 5.B). But

$$\operatorname{tr} \rho_1 = 0 \quad (5.4.15)$$

as a consequence of (5.3.14′), showing that this matrix is something like a fluctuation around the equilibrium state. Let us finally observe that in this model we cannot pretend that ρ_* would be the equilibrium state of the canonical ensemble. To obtain that result, obviously we must couple the oscillators among themselves and the model necessarily will be much more complicated. To mimic a canonical ensemble at temperature T in this model, the best we can do is to make the following choice:

$$\rho_{0,0;\omega} = \frac{Z}{T^{3/2}} e^{-\omega/T} \tag{5.4.16}$$

With this choice we have the correct equilibrium distribution and the field produces the irreversible evolution toward this equilibrium. The evolution of $\rho(t)$ then reads

$$\rho(t) = \int_0^\infty \left[\frac{Z}{T^{3/2}} e^{-\omega/T} \rho_*^{(\omega)} + e^{-2\gamma t} f(\omega) \rho_1^{(\omega)} \right] d\omega \tag{5.4.17}$$

where $f(\omega)$ is an arbitrary function, so the conclusions are essentially the same as in the last subsection.

5.5. Appendix 5A. Hardy Class Functions (Bohm and Gadella, 1989)

A complex function $f(\omega)$ on R^+ is a Hardy class function from above (below) if (i) $f(\omega)$ is the boundary value of a function $f(z)$ of the complex variable $z = x + iy$ which is analytic in the half-plane $y > 0$ ($y < 0$); (ii) $\int_{-\infty}^{+\infty} |f(x + iy|^2\, dx < k < \infty$, for all y with $0 < y < \infty$ ($-\infty < y < 0$).

5.6. Appendix 5B. Computation and Conservation of Norm, Trace, and Energy

The trace of a density matrix ρ is

$$tr\ \rho = \int_0^\infty \langle \omega, \text{ret} | \rho | \omega, \text{ret} \rangle\, d\omega \tag{5.B.1}$$

and is invariant under changes of basis. Repeating the procedure we used to go from (5.3.6) to (5.3.12), we can obtain

$$tr\ \rho = \langle z_1, + | \rho | z_1, - \rangle + \int_0^\infty \langle \omega, + | \rho | \omega, - \rangle\, d\omega \tag{5.B.2}$$

As the extended dynamics theory deals with states which vanish when $t \to \infty$, one might be concerned about the conservation of norms, traces, or energy in this theory. There is actually no problem, since we can state the following results. (i) From (5.3.14) and (5.3.14′) we see that unstable density matrices such as $|z_1, -\rangle\langle z_1 -|$ or $|z_1, +\rangle\langle z_1, +|$ have null trace; this is possible since we are not working in Hilbert space. (ii) Using (5.3.15) and (5.3.14′), we can see that the trace of (5.B.2) is conserved as the usual trace of (5.B.1). (iii) If the trace is conserved, the norm of pure states is also conserved. (iv) The mean value of the energy in one of these unstable states such as $|z_1, -\rangle\langle z_1, -|$ reads

$$\langle H\rangle = \langle z_1, -|H|z_1, -\rangle = z_1\langle z_1, -|z_1, -\rangle = 0 \qquad (5.B.3)$$

and therefore the energy of the states vanishes when $t \to \infty$ is zero, creating no problems with energy conservation.

6. COURSE GRAINING AND TRACE. TIME ASYMMETRY

6.1. Coarse Graining

Let us return, for a while, to the classical regime. Usually coarse-graining is based on the fact that the dynamical variables cannot be measured with infinite precision, *i.e.*, there is always some error and, furthermore, we cannot compute with an infinite number of digits. There may also be a fundamental graininess in nature, but this graininess has not yet been found, either theoretically or experimentally. Coarse-graining may be introduced by partitioning the space X into a finite (or discrete) number of cells A_i satisfying

$$\bigcup_i A_i = X, \qquad A_i \cap A_j = \emptyset \qquad \text{if} \quad i \neq j \qquad (6.1.1);$$

This partition is arbitrary, but it must be nontrivial with respect to some measure μ, namely

$$0 < \mu(A_i) < \mu(X) \qquad (6.1.2)$$

for all values of i. For every density ρ within each cell A_i of the partition, we can compute the average of ρ as

$$\langle\rho\rangle_i = \frac{1}{\mu(A_i)} \int_{A_i} \rho(x)\, \mu(dx) \qquad (6.1.3)$$

and the coarse-grained density with respect to the partition is given by

$$\tilde{\rho}(x) = \sum_i \langle\rho\rangle_i\, 1_{A_i}(x) = \left\{\sum_i \frac{1}{\mu(A_i)} |1_{A_i})(1_{A_i}|\right\}|\rho) = \Pi\rho(x) \qquad (6.1.4)$$

where 1_{A_i} is the characteristic function of the cell A_i and Π is the projector defined by the partition. Π is a projector since

$$\begin{aligned}\Pi^2 &= \left[\sum_i \frac{1}{\mu(A_i)} |1_{A_i})(1_{A_i}|\right]\left[\sum_j \frac{1}{\mu(A_j)} |1_{A_j})(1_{A_j}|\right] \\ &= \sum_{ij} \frac{1}{\mu(A_i)} \frac{1}{\mu(A_j)} |1_{A_i})(1_{A_j}|\mu(A_i)\delta_{ij} \\ &= \sum_i \frac{1}{\mu(A_i)} |1_{A_i})(1_{A_i}| = \Pi\end{aligned}$$

Following our reasoning in the introduction, or from Theorem 4.3.1, we can deduce the following result in the case of a finite number partitions (for a discrete number see Lebowitz (1994):

Theorem 6.1.1. If P_t is a ρ_*-mixing Markov operator with a unique stationary density ρ_* and $\{A_i\}$ is a nontrivial partition of the phase space X, then

$$\lim_{t\to\infty} (P_t\rho)^{\sim} = \lim_{t\to\infty} \tilde{\rho}_* \tag{6.1.5}$$

for all initial densities.

Thus we have obtained our coarse-graining strong limit. Now we can consider the transformation

$$P_t\tilde{\rho}(x) = (P_t\rho(x))^{\sim} \tag{6.1.6}$$

From (6.1.5) we see that this transformation has a strong limit (*i.e.*, in the norm) and therefore, according to Theorem 4.4.1, it is exact. Then, by using Theorem 4.4.2, we can conclude the following about the entropy:

Theorem 6.1.2. If P_t is a reversible ρ_*-mixing Markov operator with a unique stationary density ρ_*, and $\{A_i\}$ is a nontrivial partition of the phase space X, then

$$\lim_{t\to\infty} H_C((P_t\rho(t))^{\sim}|\tilde{\rho}_*) = 0 \tag{6.1.7}$$

for all initial densities ρ.

But we must realize that the way in which the conditional entropy converges to zero depends on the way in which the coarse-graining is carried out (Lasota and Mackey, 1985). It can be proved that the rate of convergence of entropy to equilibrium slows as the measurement techniques improve and the coarse-graining becomes finer (!!!). Such phenomena have not been observed. Thus it is most unlikely that trivial coarse-graining plays a role in determining thermodynamic behavior; if a natural graininess, it is not found. Candidates for this natural and universal graininess would be: (i) The graininess produced by operators $\tilde{\Pi}$ and $\hat{\Pi}$, which are introduced using extended dynamics methods (Castagnino *et al.*, 1996; Balescu, 1975). However, this is really only the old, coarse-graining version of the extended dynamics method. (ii) The graininess introduced by the universe event horizon (Hu *et al.*, 1992b). (iii) The graininess introduced by Planck's constant; *e.g.*, it appears to be absolutely impossible to measure lengths smaller than the Planck length; the physics related to this graininess is still a subject of intense research, however. More general projectors than those defined in (6.1.4) can be used, as we have seen in the introduction, since any projector will do the

job done in (1.2.1.2) and (1.2.1.3). A theory that uses one of these generalized projectors will also be called, by extension, a "coarse-graining" theory. Coarse-graining can be also used in the quantum case. Then Π is a projector over the quantum Liouville space $\mathcal{L} = \mathcal{H} \otimes \mathcal{H}$. Using the quantum Theorem 5.2.1, we can obtain the same results as in the classical case, where we used Theorem 4.3.1.

6.2. Time Asymmetry in Coarse-Graining Theories

All the limits of the previous section were computed as $t \to +\infty$. But it is clear that all of these limits are also valid when $t \to -\infty$. Therefore in coarse-graining theories there is equilibrium both in the far past and the far future. This fact can easily be verified with the baker's transformation, where we have a set of infinite lines, horizontal for the far future and vertical for the far past, which will be taken to have a uniform equilibrium distribution function for any coarse-graining partition. It is also evident that, if the initial distribution function has an adequate symmetry, for example, that of the characteristic function of a square domain, the evolutions toward the past and toward the future will be strictly symmetric. But this will not be the case if the initial distribution function is not symmetric [other calculations concerning the baker's transformation behavior can be found in Lasota and Mackey (1985)].

Continuing to the quantum case, we see a quite similar phenomena, even with no coarse-graining. Let us consider the state $|1\rangle$ of the Friedrichs model (which can be considered as a symmetric initial condition, as the characteristic function of a square domain, in the case of the baker's transformation). The behavior of the survival probability $P(t)$, as shown in Fig. 3, is completely symmetric with respect to $t = 0$. Thus, classically, if we use coarse-graining techniques, or quantum mechanically, if we use only states of the Hilbert space, then we will find that the past is only conventionally different than the future. How, then, may we distinguish past from future? The answer is by the method which was explained in the introduction: Take the time $t = 0$. Consider the set of evolutions of the system for $t > 0$ (for all possible initial conditions) and let us call it $\mathcal{H}_-$ in the quantum case (or $\mathcal{L}_-$ in the classical case). It is identical to the set of evolutions for $t < 0$ (for all possible initial conditions), which we will call $\mathcal{H}_+$ (or $\mathcal{L}_+$). The existence T or K of the mathematical transformation relating the future with the past evolutions shows that these sets of evolutions are identical:

Theorem 6.2.1. For every evolution $\rho(t) \in \mathcal{L}_-$ $(t > 0)$ there is a time-symmetric evolution $\rho(-t) \in \mathcal{L}_+$ $(t < 0)$ if the evolution equation is reversible.

Proof. From the definition of a reversible evolution (2.3.21), for every $\rho(t) \in \mathcal{L}_-$ there is a physical evolution $\rho(-t) \in \mathcal{L}_+$ defined by

$$\rho(-t) = \mathcal{K}\rho(t), \qquad t > 0 \quad \text{QED}$$

We also have

$$T: \quad \mathcal{L}_- \to \mathcal{L}_+, \qquad T: \quad \mathcal{L}_+ \to \mathcal{L}_-$$

These two sets of evolutions $\mathcal{L}_-$, $\mathcal{L}_+$ (or $\mathcal{H}_-$, $\mathcal{H}_+$) are the two mathematical structures introduced in Section 1.1. Since they are identical [cf. equation (6.2.2)] it is irrelevant which one we choose. So let us choose just one of these structures to build our theory, say $\mathcal{L}_-$, and discard the other. Now we can say that the theory begins at $t = 0$ and goes toward the future for $t > 0$ (or toward the past, since the choice of one word or the other is just as conventional as the choice between $\mathcal{L}_-$ and $\mathcal{L}_+$). It is evident that this theory developed in the period $0 \leq t < \infty$ will fulfill all our requirements, *provided that we forget about the period* $-\infty < t \leq 0$. These are the characteristics of the resulting theory if we use coarse-graining and usual Hilbert-space quantum states. Even if successful in many respects, it does admittedly make one a little uneasy.

6.3. Traces

Let us now consider only the classical case. Let X and Y be two topological Hausdorff phase spaces, φ: $Y \to X$ a given continuous function on X, and S_t: $Y \to Y$ a dynamical system operating in phase space Y. A function h: $R \to X$ is a trace of the dynamical system if there is a point y in space Y such that $h(t) = \varphi(S_t(y))$ for all times t (this usage of the word trace should not be confused with that referring to the trace of a matrix). It can be proved that every continuous function in a space X is the trace of a single dynamical system operating in a phase space Y, and thus we find the following quite surprising result:

Theorem 6.3.1 Lasota and Mackey. (1985). Let the phase space X be an arbitrary, but topological, Hausdorff space. Then there is a second phase space Y, also topological and Hausdorff, a dynamical system S_t operating in Y, and a continuous function φ: $Y \to X$ such that every continuous function h: $R \to X$ is the trace of S_t. [A topological space is Hausdorff (or separable) if any two distinct points possess disjoint neighborhoods.]

That is, for every h there is a point y in phase space Y such that $h(t) = \varphi(S_t(y))$, for all times t. Let us now consider the trajectories of a dynamical system: if we have a dynamical system S_t operating in a phase space Y, then only three possible types of trajectories can be observed: (i) The trajectory is a fixed point x_* such that $S_t x_* = x_*$ for all t. (ii) The trajectory is a nonintersecting curve, with the property $S_t(x) \neq S_{t'}(x)$ if $t \neq t'$. (iii) A periodic trajectory such that $S_t(x) = S_{t+T}(x)$, for all times t, with T the period. Nothing,

however, prevents the existence of nonperiodic intersecting trajectories $h(t)$ in space X if φ: $Y \to X$. Thus we can demonstrate:

Theorem 6.3.2. Let the phase spaces X and Y be topological Hausdorff spaces and h: $R \to X$ an intersecting and nonperiodic trace of a dynamical system S_t: $Y \to Y$. Then the entropy of densities evolving under the action of h is either constant or increasing.

Proof. The proof is based on the trivial observation that if h is intersecting and nonperiodic, then at every intersection point x on the trajectory h the inverse $h^{-1}(x)$ is not unique. Therefore the trace h is the trajectory of a semidynamical system, and, since semidynamical systems are irreversible, from Theorems 3.6.1 and 3.6.2, the entropy is either constant or increasing. QED

Thus the simple act of trace, taking a trace of a dynamical system (with time-constant entropy), may be sufficient to generate a system in which the entropy is increasing. For certain classes of traces, however, much more can be said. Let X and Y be two different phase spaces with normalizable measures μ_* and ν_* and associated densities ρ_* and σ_*, respectively, and let T_t: $X \to X$ and S_t: $Y \to Y$ be two measure-preserving transformations. If there is a transformation φ: $Y \to X$ which is also measure preserving, *i.e.*, if

$$\nu_*(\varphi^{-1}(A)) = \mu_*(A) \tag{6.3.1}$$

for all subsets A of the phase space Y such that $T_t \circ \varphi = \varphi \circ S_t$, then T_t is called a factor of S_t. From this definition the trajectory of the factor T_t is a trace of the system S_t. Then we have the following:

Theorem 6.3.3 (Rochlin, 1969). Every ρ_*-exact transformation is the factor of a Kolmogorov automorphism.

This theorem indicates precisely what we must do if we want to find an exact transformation with all its nice properties: (i) We must show that the system we are working with is a Kolmogorov system. This can be difficult from the mathematical point of view, but as chaos is very frequent in nature, it is not a very restrictive physical condition. (ii) Then, according to Theorem 6.3.3, every measure-preserving factor will produce an exact transformation. The problem is simply to find the most convenient one. As an example, let us again consider the baker's transformation. It can be proved that this transformation is a Kolmogorov automorphism, endowed with a constant entropy. However, the system corresponding to coordinate x is a factor of the baker's transformation. Also, it is identical to the dyadic Renyi transformation:

$$T(x) = 2x \pmod 1 \tag{6.3.2}$$

which is uniformly exact and whose entropy smoothly increases to zero by Theorem 4.4.2.

We have shown that coarse-graining produces no substantial difference between past and future. This is not the case with traces, as we can see from the baker's transformation, in which that x side of a parallelogram will always increase toward the future and decrease toward the past. Thus, coarse-graining does not produce time asymmetry, whereas traces do produce this phenomenon. Let us now summarize the entire picture: We have projectors Π such as those introduced in Section 6.1, namely

$$\Pi: \quad \mathcal{L}_Y \to \mathcal{L}_X, \qquad \Pi^2 = \Pi \tag{6.3.3}$$

where $\mathcal{L}_Y = \mathcal{L}$ is the state space and $\mathcal{L}_X$ is the space of relevant states. Π does not have an inverse Π^{-1}, since $\Pi^2\,\Pi^{-1} = \Pi\Pi^{-1}$ yields $\Pi = 1$. We have traces

$$\varphi: \quad Y \to X \tag{6.3.4}$$

which map between phase spaces. φ can have an inverse, and in this case $\varphi(Y)$ is dense in X, or it does not have an inverse when the dimension of X is smaller than the dimension of Y, for example, in the case of (6.3.2). Finally, let us remark that traces define a mapping in the corresponding Liouville spaces. Let $\mathcal{L}_X$ and $\mathcal{L}_Y$ be the Liouville spaces corresponding to the phase spaces X and Y. Then to the mapping φ: $Y \to X$ corresponds the mapping

$$\Lambda^{-1}: \quad \mathcal{L}_Y \to \mathcal{L}_X \tag{6.3.5}$$

(the -1 is just a matter of convention) defined by

$$\Lambda^{-1}\rho(y) = \rho(\varphi^{-1}(x)) \tag{6.3.6}$$

In the next subsection we will study even more general mappings.

6.4. Generalized Traces

In Section 6.3 we were forced to work only in the classical case, since we used phase space. We would now like to generalize the trace notion in order to be able to also consider the quantum case. A generalized trace is given by (6.3.3) if (6.3.4) is not fulfilled; *i.e.*, it is a mapping between Liouville spaces not originated by a mapping between phase spaces. As it is a mapping like (6.3.3), it is similar to a "projector with an inverse." But now spaces $\mathcal{L}_Y$ and $\mathcal{L}_X$ can be classical or quantum Liouville spaces. These generalized traces are typical of the extended dynamics formalism, and attempt to show it as a kind of generalization of the coarse-graining one. Let us consider the particular case $\mathcal{L}_Y = \Phi_+^\times$, $\mathcal{L}_X = \mathcal{L}$. Let us also consider the basis $\{|1\rangle, |\omega\rangle\}$ of (5.3.1), which we shall call $\{|i\rangle\}$, and such that $H|i\rangle = z_i|i\rangle$, $z_i \in C$. Let us define the basis $\{|ij)\}$, $|ij) = |i\rangle\langle j|$. Let us also consider the basis $\{|z_1, -\rangle, |\varphi, -\rangle\}$ of (5.3.23), which we will call $\{|i, -\rangle\}$, and, in

the same fashion, let us define the basis $\{|ij, -)\}$, $|ij, -) = |i, -\rangle\langle j, -|$. Using the basis $\{|z_1, +\rangle, |\omega, +\rangle\}$, we can, in addition, define a basis $\{|ij, +)\}$. We can then define a generalized trace as

$$\Lambda^{-1}: \quad \Phi_+^\times \to \mathscr{L} \tag{6.4.1}$$

$$\Lambda = \sum_{ij} |ij, -)(ij|, \qquad \Lambda^{-1} = \sum_{ij} |ij)(ij, +| \tag{6.4.2}$$

In other words, Λ is the transformation which describes the correspondence of each state ρ of space $\mathscr{L}$ to a functional in space $\Phi_+^\times$. Λ looks like just a "change of basis." But Λ is actually much more than a change of coordinates, since it takes vectors in one space to vectors in another. Therefore to weak limits in $\mathscr{L}$ correspond strong limits in $\Phi_+^\times$ and the generalized trace Λ embodies the solution of our problem: going from weak limits to strong limits can be considered as the symbol which synthesizes the fine-graining technique.

Some observations are in order:

(i) Since Λ is a generalized trace, therefore, as a trace, it contains time asymmetry. In fact, Λ defined in (6.4.1) is related to damping phenomena which produce equilibrium toward the future, and should be called Λ_+. We can also define a Λ_- related to creation phenomena implying equilibrium in the far past, namely

$$\Lambda_-: \quad \Phi_-^\times \to \mathscr{L} \tag{6.4.3}$$

$$\Lambda_- = \sum_{ij} |ij, +) (ij|, \qquad \Lambda_-^{-1} = \sum_{ij} |ij)(ij, +| \tag{6.4.4}$$

If we choose Λ_- rather than Λ_+, we are creating also a time asymmetry. In order to see the relation of the two Λ's, let us introduce the star conjugation

$$A^* = \mathscr{K} A^\dagger \mathscr{K}^\dagger \tag{6.4.5}$$

Then it is easy to see that

$$\Lambda_+ = \Lambda_-^*, \qquad \Lambda_- = \Lambda_+^* \tag{6.4.6}$$

$$\Lambda_-\Lambda_-^* = \Lambda_+\Lambda_+^* = 1, \qquad \Lambda_-^{-1} = \Lambda_-^* = \Lambda_+, \qquad \Lambda_+^{-1} = \Lambda_+^* = \Lambda_- \tag{6.4.7}$$

For these last equations we can say that the Λ's are star-unitary.

(ii) Using the generalized trace Λ, we do not lose any information. [In the case of usual traces we lose information if dim $X <$ dim Y, as in the example of the baker's transformation before, (6.3.2), but in the case of the Λ trace the dimension of the two spaces is the same and $\mathscr{L}$ is dense in $\Phi_+^\times$]. However, it can be demonstrated that this generalized trace Λ somehow

renormalizes the infinite amount of information contained in $\mathscr{L}$ (Ordonez, 1997).

(iii) A generalized trace is not a trace, so there is not a mapping in the corresponding phase spaces linking corresponding trajectories. In this sense, in using extended dynamics techniques, trajectories lose all their importance and even have no meaning Prigogine and Petrosky, 1993; Prigogine, 1993).

(iv) The Λ trace allows us to define a Hilbert space in which the time evolution is irreversible. In fact, using the bases we have introduced, we can deduce that

$$1 = \sum_{ij} |ij, -)(ij, +| \tag{6.4.8}$$

$$L = \sum_{ij}; (z_i - z_j^*)|ij, -)(ij, +| \tag{6.4.9}$$

whereas $\text{Im}\, z_i = -\gamma_i/2 < 0$, we have $\text{Im}(z_i - z_j^*) \leq 0$. The time evolution operator is $U(t) = e^{-iLt}$ and $UU^\dagger = 1$, *i.e.*, U is unitary. Let us now define a modified Liouvillian:

$$G = \Lambda^* L \Lambda = \sum_{ij} (z_i - z_j^*)\, |ij)(ij| \tag{6.4.10}$$

which induces an evolution $W(t) = e^{-iGt}$ such that $WW^\dagger \neq 1$, and therefore is not unitary, but star-unitary, $WW^* = 1$. The two evolution are related by

$$W(t) = \Lambda^* U(t) \Lambda \tag{6.4.11}$$

We can also define Λ-density matrices as being related by the Λ-trace (6.4.1):

$$\rho_\Lambda(t) = \Lambda^* \rho(t), \qquad \rho(t) = \Lambda \rho_\Lambda(t) \tag{6.4.12}$$

where $\rho(t) \in \Phi_+^\times$, $\rho_\Lambda(t) \in \mathscr{L}$, which evolves as

$$\rho(t) = U(t)\rho(0), \qquad \rho_\Lambda(t) = W(t)\rho_\Lambda(0) \tag{6.4.13}$$

Equation (5.3.24), translated into the ρ_Λ language, read

$$\rho_\Lambda(t) = \rho_{\Lambda^*} + e^{-\gamma_1 t/2}\rho_{\Lambda 1}(t) + e^{-\gamma_1 t}\rho_{\Lambda 2}(t) \tag{6.4.13$'$}$$

which can also be obtained from equation (6.4.13), since the $\rho_\Lambda(t)$ evolves under the action of the operator e^{-iGt} and G has complex eigenvalues [cf. (6.4.11]. Then the space $\Phi_+^\times$ of the ρ'*s* can be considered as an ideal reversible world of reversible equations, namely the ideal world of Newton, endowed with unitary evolutions [David Bohm would say that this is the space of implicate order (Ordonez, 1997)], while the space $\mathscr{L}$ of the ρ_Δ is the real, physical, irreversible world of Boltzmann, endowed with nonunitary evolution [just star-unitary; David Bohm would say that this is the space of explicate order (Ordonez, 1997)]. Λ establishes a canonical mapping between these

two worlds [David Bohm would say a "metamorphosis" (Ordoñez, 1997)]. Even if the ρ_Δ's live in the ordinary Liouville space, they evolve with a nonunitary law [cf. (6.4.13)], so that the Λ-trace achieves the dream of physicists: it creates an ordinary Hilbert space where the evolutions are nonunitary and irreversible. Precisely, the essence of the fine-graining formalism was to maintain the time-symmetric primitive equations [with operator evolution $U(t)$] and to obtain time asymmetry by choosing a typical time-asymmetric space $\mathscr{L}_Y = \phi^{\times}_{+}$. The Λ-trace exchanges these roles. We get a time-asymmetric equation [with evolution $W(t)$] in a time-symmetric space $\mathscr{L}$ as in the coarse-graining case. But of course the physics remains the same.

(v) Using the Λ, Lyapunov variables can be found very easily, since

$$(\rho(t)|\rho(t)) = (\rho(0)|U^\dagger U|\rho(0)) = (\rho(0)|\rho(0)) = \text{const} \tag{6.4.14}$$

therefore it is not a Lyapunov variable, but

$$Y = (\rho_\Lambda(t)|\rho_\Lambda(t)) = (\rho_\Lambda(t)|W^\dagger W|\rho_\Lambda(0)) = \text{var} \tag{6.4.15}$$

or

$$Y = (\rho(t)|(\Lambda^*)^\dagger \Lambda^*|\rho(t)) = (\rho(t)|M|\rho(t)) = \text{var}, \qquad M = (\Lambda^*)^\dagger \Lambda^* \tag{6.4.15'}$$

Precisely, if

$$|\rho_\Lambda(0)) = \sum_{ij} \rho_{ij}|ij) \tag{6.4.16}$$

the corresponding time evolution is

$$|\rho_\Lambda(t) = \sum_{ij} \rho_{ij} e^{-i(z_i - z_j^*)}|ij \tag{6.4.17}$$

But $\text{Im}(z_i - z_j^*) = -\Gamma_{ij} \leq 0$, so

$$(\rho_\Lambda(t)|\rho_\Lambda(t)) = \sum_{ij} |\rho_{ij}|^2 e^{-\Gamma_{ij} t} \tag{6.4.18}$$

is always decreasing and is therefore a Lyapunov variable. In more general cases than that of (6.4.1)–(6.4.2), it can be proved that every rigging corresponds to a Λ-trace and vice versa (Ordoñez, 1997).

6.5. Time Asymmetry in Extended Dynamics Theories

Let us begin to compute the conditional entropy $H_C(\rho|\rho_*)$ [cf. equations (3.6.1) and (3.6.2)] and in the case of the classical evolution (4.5.1.12). [In the quantum case we have the time evolution (5.4.13).] If we want to use the classical equation for H_C and we have a quantum density matrix, we must

first transform this quantum matrix to the corresponding classical distribution function, using (6.A.1). But we can directly use the definition of H_C, if we define the logarithm of a quantum density matrix as the operator whose eigenvalues are equal to the logarithms of the eigenvalues of the primitive operator (Landau and Lifshitz, 1958). As the Wigner integral is linear, the quantum equivalent of (5.4.13) is the same equation. Then

$$H_C(\rho|\rho_*) = -\int_x (\rho_* + e^{-\gamma t}\rho_1) \log\left(1 + e^{-\gamma t}\frac{\rho_1}{\rho_*}\right) dx \qquad (6.5.1)$$

Bearing in mind that $|\rho_1| << \rho_*$ for $t >> \gamma^{-1}$, we can expand the logarithm and, since $tr\ \rho_1 = 0$, taking into account (6.A.3), we have

$$H_C(\rho|\rho_*) = -e^{-2\gamma t}\int_X \frac{\rho_1^2}{\rho_*} dx \qquad (6.5.2)$$

which is negative, growing, and has a vanishing limit when $t \to \infty$; so it has all the properties required to formulate the second law of thermodynamics in its third-order form. But we have obtained this satisfactory conclusion because we worked with the operator test function space Φ_+, and the quantum states belong to space $\Phi_-^\times$ [albeit with some remaining mathematical problems, since we are computing the log of a vector of a rigged Hilbert space; these problems can be solved, in principle, if we use the generalized trace Λ of the previous section, and if we substitute the ρ'*s* by ρ_Λ'*s*, since these last density matrices belong to $\mathscr{L}$, so they can be used without problems, but they retain the evolution properties of the ρ'*s*, namely the damping factors of (6.5.1) and (6.5.2) as in (6.4.13′)]. Then, explicitly, we define

$$H_C(\rho|\rho_*) = -\int_X \rho_\Lambda \log\frac{\rho_\Lambda}{\rho_{\Lambda^*}} dx)$$

So we can go now to the central problem of the origin of time asymmetry in extended dynamics theories.

Let us now consider an isolated system which is our universe; we can know nothing of the exterior of the system and the system cannot interact with anything outside itself. If the time evolution equations of a theory are time-symmetric, it is quite impossible to break this symmetry by rigorous mathematical manipulations; symmetry will always appear one way or another. Nevertheless, the examples we have given demonstrate that normally in these theories we can find two extensions of Liouville space $\mathscr{L}$. They are the rigged Hilbert spaces $\Phi_-^\times$ and $\Phi_+^\times$ which are defined using the test spaces Φ_- and Φ_+ (usually these test spaces are nuclear spaces which can be consid-

ered as the spaces corresponding to the operators of the measurement devices, as explained in Appendix 4A). Time symmetry implies that these spaces are related by

$$\mathcal{H}: \quad \Phi_{-}^{\times} \to \Phi_{+}^{\times}, \qquad \mathcal{H}: \quad \Phi_{+}^{\times} \to \Phi_{-}^{\times} \tag{6.5.3}$$

and that therefore they are identical. To choose one or the other is irrelevant, as irrelevant as the throw of a die with the same number on each of its faces. So if we choose one space or the other, the physics does not change. Both spaces are only conventionally different. Any possible difference could only occur for the exterior of the system and there is no possibility of interaction with it. Nevertheless, in each space $\Phi_{\mp}^{\times}$ or $\Phi_{+}^{\times}$, the future is substantially different from the past, since there is equilibrium in only one of these directions and we can call this direction the future. In choosing one of the spaces, we establish a time asymmetry, and we can formulate the second law of thermodynamics as we have done; our problem is solved (cf. the solution with the coarse-graining case, which is not so different). We can say the same when we speak of the choice of the generalized traces Λ_{-} or Λ_{+} and work within the Liouville space $\mathcal{L}$, the space of physical states. The same trick can be done in various different ways:

(i) In Antoniou and Tasaki (1991, 1993a, b) and Antoniou *et al.* (1995) two semigroups are defined, each related to a rigged Hilbert space, and one of these semigroups is arbitrarily chosen. One semigroup is obtained by expanding the solution of the evolution equation in a basis of $\Phi_{+}^{\times}$; the evolution turns out to be well defined for $t \in (-\infty, +\infty]$ and not well defined for $t \to -\infty$. The other semigroup has the reverse properties.

(ii) In Balescu (1975) a projector $\Pi_{-} = \Sigma_i \mid ii, -)(ii, +\mid$ is defined and taken to be the projector on the really relevant space. But $\Pi_{+} = \mathcal{H}\Pi_{-}$ is identical to Π_{-}, so that we must choose one or the other as in the previous cases.

So in all these cases we must make a conventional choice to find a mathematical structure, a space, semigroup, projector, or whatever, such that, using this structure, the future exhibits substantially different properties from the past. One might say that we have not explained time asymmetry, but have merely introduced it by making an arbitrary choice. To answer this criticism we must remember that physics never really explains. It merely finds the mathematical structure most adequate to explaining physical phenomena; for example, the most adequate mathematical space, the most adequate mathematical equations, and so on. The curvature of space-time does not explain gravity, the Riemannian manifold happens to be the best mathematical structure to deal with gravity. Analogously, it simply turns out that the most adequate mathematical space to explain time asymmetry and the second law of thermodynamics is a rigged Hilbert

space, not the usual Hilbert space: *So the relevant important choice is between mathematical structures* $\mathscr{L}$ *and* $\Phi^{\times}_{-}$ (*or equivalently* $\Phi^{\times}_{+}$). *The particular choice from these last two rigged spaces,* $\Phi^{\times}_{-}$ *and* $\Phi^{\times}_{+}$, *is, on the contrary, irrelevant and physically unimportant*

6.6. Comparison between Extended Dynamics and Coarse-Graining

As we can see, coarse-graining and extended dynamics are very similar. Both are based on Theorem 4.3.1, which concerns the weak limit of mixing evolutions. Coarse-graining obtains a strong limit via a projection, extended dynamics obtains the "strong limit" using functionals. Both obtain their arrow of time by defining a pair of time-symmetric structures: (i) The pair $\mathscr{L}_{-}$, $\mathscr{L}_{+}$ of $t > 0$ and $t < 0$ evolutions in the case of coarse-graining. (ii) The pair of rigged Hilbert spaces $\Phi^{\times}_{-}$, $\Phi^{\times}_{+}$ in the case of fine-graining. In both methods one of these structures is conventionally chosen. The main weakness of coarse-graining is that the projector is not defined in a canonical way. The main weakness of extended dynamics is that we are forced to enlarge the space, and that we do not know the exact nature of the objects we must add. Are these ideal unstable states just useful mathematical tools (like the Fadeev–Popov ghost), or real physical objects? The answer to this question depends on our point of view when studying the problem. In fact:

(i) Any decaying state was always created by a creation process. The quantum state corresponding to the creation process followed by decay belongs to $\mathscr{H}$ [like the vector $|1\rangle$ of the Friedrichs model with the survival probability (5.3.19), that of Fig. 3]. Nevertheless, if the lifetime of the decaying state is very long, we are used to neglecting the creation process and to considering the state just as a decaying state with exponentially decaying survival probability [as in (5.3.20)]. This is the state $|z_1\,-\rangle$ belonging to $\phi^{\times}_{+}$. So the quantum theory, which uses these states, could be considered as an effective theory where creation process are neglected. We can say the same for classical theories. In the baker's transformation, a regular density with a regular support will have a creation process and a symmetric decaying process toward equilibrium in much the same way as state $|1\rangle$. But if we study the time evolution of a "horizontal Dirac comb" state, we will find that these ideal states have no creation process, as the state $|z_2\,-\rangle$.

(ii) Nevertheless, what may be merely a useful simplification when considering states with long lifetime could be a rigorous fact in the case of the universe, where this process must be necessarily neglected, since we do not know its creation process.

(iii) So the new unstable states added to physical space are similar to plane waves; they are eternal objects with no creation process and, in fact, if we define plane waves in a rigorous way, we need a rigged Hilbert space

to do it. From this point of view, coarse-graining physicists appear to be rather stubborn people who work only with waves packets, and refuse to use plane waves because they "are not physical objects."

(iv) If we allow time to go to infinity and we wish to consider the rigorous equilibrium state at infinite time, this state belongs to $\Phi_+^\times$ as in the case of the baker's transformation, and so we are forced to work with an extended dynamics theory. But if we content ourselves with approximate equilibrium states at finite time, arguing that $t \to \infty$ is physically impossible, we do not need these states. Thus the real nature of the new states is open to discussion. However, an extended-dynamics physicist can take a conservative attitude and consider the new states simply as idealized states, or just as useful mathematical devices as are plane waves. Is the choice of the extended dynamics or coarse-graining just a matter of taste, or there are physical or mathematical reason to choose one or the other? The reader must decide for him or herself.

6.7. Appendix 6A: Wigner Function Integral (Balazs and Voros, 1990; Hillery *et al.*,1984)

We have repeatedly jumped back and forth between the classical and quantum cases. It is therefore interesting to present a theory to formalize these jumps and to base some applications on it. Let ρ be a density matrix of Liouville space $\mathscr{L} = \mathscr{H} \times \mathscr{H}$ and let $\{|q\rangle\}$ be the configuration or position basis of the Hilbert space $\mathscr{H}$. The corresponding Wigner function reads

$$\rho_w(q, p) = \pi^{-1} \int \langle q + \lambda|\rho|q - \lambda\rangle e^{2i\lambda p}\, d\lambda \tag{6.A.1}$$

It can be proved that

$$L\rho_W(q, p) = \pi^{-1} \int \langle q + \lambda|L\rho|q - \lambda\rangle e^{2i\lambda p}\, d\lambda + O(\hbar) \tag{6.A.2}$$

where L is respectively the classical and quantum Liouville operator. In the classical limit, $\hbar \to 0$, and therefore ρ_W can be considered as the classical distribution function corresponding to ρ. As in the classical regime, for practical reasons, we work in this limit, and we will consider (6.A.1) to be the relation between the quantum density matrix and the classical distribution function. In fact, even if ρ_W is not generally positive definite, using the Wigner integral from the classical equation, we can pass to the quantum equation and vice versa, as a few examples will show. For example, let us observe that

$$\|\rho_W\| = \iint \rho_W(q, p)\, dq\, dp$$
$$= \int dq \int \langle q + \lambda|\rho|q - \lambda\rangle\, \delta(\lambda)\, d\lambda = tr\, \rho \qquad (6.A.3)$$

so the quantum trace corresponds to the classical norm. Also

$$(\rho_W|O_W) = \iint \rho_W(q, p) O_W(q, p)\, dq\, dp$$
$$= \pi^{-2} \int dq \iiint \langle q + \lambda|\rho|q - \lambda\rangle\langle q + \mu|O|q - \mu\rangle$$
$$\times e^{2ip(\lambda+\mu)} dp\, d\lambda\, d\mu \qquad (6.A.4)$$
$$\pi^{-1} \int dq \int \langle q + \lambda|\rho|q-\lambda\rangle\langle q - \lambda|O| - 1\lambda\, d\lambda \cong tr(\rho O)$$

Therefore to the inner product in classical Liouville space there corresponds the inner product in the quantum Liouville space. This fact completes the analogy between classical and quantum spaces implemented by the Wigner integral.

As an exercise, we can compute the classical distribution function corresponding to the density matrices $\rho_1(t)$ and $\rho_2(t)$ of (5.3.26) and (5.3.27). Since these equations will be used in Section 8, where we will use (6.5.2) to compute the entropy neglecting $O(\lambda)$, we will also do so in this exercise:

$$\rho_{W1}(t) = \pi^{-1} \int \langle q + \lambda| \int_0^\infty (\rho_{1\omega} e^{-1}(\omega^1 - \omega)t\, |1\rangle\langle\omega|$$
$$+ \text{h.c.})\, d\omega|q - \lambda\rangle e^{2ip\lambda}\, d\lambda \qquad (6.A.5)$$

where

$$\langle q|\omega\rangle = \frac{1}{\sqrt{2\omega}}\, e^{-i\sqrt{\omega q}} \qquad (6.A.6)$$

and where $2m = 1$, $\hbar = 1$, etc. After a simple calculation we obtain

$$\rho_{W1}(q, p, t) \approx \rho_{1,(2p-\sqrt{\omega 1})^2}\, e^{4ip(\sqrt{\omega}_1 - p)t} e^{2i(p+\sqrt{\omega}_1)q} + \text{h.c.} \qquad (6.A.7)$$

We can see that the main values of this distribution function are obtained when $p = \sqrt{\omega_1}$ since for other values there are rapid oscillations. Analogously,

$$\rho_{w2}(t) = \pi^{-1}\, \rho_{11} \int \langle q + \lambda|1\rangle\langle 1|q - \lambda e^{2ip\lambda}\, d\lambda \approx \delta(p - \sqrt{\omega_1}) \qquad (6.A.8)$$

Therefore, in this case also, all the effect is concentrated around the energy ω_1, which, in the application of Section 8, will correspond to the characteristic energy of nuclear reactions.

7. ENTROPY IN CURVED SPACE-TIME

We have mentioned cosmology in two contexts: (i) The cosmological event horizon could be a way to explain a universal graininess of nature. (ii) Extended dynamics time asymmetry is explained using a system with no exterior, namely the universe. Furthermore, we must investigate a cosmological arrow, so we cannot avoid cosmology in a complete discussion of this subject. Many years ago Mach taught us that most basic physical facts can only be explained if we consider the universe as a whole. For example, if we wish to explain why a system is either an inertial one or not, we must consider the whole universe; the system will be inertial if it is in uniform translatory motion with respect to the matter of the whole universe. The arrows of time are not exceptions, since they have a global nature. In fact, sugar lumps disolve in coffee everywhere in the same time direction, here and in the Andromeda nebula. We must explain why this is so, and we will find the explanation only if we define the arrow of time in global cosmological models. Thus, since cosmological models are presented in these spaces (Tolman, 1987), let us begin by studying the notion of entropy in curved space-time.

7.1. Thermodynamics in Special Relativity

For phenomenological reasons we can assume that the laws of thermodynamics are valid in the special-relativistic proper system of coordinates S^0. From the relativity principle we then know that these laws are also valid in every inertial system S in translatory uniform motion with respect to S^0, provided that the quantities involved in these laws can be transformed conveniently. In other words, we would like to obtain the "Lorentz transformation" which makes invariant the following laws:

(i) The first law:

$$\Delta E = \Delta Q - \Delta W \tag{7.1.1}$$

where E is the energy, Q the heat, and W the work.

(ii) The second law:

$$\Delta S \geq \frac{\Delta Q}{T} \tag{7.1.2}$$

where S is the entropy, T is the temperature, and the equality only holds for reversible evolution.

Let us then (a) suppose that the pressure is isotropic, i.e., it is equal in all directions, and (b) for simplicity, temporarily use axes chosen in such a way that the velocity u of the system S with respect to the system S^0 is parallel to the x axis. Then, from ordinary special relativity, we know the coordinate transformation equations for the following mechanical quantities:

(i) For the volume v

$$v = v_0 \sqrt{1 - u^2} \tag{7.1.3}$$

(ii) For the pressure p

$$p = p_0 \tag{7.1.4}$$

(iii) For the energy E

$$E = \frac{E_0 + p_0 v_0}{\sqrt{1 - u^2}} \tag{7.1.5}$$

(iv) For the work W

$$dW = \sqrt{1 - u^2}\, dW_0 + \frac{u^2}{\sqrt{1 - u^2}}\, d(E_0 + p_0 v_0) \tag{7.1.6}$$

where the quantities with subscript 0 refer to system S^0.

Then, for the covariance of the first law, (7.1.1), it is necessary and sufficient that

$$Q = \sqrt{1 - u^2} Q_0 \tag{7.1.7}$$

so we have obtained the transformation law for the heat. Let us now consider a thermic system at S^0. We can accelerate this thermic system up to the velocity u in a reversible and adiabatic way, so that the entropy of the system is not modified. We obtain

$$S = S_0 \tag{7.1.8}$$

Finally, from equations (7.1.7) and (7.18) it is evident that the second law, (7.1.2), will be covariant if

$$T = \sqrt{1 - u^2} T_0 \tag{7.1.9}$$

So we have obtained the change of coordinate equation of all the basic thermodynamic quantities. Let us now find the corresponding equations in four-dimensional language. The first law is just a form of the statement of conservation of energy, and therefore its four-dimensional version will be

$$\partial_\mu T^{\mu\nu} = 0 \tag{7.1.10}$$

where $T^{\mu\nu}$ is a convenient energy-momentum tensor ($\mu, \nu, \ldots = 0, 1, 2,$

3). To deduce the four-dimensional form of the second law, let us consider a small volume of a thermodynamic fluid ν and let us call ϕ the entropy density at the point where this element of volume is located in such a way that $\phi\nu$ is the entropy of the element. If δt is a small time period, the second law reads

$$\frac{d}{dt}(\phi v)\delta t \geq \frac{\delta Q}{T} \tag{7.1.11}$$

or ($i, j, \ldots, = 1, 2, 3$)

$$\left(\frac{d\phi}{dt}\, v + \phi\, \frac{dv}{dt}\right)\delta t = \left(u_i \partial_i \phi + \frac{\partial \phi}{\partial t} + \phi \partial_i u_i\right) v \delta t \geq \frac{\delta Q}{T} \tag{7.1.12}$$

where $u^i = dx^j/dt$. Combining terms, we have

$$\left(\partial_i\left(\phi\, \frac{dx_i}{dt}\right) + \frac{\partial \phi}{\partial t}\right) v \delta t \geq \frac{\delta Q}{T} \tag{7.1.13}$$

but

$$\frac{ds}{dt} = \sqrt{1 - u^2} \tag{7.1.14}$$

and from equations (7.1.3) and (7.1.8)

$$\phi = \frac{\phi_0}{\sqrt{1 - u^2}} \tag{7.1.15}$$

$$\frac{\delta Q}{T} = \frac{\delta Q_0}{T_0} \tag{7.1.16}$$

Thus, we obtain

$$\partial_\mu\left(\phi_0\, \frac{dx^\mu}{ds}\right) \delta \upsilon \geq \frac{\delta Q_0}{T_0} \tag{7.1.17}$$

where $\delta\upsilon = \upsilon\delta t$ is the coordinate four-dimensional volume element, which is equal to

$$\delta v_0 = \frac{v}{\sqrt{1 - u^2}} \sqrt{1 - u^2}\, \delta t = \delta v$$

[cf. (7.1.3), (7.1.14)], the proper four-dimensional volume element; so we can use either one. Thus, if we define the flow of proper entropy or "entropy vector" as

$$S^{\mu} = \phi_0 \frac{dx^{\mu}}{ds} \tag{7.1.18}$$

then we obtain the four-dimensional version of the second law:

$$\partial_{\mu} S^{\mu} \delta v_0 \geq \frac{\delta Q_0}{T_0} \tag{7.1.19}$$

which is valid for all inertial systems, and we can write $\delta\upsilon$ instead of $\delta\upsilon_0$.

7.2. Thermodynamics in General Relativity

Using the transcription rules to go from special relativity to general relativity, namely

$$\eta_{\mu\nu} \to g_{\mu\nu}, \qquad \partial_{\mu} \to \nabla_{\mu}, \qquad \delta\upsilon \to \sqrt{-g}\,\delta\upsilon \tag{7.2.1}$$

the first law, (7.1.3), reads

$$\nabla_{\mu} T^{\mu\nu} = 0 \tag{7.2.2}$$

or, introducing the tensor density $\mathsf{T}^{\mu\nu} = \sqrt{-g}\,T^{\mu\nu}$, we have the ordinary divergence:

$$\partial_{\mu}(\mathsf{T}_{\mu\nu} + t^{\mu\nu}) = 0 \tag{7.2.3}$$

where $t^{\mu\nu}$ is the pseudo tensor density of potential energy-momentum. This would be the general-relativistic covariant equation with the closest resemblance to (7.1.3). Using the transcription rules on the second law, (7.1.19), we obtain

$$\nabla_{\mu} S^{\mu} \sqrt{-g}\,\delta\upsilon \geq \frac{\delta Q_0}{T_0} \tag{7.2.4}$$

where, as all the factors are scalars, we have in fact obtained an equation valid for all coordinate systems. Introducing the density $\mathsf{S}^{\mu} = \sqrt{-g}S^{\mu}$, since

$$\nabla_{\mu} S^{\mu} = \frac{1}{\sqrt{-g}}\,\partial_{\mu}\sqrt{-g}S^{\mu} = \frac{1}{\sqrt{-g}}\,\partial_{\mu}\mathsf{S}^{\mu}$$

this last equation reads

$$\partial_{\mu}\mathsf{S}^{\mu}\,\partial\upsilon \geq \frac{\delta Q_0}{T_0} \tag{7.2.5}$$

which, again, is the general-relativistic covariant equation with the closest resemblance to the special-relativistic second law (7.1.19). Of course these are not the unique covariant generalizations of the thermodynamic laws of

general relativity, but they are the simplest and they lead to successful applications.

7.3. Thermodynamics in Cosmology

Let us consider a Robertson–Walker metric:

$$ds^2 = dt^2 + a^2 d\sigma^2 \tag{7.3.1}$$

where $d\sigma$ is the comoving arc length, and a the scale factor or the radius of the universe. If the energy-momentum tensor corresponds to an isotropic fluid with density ρ_{00} and pressure p_0, the first law reads

$$\frac{d}{dt}(\rho_{00} a^3 \delta\sigma) + p_0 \frac{d}{dt}(a^3 \delta\sigma) = 0 \tag{7.3.2}$$

where $\delta\sigma$ is a comoving-coordinate, three-dimensional volume. If we consider a comoving, thermic fluid, there will be no exchange of heat among the comoving volumes, and $u^\mu = dx^\mu/ds = (1, 0, 0, 0)$, so the second law, as expressed by (7.2.5), reads

$$\partial_\mu(\phi_0 u^\mu \sqrt{-g}) \geq 0 = \frac{d}{dt}(\phi_0 a^3) \geq 0 \tag{7.3.3}$$

where ϕ_0 is the proper entropy density and a the scale factor or radius of the universe. If we multiply this equation by the constant-coordinate, comoving volume $\delta\sigma$, we obtain

$$\frac{d}{dt}(\phi_0 a^3 \delta\sigma) \geq 0 \tag{7.3.4}$$

This equation gives the recipe to compute the entropy in the comoving frame of a Robertson–Walker metric: *multiply the local proper entropy density by the proper volume.* This is, of course, a very reasonable and natural result, perhaps so natural that it can simply be assumed from the outset, but now, however, it has been rigorously proved. Let us check this result with just one calculation: We know that, in a radiation-dominated universe, the temperature obeys the law

$$T = T_0 \frac{a_0}{a} \tag{7.3.5}$$

which can be obtained by integrating equation (7.3.2) if we take $p_0 = \frac{1}{3}\rho_{00} \sim T_4$, namely the radiation state equation, and that the entropy of a blackbody radiation is given by the formula

$$S = \frac{4}{3} C_S T^3 V \tag{7.3.6}$$

where C_S is the Stefan coefficient, T the temperature, and V the volume. If we substitute these two last equations in (7.3.4), we see that the evolution of a radiation-dominated universe is reversible, as should be expected. From these considerations we can deduce that the only effect produced by the expansion of the universe in isotropic models is that the temperature decreases. This is the only effect we need to take into account from now on.

8. THE COSMOLOGICAL PROBLEM

8.1. The Problem of Time Asymmetry

The problem of the existence of the arrows of time or, in other words, the time asymmetry of the universe may be stated, as explained in the introduction, by asking the following questions:

(i) *Why is there time asymmetry in the universe if all the relevant physical laws are time-symmetric?* In fact, the universe has several time asymmetries, corresponding to the various arrows of time, thermodynamic, electromagnetic, psychological, etc., whereas its main laws are time-symmetric. (ii) *Why do all the arrows of time point to the same direction?*

In this section we would like to answer these questions, using our mathematical framework for the problem. Let us first review the main equation of Section 2. If the state of a physical system is described by ρ (classically ρ being the distribution function or quantum mechanically the density matrix) we will call $\rho^{\text{rev}} = \mathcal{K}\rho$ the state with reversed initial conditions [e.g., if K is the Wigner operator of quantum mechanics, then $\mathcal{K}\rho = K\rho K^{\dagger}$ (Messiah, 1962; Castagnino *et al.*, 1996; Castagnino and Gunzig, 1997)]. We will say that the conditions at $t = 0$ are time-symmetric if $\rho^{rev}(0) = \mathcal{K}\rho(0) = \rho(0)$ and time-asymmetric otherwise. If $\rho(t)$ is the state of the universe at time t, then the universe would have a time-symmetric evolution with respect to $t = 0$ if [cf. (2.2.21)]

$$\mathcal{K}\rho(t) = \rho(-t) \tag{8.1.1}$$

But the universe has, in fact, a time-asymmetric evolution, at least with respect to some instant of time which we call $t = 0$, such that

$$\mathcal{K}\rho(t) \neq \rho(-t) \tag{8.1.2}$$

If the evolution equations embodied in the Liouvillian universe, operator L, are time-symmetric, namely [cf. (2.2.11)]

$$\mathcal{K}L\mathcal{K}^{\dagger} = L \tag{8.1.3}$$

then corresponding to time-symmetric conditions at $t = 0$ there is a time-symmetric evolution (8.1.1) and to time-asymmetric conditions there corresponds a time-asymmetric evolution such as (8.1.2). In fact

$$\rho(t) = e^{-iLt}\rho(0) \tag{8.1.4}$$

and therefore, if the $t = 0$ condition is time-symmetric, we have

$$\mathcal{K}\rho(t) = e^{i\mathcal{K}L\mathcal{K}^{\dagger}t}\mathcal{K}\rho(0) = e^{iLt}\mathcal{K}\rho(0) = \rho(-t) \tag{8.1.5}$$

since $\mathcal{K}$ is an antilinear operator (namely $\mathcal{K}i = -i$). In the same way, the time-asymmetric case can be demonstrated. Then the observed time asymmetry of the universe evolution which obeys (8.1.2) can only be explained in two alternative ways: (i) Equation (8.1.3) is actually not exact, and there is a small, but relevant, time-asymmetric term in the Liouvillian (caused perhaps by the weak interactions), or (ii) we have

$$\mathcal{K}\rho(0) \neq \rho(0) \tag{8.1.6}$$

or in other words the initial state of the universe is not time-symmetric.

So, if we reject weak interactions, or any inventive manipulation of the otherwise time-symmetric physical laws as the origin of time asymmetry, then we must necessarily consider (8.1.6) as the only possible cause of this phenomenon. Since, in principle, asymmetry is a more generic property than symmetry (just as complex numbers are more frequent than real ones), then (8.1.6) seems very natural and therefore this will be the idea that we shall adopt in this section. If (8.1.6) is valid, then from (8.1.5) we have

$$\mathcal{K}\rho(t) \neq e^{iLt}\rho(0) = \rho(-t) \tag{8.1.7}$$

i.e., equation (8.1.2), is the equation we must prove. Finally, let us remark that the same explanation can be used to explain the other two fundamental asymmetries of nature, P and C. In fact, if

$$P\rho(0) \neq \rho(0), \quad C\rho(0) \neq \rho \tag{8.1.8}$$

we will have

$$P\rho(t) \neq \rho(t), \quad C\rho(t) \neq \rho(t) \tag{8.1.9}$$

even if

$$PLP^{\dagger} = L, \qquad CLC^{\dagger} = L \tag{8.1.10}$$

That is, equation. (, ,) can be demonstrated if we postulate the existence

of a small fluctuation between the amounts of matter and antimatter at the beginning of the universe.

8.2. Entropy, Fluctuations, and Irreversibility

Let us first study the thermodynamic arrow of time, and consider the entropy S as the state function representing most eloquently the thermodynamic state of the universe. (S can be computed using coarse-graining entropy, or extended dynamics entropy.) We know that the vast majority of possible states of the universe will be near the equilibrium state ρ_* and will have the equilibrium entropy S_*. Nevertheless we know that fluctuations around the equilibrium state, those less probable, unstable states near the equilibrium, will spontaneously appear, and we also know that the entropy in these fluctuation states will be smaller than S_*. In any case, the steady equilibrium state satisfies Liouville equation

$$L\rho_* = 0 \tag{8.2.1}$$

For simplicity let us consider that there is just one equilibrium state in the universe, as is very likely since the universe appears to be chaotic and therefore it is at least ergodic. Then from (8.1.3) we have

$$L\mathcal{K}\rho_* = \mathcal{K}L\mathcal{K}^\dagger \mathcal{K}\rho_* = 0 \tag{8.2.2}$$

So

$$\mathcal{K}\rho_* = \rho_* \tag{8.2.3}$$

Thus, if $\rho(0) = \rho_*$, then we will have a time-symmetric evolution and no thermodynamic arrow of time. (In fact, the universe will always remain in state ρ_*.) But in general, for an unstable nonequilibrium-fluctuation state ρ, we will have that $\mathcal{K}\rho \neq \rho$. Therefore it is enough to assume that the universe began (at $t = 0$) in one of these states, and we will have a time-asymmetric evolution and a thermodynamic arrow of time. This is because the initial entropy is $S < S_*$, and therefore there will be growth of entropy toward both the past and the future of $t = 0$, since entropy will try to reach the equilibrium entropy in both directions. [In the exceptional case that the initial nonequilibrium unstable state is such that $\mathcal{K}\rho(0) = \rho(0)$ at $t = 0$, it would be $\mathcal{K}\rho \neq \rho$ at a different time, close to $t = 0$, which can, in addition, be taken to be the origin of time in (8.1.7).] Then it is enough to suppose that the universe began in a nonequilibrium unstable state to obtain the thermodynamic arrow of time and the second law of thermodynamics, if we conventionally only consider times $t \geq 0$ (and conventionally name this period the "future" of $t = 0$). This low-entropy, initial state of the universe could be considered to be a fluctuation. In fact, irregular fluctuations of the equilibrium entropy are

present in systems with a finite number of particles (Landau and Lifshitz, 1958), but they vanish if this number goes to infinity. Then, these fluctuations cannot be considered if we work with a distribution ρ in Liouville space, as we have done in this work, because these distributions are probabilities computed assuming an infinite number of particles (or an infinite number of copies of the system).

Fluctuations can be introduced in several ways: for example, (i) using Boltzmann entropy as in Lebowitz (1994) i or (ii) working in a rigged space, where the distribution corresponding to a finite number of particles, namely ρ's, built using a finite number of Dirac's deltas, can be considered, *etc.* We will not discuss this subject further here. Regarding this solution to the problem of the initial low-entropy state of the universe, it could be argued that the initial fluctuation is very unlikely to occur Prigogine, 1980), since the universe is very large, perhaps even infinite. Nevertheless, we will prove in the next section, that this conjecture is unnecessary, since the initial instability is naturally produced by the expansion of the universe, so that actually no fluctuations are needed. This is why we do not discuss fluctuations in this work.

For isolated subsystems within the Universe, time asymmetry can be obtained in a similar way. In fact, we are used to imagining that these subsystems (a Gibbs ink drop spreading in a glass of water, perfume spreading in a room, and so on) begin in an unstable initial state with low entropy (a concentrated ink drop, all the perfume inside the bottle, *etc.*). However, these initial states are always produced, not by unlikely fluctuations, but by external agencies (*i.e.*, the ink or perfume factories). They use energy to produce these concentrations which they obtain from other subsystems in unstable initial states (chemically unstable coal or nuclear-unstable isotopes, *etc.*) and which, in turn, obtain their energy, via a chain of unstable states (such as those of the stars), ultimately from the universe's initial, unstable state. Therefore, we conclude that all time-asymmetric processes have a cosmological origin. The only difference is that, in the case of a subsystem, we have a reason to consider only times $t \geq t_0$, with the time $t_0 = 0$ being the time of creation of the initial unstable state of the subsystem, since time $t < t_0$ corresponds to a period before the creation of the unstable states by the external agency (the concentration period of the ink drop or perfume), where the subsystem is not isolated. The subsequent diffusion of the ink drop, perfume, *etc.*, will produce the growth of thermodynamic entropy. The quantum equivalent of this reasoning may be found in Castagnino *et al.* (1996), Castagnino and Laura (1983), Laura and Castagnino (1997), Castagnino *et al.* (1997). The creation of a low-entropy state is therefore produced either by an initial fluctuation, as in the case of the universe (however, we will see

in the next subsection that the initial fluctuation is not necessary, or by an external agency, as in the case of subsystems within the universe. Thus, neglecting fluctuations for a moment, we will consider the appearance of a low-entropy state not produced by an external agency to be a "conspiracy." We can then conclude that conspiracies do not exist in nature. In fact, let us consider a system in a low-energy, unstable state produced by external agencies (*e.g.*, a shop with glasses, plus elephant). Any process within the system will produce a growth of entropy (for example, when the elephant enters the store and breaks all the glasses). This is an irreversible process. In fact, its time-reversed process (the film of the motion of the elephant played backward) is full of conspiracies and therefore does not exist in nature. Irreversibility, therefore, can also be explained in this way by our formalism.

8.3. The Problem of the Coordination of the Arrows of Time

Now we must solve the second problem which we stated by the following question: (ii) *Why do all the arrows of time point in the same direction?* We would also like to show that the initial fluctuation is not strictly necessary. To solve these problems we shall consider the cosmological arrow of time, that is, the growth of the radius or scale factor of the universe a, to be the master arrow of time, defining the direction of all the others. First, we will show that the thermodynamic arrow of time, that is, the tendency to obtain a final equilibrium, points in the same direction as the master arrow. Let S^* be the equilibrium entropy and let $S(t)$ be the actual entropy of the matter and radiation within the universe at time t. The entropy gap

$$\Delta S = S_* - S(t) \tag{8.3.1}$$

would be minus the conditional entropy $-H_c(\rho|\rho^*)$, according to equation (3.6.2), in full agreement with general relativity, if we take into account the change of the universe's temperature as explained in Section 7, namely

$$\Delta S = \int_X \rho(x) \log \frac{\rho(x)}{\rho_*(x)}\, dx \tag{8.3.2}$$

where $\rho(t, x)$ and $\rho^*(x)$ are the corresponding local distribution functions, and X is the phase space $x \in X$ a point of this space. The distribution functions are normalized as

$$\int_X \rho\, dx = 1, \qquad \int_X \rho^*\, dx = 1 \tag{8.3.3}$$

Now we can take, as in equation (5.3.24),

$$\rho(t) = \rho_* + (\rho_1 + \rho_2 e^{-\partial t/2}) e^{-\partial t/2} = \rho_* + \rho_\Delta e^{-\partial t/2} \tag{8.3.4}$$

where the second term on the rhs is a sort of correction around the equilibrium term, with a damping factor which has a characteristic time $\approx \gamma^{-1}$. We will only consider the universe's evolution after the decoupling time, when the universe is matter-dominated, and $\gamma^{-1} = t_{NR}$ is the characteristic time of nuclear reactions, which cause the matter within the stars to evolve toward thermal equilibrium with the cosmic microwave background. Equation (8.3.4) can be considered merely as a phenomenological equation, which can be obtained if we use coarse-graining techniques and we neglect the Zeno and Khalfin effects; however, we know that there is a rigorous way to eliminate these effects, by using the rigged Hilbert space formalism (extended dynamics techniques) as in (5.3.24). $\rho_{\Delta=\rho_1+\rho_2} e^{-\partial t/2}$ is normalized as

$$e^{-\partial t/2} \int_X \rho_\Delta \, dx = \int_X \rho \, dx - \int_X \rho_* \, dx = 0 \tag{8.3.5}$$

This normalization is also a consequence of equation (5.3.14′). We will take $|\rho_\Delta| << \rho_*$, or $t >> \gamma^{-1}$; in other words, the fluctuation is small compared to the equilibrium distribution function. Then, on expanding the logarithm and neglecting unimportant terms, as in (6.5.2), the entropy gap ΔS reads

$$\Delta S \approx e^{-\gamma t} \int_X \frac{\rho_\Delta^2}{\rho^*} \, dx > 0 \tag{8.3.6}$$

Thus, when $\gamma = 0$, the growth of entropy variation disappears. To compute the time derivative of ΔS, let us use the model of (5.3.31) and (5.4.17), described at the end of Section 5.3. Then the last equation reads

$$\Delta S \approx e^{-\gamma t} \int_X \frac{T^{3/2}}{Z} e^{\omega/T} \rho_\Delta^2 \, dx \tag{8.3.7}$$

where we have explicitly shown the time variation in the first exponential function and $T(t)$. The rest of the quantities are time constant, since we can neglect the second time-varying term of ρ_Δ with respect to the first, time-constant one (otherwise we could keep both terms with a small modification to the formulas). ρ_1 and ρ_2 are independent of the temperature because they are related only to the nuclear reaction processes. From Castagnino and Laura (1997) [or from (6.A.7) and (6.A.8)] we can introduce a reasonable simplification by supposing that the only important values of the last integral are those around ω_1, the characteristic energy of the nuclear processes. Then

$$\Delta S = C e^{-\gamma t} T^{3/2} e^{\omega_1/T} \tag{8.3.7}$$

where C is a time-independent constant. The temperature evolution will be

dominated by the radiation within the universe and will therefore follow equation (7.3.5) so that

$$\Delta S = C' e^{-\gamma t} a^{-3/2} e^{\omega a / T_0 a_0} \tag{8.3.8}$$

where C' is another time-independent constant. Now we can compute the time derivative, which reads

$$\Delta \dot{S}(t) = \left(-\gamma - \frac{3\dot{a}}{2a} + \frac{\omega_1 \dot{a}}{T_0 a_0} \right) \Delta S \tag{8.3.9}$$

where $\dot{a}/a = H(t) \approx t_U^{-1}$ is the Hubble coefficient. Since we are in the matter-dominated period, we have

$$a = a_0 \left(\frac{t}{t_0} \right)^{2/3} \tag{8.3.10}$$

and thus

$$\Delta \dot{S} = \left(-\gamma - t^{-1} + \frac{2\omega_1}{3T_0 t_0} \left(\frac{t_0}{t} \right)^{1/2} \right) \Delta S \tag{8.3.11}$$

Equation (8.3.8) shows two antagonistic effects (Fig. 6). The universe's gravitational field, embodied in the positive coefficient (and in the term t^{-1}), is the external agency which mostly tries to take the system away from equilibrium, whereas the nuclear reaction, embodied in γ, tries to convey the system toward equilibrium (but the gravitational term t^{-1} tries to establish equilibrium). These two effects are equal at critical times t_{cr} such that

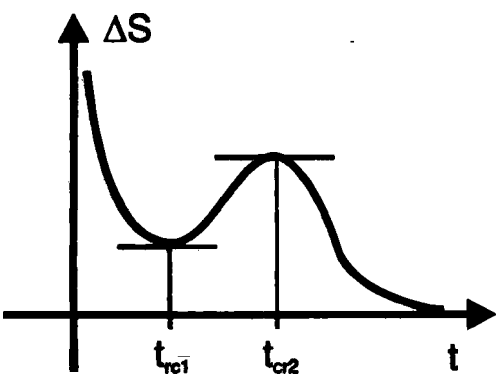

Fig. 6. Plot of ΔS showing the minimum and the maximum.

$$\gamma t_0 + \left(\frac{t_0}{t_{cr}}\right) = \frac{2\omega_1}{3T_0}\left(\frac{t_0}{t_{cr}}\right)^{1/3} \tag{8.3.12}$$

Usually this equation will have two pnoindeno roots $t_{cr1} < t_{cr2}$ (Fig. 7). It is premature to give physical numerical values to the parameters of the model. In fact, this model is extremely simplified, since it is based on a homogeneous space geometry, while the decaying processes are produced within the stars. So what we really need is an inhomogeneous geometry to properly describe the phenomenon. However, with reasonable numerical values (essentially taking $\omega_1 >> T_0$, $\gamma^{-1} \approx t_0$), we can obtain the following conclusions;

(a) The first root is in the region $t << t_0$, so that the first term on the lhs of the last equation can be neglected to obtain $t_{cr1} = t_0(3T_0/2\omega_1)^{3/2}$ (this quantity, with minus sign, gives the third negative root). For these times the entropy gap has a minimum. (b) The second root is in the region $t >> t_0$, so that the second term on the lhs can be neglected to obtain $t_{cr2} = (2\omega_1 t_{NR}/3T_0 t_0)^3\ t_0$. For these times the entropy gap has a maximum.

Then we can conclude the following:

(i) If $t < t_{cr1}$, then the second term on the lhs of equation (8.3.12) dominates $\Delta\dot{S} < 0$, and there is a big value for the entropy gap which is rapidly thermalized.

(ii) If $t_{cr1} < t < t_{cr2}$, then $\Delta\dot{S} > 0$, the rhs of (8.3.12) dominates, and there will be a growth of the entropy gap, produced by the universe's expansion, which drives the universe away from equilibrium. There will be a growth of complexity during this period, allowing particles, atoms, molecules, galaxies, stars, planets, and living beings to appear.

(iii) On the contrary, if $t > t_{cr2}$, then $\Delta\dot{S} < 0$, the first term on the l.h.s. of (8.3.12) dominates, the entropy gap will diminish, and the universe goes toward its final equilibrium state, driven by the nuclear reaction processes,

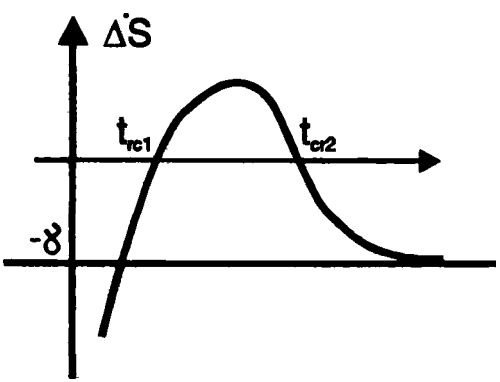

Fig. 7. Plot of $\Delta\dot{S}$ showing the two roots.

in agreement with (Davies, 1994). All structure within the universe decays and disappears. Therefore

$$\lim_{t\to\infty} \Delta S = 0 \tag{8.3.13}$$

Numerical estimates show that $t_{cr1} << t_0 << t_{cr2}$ in such a way that the first period can be, in some sense, neglected since this is the period is before the decoupling time. [Also, $t_{cr2} >> t_0$ as in (Reeves, 1993)]

(iv) Equation (8.3.8) shows how the universe expansion creates, in a continuous manner, the universe's instability and complexity. This fact makes the initial fluctuation hypothesis unnecessary. The instability is created toward the future, defined as the direction of the universe's expansion. Equation (8.3.8) also shows how the local nuclear reactions try to restore equilibrium, *in the same time direction.* The thermodynamic arrow of time is the local tendency to thermodynamic equilibrium (and not the total entropy gap growth). Therefore the thermodynamic arrow coincides with the cosmological arrow.

(v) All this reasoning is also valid before the recombination time. In that case we would use a much bigger γ, because in that period we would have to consider reactions with much smaller characteristic times, smaller in fact than the recombination time. Since the period $t < t_{cr1}$ probably lies in this period, then perhaps the universe reaches also a thermodynamic equilibrium, and we can use the arguments of Zeh (1989), to show that the electromagnetic arrow of time coincides with the cosmological one. Also, the damping factor $e^{-\gamma t}$, can be obtained if we consider a pole in the lower half-plane of the unphysical sheet of the complex energy plane; thus we must use the upper rim of the positive real axis cut and retarded solutions therefore as in electromagnetism (Zadella and Zudin, 1996).

(vi) Finally, we ourselves are merely a subsystem with unstable initial state, produced by external agencies, just as the ink drop or the bottle of perfume, and therefore our thermodynamic arrow, which can be identified with our psychological arrow, is aligned with the cosmological one. Therefore all arrows of time point in the same direction (see also Aquilano and Castagnino, 1996a.

(vii) Hence we have given, in a mathematical formalism, answers to the two main questions concerning the universe's time asymmetry. We believe that the solution presented is quite satisfactory, and that all that remains is to study many more physical examples using the extended dynamics method, and to add some mathematical refinement (such as that of Castagnino and Gunzig, 1997). Once such examples have been studied and these refinements made, we will have a definitive and rigorous answer to these long-standing fundamental questions.

9. CONCLUSIONS

As a result of all this explanation and discussion, we believe that we can draw the following conclusions:

(i) There is no compelling, local-physical motivation to prefer one choice of technique over the other. Therefore, it is not easy to see how to find a local cross-experiment to settle the matter. Probably this cross-experiment does not even exist, so that both techniques really are physically equivalent.

(ii) Coarse-graining is more "physical," since it works directly in the usual Hilbert space. The price to pay is the introduction of an object which is actually alien to the theory, the projector. This projector is essentially arbitrary, so that coarse graining will not have a deep physical meaning unless or until a natural graininess is found.

(iii) Extended dynamics is more "mathematical" in that it works in rigged Hilbert space. After paying this price, however, we are not forced to introduce any object alien to the theory. In this sense, extended dynamics is purer and cannot really be distinguished from no-graining. Extended dynamics therefore appears to be conceptually superior even if, from the operational point of view, coarse-graining might be considered more convenient. In any case, extended dynamics also has its ambiguities, *e.g.*, the choice of the test function space, even if it seems more probable that we would find a canonical choice of this space in the future than a canonical choice of the coarse-graining projector.

(iv) For conceptually difficult aspects of physics, such as cosmology or quantum measurement theory, it is advisable to use extended dynamics, since it is conceptually superior to coarse-graining. Perhaps one should still hold out hope for a global cross-experiment which may ultimately base the appropriateness of one technique or the other on sound cosmological reasoning.

ACKNOWLEDGMENTS

This work was partially supported by grants CI1*-CT94-0004 of the European Community, PID-0150 of CONICET (National Research Council of Argentina), EX-198 of Buenos Aires University, and by the British Council and the OLLAM and Antorchas Foundations.

REFERENCES

Anosov, D. V. (1963). *Soviet Mathematics-Doklady*, **4**, 1153.
Antoniou, I., and Prigogine, I. (1993). *Physica A*, **192**, 443.
Antoniou, I., and Tasaki, S. (1991). *Physica A*, **190**, 303.
Antoniou, I., and Tasaki, S. (1993a). *International Journal of Quantum Chemistry*, **46**, 427.

Antoniou, I., and Tasaki, S. (1993b). ULB preprint.
Antoniou, I., Laura, R., Tasaki, S., and Suchanecki, N. (1995). ULB preprint.
Aquilano, R., and Castagnino, M. (1996a). *Modern Physics Letters A*, **11**, 755.
Aquilano, R., and Castagnino, M. (1996b). *Astrophysics and Space Science*, **238**, 159.
Arnold, V. I., and Avez, A. (1968). *Ergodic Problems of Classical Mechanics*, Benjamin, New York.
Balazs, N., and Voros, A. (1990). *Annals of Physic*, **199**, 123.
Balescu, R. (1975). *Equilibrium and Non-Equilibrium Statistical Mechanics*, Wiley, New York.
Ballentine, L. E. (1990). *Quantum Mechanics*, Prentice-Hall, Englewood Cliffs, New Jersey.
Bohm, A. (1979). *Quantum Mechanics: Foundations and Applications*, Springer-Verlag, Berlin.
Bohm, A., and Gadella, M. (1989). *Dirac Kets, Gamow Vectors, and Gel'fand Triplets*, Springer-Verlag, Berlin.
Caldeira, A., and Leggett, A. (1995). *Physical Review D*, **31**, 1059.
Castagnino, M., and Gunzig, E. (1997). A landscape in time-asymmetry, *International Journal of Theoretical Physics*, in press.
Castagnino, M., and Laura, R. (1983). The cosmological essence of time asymmetry, in *Proceedings SILARG VIII*, W. Rodrigues, ed., World Scientific, Singapore.
Castagnino, M., and Laura, R. (1997). A minimal irreversible quantum mechanics, *Physical Review A*, in press.
Castagnino, M., Gaioli, F., and Gunzig, E. (1996). *Foundations of Cosmic Physics*, **16**, 221.
Castagnino, M., Gadella, M., Gaioli, F., and Laura, R. (1997). Gamov vectors and time asymmetry, *Fortschritte der Physik*, submitted.
Castagnino, M., Gunzig, E., and Lombardo, F. (1995). *General Relativity and Gravitation*, **27**, 257.
Davies, P. C. (1994). Stirring up trouble, in *Physical Origin of Time Asymmetry*, J. Halliwell *et al.*, eds., Cambridge University Press, Cambridge.
Gadella, M., and Rubin, G. E. (1996). *International Journal of Quantum Chemistry*, **58**, 441.
Gel'fand, I., and Shilov, G. (1968). *Generalized Functions*, Academic Press, New York.
Halmos, P. R. (1956). *Lectures on Ergodic Theory*, Chelsea, New York.
Hillery, M., O'Conell, R. F., Scully, M. D., and Wigner, E. P. (1984). *Physics Reports*, **106**, xxx.
Hu, B. L., Paz, J. P., and Zhang, Y. (1992a). *Physical Review D*, **45**, 2843.
Hu, B. L., Paz, J. P., and Zhang, Y. (199b7). Quantum origin of noise on fluctuation in cosmology, in *Proceedings Conference on the Origin of Structure in the Universe*, World Scientific, Singapore.
Hu, B. L., Paz, J. P., and Zhang, Y. (1993). *Physical Review D*, **47**, 1776.
Landau, L. D., and Lifshitz, E. M. (1958). *Statistical Physics*, Pergamon Press, Oxford.
Lasota, A., and Mackey, M. C. (1985). *Probabilistic Properties of Deterministic Systems*, Cambridge University Press, Cambridge.
Laura, R., and Castagnino, M. (1997). Functional approach for quantum systems with continuous spectrum, *Physical Review E*, in press.
Lebowitz, J. L (1994). Time's arrow and Boltzmann's entropy, in *Physical Origin of Time Asymmetry*, J. Halliwell *et al.*, eds., Cambridge University Press, Cambridge.
Mackey, M. C. (1989). *Reviews of Modern Physics*, **61**, 981.
Messiah, A. (1962). *Quantum Mechanics*, North-Holland, Amsterdam.
Misra, B., Prigogine, I., and Courbage, M. (1979). *Physica A*, **98**, 1.
Ordonez, A. (1997). Rigged Hilbert spaces associated with Misra–Prigogine–Courbage theory of irreversibility, *Physica A*, in press.
Prigogine, I. (1980). *From Being to Becoming: Time and Complexity in Physical Sciences*, Freeman, San Francisco.

Prigogine, I. (1993). Time, dynamics, and chaos, in *Nobel Conference XXVI, Chaos: The New Science*, John Holte, ed., Gustavus Adolphus College, St. Peter, Minnesota.
Prigogine, I., and Petrosky, T. (1993). *Physics Letters A*, **182**, 5.
Prigogine, I., George, C., Henin, F., and Rosenfeld, L. (1980). *Chemica Scripta*, **4**, 5.
Reeves, H. (1993). The growth of complexity in expanding universes, in *The Anthropic Principle, Proceedings Second Venice Conference on Cosmology*, F. Bertolo and U. Cino, eds., Cambridge University Press, Cambridge.
Rochlin, V. A. (1969). *American Mathematical Society Translations (2)*, **39**, 1.
Roman, P. (1965). *Advanced Quantum Theory*, Addison-Wesley, Reading, Massachusetts.
Sachs, R. G. (1987). *The Physics of Time Reversal*, University of Chicago Press, Chicago.
Schild, P. (1979). *The Theory of Bernoulli Shift*, University of Chicago Press, Chicago.
Sudarshan, E. C. G., Chiu, C. B., and Gorini, V. (1978). *Physical Review D*, **18**, 2914.
Tabor, G. (1980). *Chaos and Integrability in Non-linear Dynamics*, Wiley, New York.
Tolman, R. C. (1987). *Relativity, Thermodynamics, and Cosmology*, Dover, New York.
Voigt, J. (1981). *Communications in Mathematical Physics*, **81**, 31.
Walter, P. (1982). *An Introduction to Ergodic Theory*, Springer-Verlag, Berlin.
Zeh, D. (1989). *The Physical Bases of the Direction of Time*, Springer-Verlag, Berlin.
Zwanzig, R. W. (1960). *Chemical Physics*, **33**, 1388.
Zwanzig, R. W. (1961). Statistical mechanics of irreversibility, in *Lectures in Theoretical Physics III*, W. E. Britten *et al.*, eds., Interscience, New York.