

---

## The Quantum Mechanical Arrow of Time

The dynamics of probability distributions on classical phase space, discussed under various aspects in Chap. 3, may be *formally* translated into quantum mechanics by means of the canonical quantization rules. Many authors of standard textbooks therefore maintain that the foundation of irreversibility in quantum mechanics is identical to that in classical physics. There could then only be quantitative differences arising from different spectral properties of the ‘corresponding’ Liouville operators. However, this approach to statistical quantum mechanics completely ignores the fundamental interpretational differences of concepts that formally correspond to one another (such as probability distributions and density operators – see Sect. 4.2). It therefore conceals essential aspects of quantum theory which may be important for irreversibility in general (recall the general discussion in the Introduction):

1. The quantum mechanical probability interpretation represents an *indeterminism* of controversial origin. Most physicists seem to regard it as an objective *dynamical* indeterminism (see Fig. 3.8), and some even as representing a fundamental arrow of time that would go beyond dynamics. Others have instead suggested that one may *explain* the unpredictability of quantum mechanical measurement results in terms of conventional statistical arguments, viz., by means of thermal fluctuations that are related to the amplification process which leads to macroscopic outcomes. If, however, this question is circumvented by interpreting the wave function as representing ‘human knowledge as an intermediate level of reality’ (Heisenberg 1956), this may exclude *any* possibility of a dynamical analysis, while Maxwell’s demon, discussed in Sect. 3.3.2, would return through the quantum back door. Therefore, the foundation of irreversibility seems to be intimately related to the *interpretation* of quantum theory (see Sects. 4.3 and 4.6). In order to clarify this situation as far as possible, one first has to analyze the dynamical formalism as it is actually *used*, and thus empirically justified.

2. The quantum theory is *kinematically nonlocal*. For example, the generic many-particle wave function  $\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ , which represents a ‘pure’ quantum state, describes *quantum correlations* that are *not* due to incomplete information (even though they may *lead to* statistical correlations in measurements). Similarly, a state of quantum field theory is given by a wave functional of fields which are defined all over space. This ‘entanglement’ is a direct consequence of the superposition principle. In quantum theory, *the state of the whole does not define states of its parts*. This is in fundamental contrast to the completely determined many-particle state of classical mechanics: a point in phase space (that is, a definite state) remains a point when projected onto a subsystem. The kinematical indeterminacy of the parts in quantum theory describes a non-trivial ‘wholeness’ of Nature, which cannot, as in classical physics, be interpreted as a mere *dynamical* interconnectedness (that may lead to *statistical* correlations in an incomplete description). Quantum nonlocality is not just a ‘spooky action at a distance’ that would affect *hidden local states*. Moreover, the absence of well defined subsystem states has nothing to do with Heisenberg’s uncertainty (or ‘indeterminacy’) relations, which signal the limited validity of *classical* concepts for describing physical states. They apply even when the true and deterministically evolving quantum states are *certain* (‘pure’).

Consequences of these basic differences between classical and quantum statistical physics will be discussed after their formal analogy has been set up in Sect. 4.1.

## 4.1 The Formal Analogy

### 4.1.1 Application of Quantization Rules

The *formal* transition from classical to quantum statistical mechanics can be based on the ‘canonical quantization rules’, which replace functions of state  $a(p, q)$  by ‘corresponding’ operators  $A = a(P, Q)$ , and Poisson brackets between them by commutators. For example, the Liouville equation (3.26) transforms as

$$i \frac{\partial \rho_\Gamma}{\partial t} = i \{H, \rho_\Gamma\} =: \hat{L} \rho_\Gamma \quad \longrightarrow \quad i \frac{\partial \rho}{\partial t} = [H, \rho] =: \hat{L} \rho . \quad (4.1)$$

It is then called the *quantum Liouville* or *von Neumann equation*. The classical probability densities  $\rho_\Gamma(p, q)$  are thus replaced by *density operators*  $\rho$ . The caret is here used to distinguish the new operators, which act on the quantum mechanical Hilbert space operators (such as density operators), from these Hilbert space operators themselves. In the formal analogy, the new ‘superoperators’ (as they are sometimes called) correspond to the operators that were defined in Sect. 3.1.2 as acting on probability densities.

The Hilbert space operators form a *new* Hilbert space if an inner product  $\langle\langle\rho_1, \rho_2\rangle\rangle := \text{Trace}\{\rho_1^\dagger \rho_2\}$  is defined for them in analogy to the inner product  $\langle\rho_{\Gamma 1}, \rho_{\Gamma 2}\rangle = \int \rho_{\Gamma 1}^*(p, q) \rho_{\Gamma 2}(p, q) dp dq$  for classical probability densities on  $\Gamma$ -space – see the text above (3.27).

Furthermore, all mean values  $\bar{a}$  of functions of state  $a(p, q)$ , defined with respect to probability densities  $\rho_\Gamma(p, q)$ , are replaced by *expectation values*  $\langle A \rangle$  of corresponding ‘observables’  $A$ :

$$\bar{a} := \int a(p, q) \rho_\Gamma(p, q) dp dq \longrightarrow \langle A \rangle := \text{Trace}\{A\rho\} . \quad (4.2)$$

Since (4.2) implies

$$\overline{\ln \rho} \longrightarrow \langle \ln \rho \rangle = \text{Trace}\{\rho \ln \rho\} , \quad (4.3)$$

the quantum mechanical entropy functional corresponding to the ensemble entropy  $S_\Gamma$  becomes *von Neumann’s entropy*,

$$S[\rho] := -k \text{Trace}\{\rho \ln \rho\} . \quad (4.4)$$

However, in spite of this formal analogy, a density operator can no longer be interpreted as representing an ensemble of states that would define ensembles of values for all functions of state (see below).

The dynamics of the statistical operators, defined by the right-hand equation of (4.1), is unitary, with a formal solution

$$\rho(t) = U(t) \rho(0) U^\dagger(t) , \quad (4.5)$$

and  $U(t) = \exp(-iHt)$  for time-independent Hamiltonians. This dynamical form warrants conservation of von Neumann entropy (4.4) under the von Neumann equation,

$$\begin{aligned} \text{Trace}\{\rho(t) \ln \rho(t)\} &= \text{Trace}\{U(t) \rho(0) U^\dagger(t) U(t) \ln \rho(0) U^\dagger(t)\} \\ &= \text{Trace}\{\rho(0) \ln \rho(0)\} . \end{aligned} \quad (4.6)$$

Since *classical* determinism (the conservation of probabilities along individual trajectories) may also be described in the form of a unitary time-dependence of probability distributions [see (3.26)], the formal argument in (4.6) may also be applied to classical ensemble dynamics.

The square of the Hilbert space norm of a density operator,

$$\|\rho\|^2 := \langle\langle\rho, \rho\rangle\rangle = \text{Trace}\{\rho^2\} = \langle\rho\rangle , \quad (4.7)$$

defines a *linear measure* of negentropy (see footnote 3 of Chap. 3). It is also conserved under the unitary dynamics (4.1). The corresponding linear entropy is often defined as  $S_{\text{lin}} = \langle(1 - \rho)\rangle$  (such that  $0 \leq S_{\text{lin}} < 1$ ). In contrast to this linear entropy, which uses the Hilbert space norm of operators, the

probability norm,  $\text{Trace}\{\rho\} = \langle 1 \rangle = 1$ , characterizes a Banach space of *trace class operators* (those with non-vanishing trace). It is preferentially used in open systems quantum mechanics (Sect. 4.4), since total probability must be conserved even under phenomenological stochastic equations of motion that describe an increase in ensemble entropy.

In further formal analogy to classical ensemble mechanics, any coarse-grained (or relevant) information measured by  $\text{Trace}\{(\hat{P}\rho) \ln(\hat{P}\rho)\}$  is in general *not* conserved under a unitary transformation. The Zwanzig projection operators  $\hat{P}$  are once again idempotent operators on the Hilbert space of density operators, with the additional properties  $\text{Trace}\{\hat{P}\rho\} = 1$  and positive  $\hat{P}\rho$  for all  $\rho$  – just as in Sect. 3.2.

Statistical operators (density operators)  $\rho$  may be *represented* by various ensembles of wave functions  $\psi_\alpha$  with probabilities  $p_\alpha$  in the form  $\rho = \sum_\alpha |\psi_\alpha\rangle p_\alpha \langle \psi_\alpha|$  (see Sect. 4.2). In the diagonal form of  $\rho$ , where the eigenstates  $\psi_\alpha$  form an orthonormal set,  $\|\rho\|^2$  is given by the sum  $\sum_\alpha p_\alpha^2$ . Its conservation (or that of  $\langle \ln \rho \rangle = \sum p_\alpha \ln p_\alpha$ ) thus reflects the *individual* conservation of these diagonal elements in the moving basis  $\psi_\alpha(t)$  – in analogy to the conservation of a comoving phase space volume in deterministic classical mechanics.

Matrix elements of the density operator with respect to a *random* basis  $\{\phi_n\}$ ,

$$\rho_{mn} = \sum_\alpha \langle \phi_m | \psi_\alpha \rangle p_\alpha \langle \psi_\alpha | \phi_n \rangle = \sum_\alpha c_{\alpha m} p_\alpha c_{\alpha n}^* \quad (4.8)$$

if  $\psi_\alpha = \sum_m c_{\alpha m} \phi_m$ , are in general small for  $m \neq n$  because of random phases of the coefficients in the sum over  $\alpha$ . Pauli (1928) referred to this *random phase approximation* when he neglected off-diagonal matrix elements while deriving his master equation (4.18) below. However, they may in general be small individually in spite of amounting to a significant effect as a whole. In contrast, the complete *neglect* of off-diagonal elements *in a certain basis*,

$$\hat{P}_{\text{diag}} \rho_{mn} := \rho_{mm} \delta_{mn} , \quad (4.9)$$

defines the most important Zwanzig projection of quantum statistical mechanics. It regards these off-diagonal elements (or any interference between the states of this basis) as ‘irrelevant’ for all practical purposes, although it does not assume them to vanish. So it has nothing to do with the usual diagonalization of Hermitean operators in their eigenrepresentation. The inequality

$$\text{Trace}\{\hat{P}_{\text{diag}} \rho \ln(\hat{P}_{\text{diag}} \rho)\} = \sum_n \rho_{nn} \ln \rho_{nn} \leq \text{Trace}\{\rho \ln \rho\} = \sum_{mn} \rho_{mn} (\ln \rho)_{nm} \quad (4.10)$$

is called *Klein’s lemma* – see (3.35). It is a consequence of the fact that  $\hat{P}_{\text{diag}}$  is a genuine projection operator (see Sect. 3.2).

An obvious (weaker) generalization of (4.9) is

$$\hat{P}_{\text{semidiag}} \rho := \sum_n P_n \rho P_n , \quad (4.11)$$

where  $\{P_n\}$  (no caret!), with  $P_m P_n = P_m \delta_{mn}$ , is a complete set of projection operators on mutually orthogonal subspaces of the Hilbert space of quantum states. In quantum field theory, projections on ‘unitarily inequivalent’ separable subspaces of Hilbert space, sometimes even regarded as ‘distinct Hilbert spaces’, are often chosen for this purpose. However, these decompositions of non-separable Hilbert spaces are no less arbitrary than any other  $\hat{P}_{\text{semidiag}}$  (though often *useful* in the case of large numbers of effective degrees of freedom). If imposed axiomatically, the relevance concept (4.11) may represent a *superselection rule* (Wick, Wightman and Wigner 1952, Jauch 1968, Hepp 1972). This observation suggests that proposed superselection rules are similarly based on some *dynamical* robustness like the ‘thermodynamically macroscopic’ variables of Chap. 3 that are usually assumed as ‘given’ – a possibility that will be further investigated and confirmed in Sect. 4.3.

#### 4.1.2 Master Equations and Quantum Indeterminism

The Hamiltonian of a quantum mechanical system is often written in the form  $H = H_0 + H_1$  in order to derive a master equation in terms of a perturbation expansion with respect to  $H_1$ . However, the main purpose of this split Hamiltonian is to define a relevance concept of type (4.9) or (4.11) by means of the eigenbasis of  $H_0$ . It may then (but need not) be *further* used for a time-dependent perturbation expansion with respect to the off-diagonal elements of  $H$  in this representation.

The dynamics of the ‘relevant’ part  $\hat{P}_{\text{diag}}\rho$  is the dynamics of the diagonal elements of  $\rho$ . According to (4.1) one has in any representation (now writing  $\hat{P}_{\text{diag}} = \hat{P}$  for short)

$$\begin{aligned} i \frac{d\rho_{mm}}{dt} &= \sum_n (H_{mn}\rho_{nm} - \rho_{mn}H_{nm}) \\ &\equiv \sum_{n(\neq m)} (H_{mn}\rho_{nm} - \rho_{mn}H_{nm}) \hat{=} \hat{P}\hat{L}(1 - \hat{P})\rho. \end{aligned} \quad (4.12)$$

Since the diagonal matrix elements of  $\rho$  do not contribute to the RHS, the first term of Zwanzig’s pre-master equation (3.44), representing  $\hat{P}\hat{L}\hat{P}$ , vanishes for this relevance concept. The terms remaining in (4.12) describe the coupling to the ‘irrelevant’ off-diagonal elements, and demonstrate that the diagonal elements are dynamically autonomous only in the trivial case (see footnote 6 of Chap. 3 regarding the quantum Zeno effect). Because of the formal analogy, the rest of Zwanzig’s method can then be applied, provided the required approximations are valid. The propagator  $\exp[-i(1 - \hat{P})\hat{L}\tau]$ , occurring in the operator  $\hat{G}_{\text{ret}}$  of the Markovian approximation (3.48), defines here a closed but highly non-trivial dynamics of the off-diagonal elements of  $\rho_{mn}$ .

Pauli’s master equation can now be obtained from (3.48) and (3.45) by using a perturbation expansion in terms of the off-diagonal elements of the

Hamiltonian for calculating  $\hat{G}_{\text{ret}} = \int_0^T \hat{G}(\tau) d\tau$ . These off-diagonal elements are thus assumed to be small, although the master equation would become trivial if they vanished exactly (that is, for  $H = H_0$ ). This last remark emphasizes the *dynamical* role of the relevance concept.

Now consider the last three factors of the RHS of the integral kernel (3.45) applied to  $\rho$ :

$$(1 - \hat{P})\hat{L}\hat{P}\rho = (1 - \hat{P})[H, \hat{P}\rho] \hat{=} H_{mn}(\rho_{nn} - \rho_{mm}) \quad \text{with } m \neq n. \quad (4.13)$$

This expression depends only on the off-diagonal elements of  $H$ . The projection  $1 - \hat{P}$  is ineffective, as  $\hat{P}\hat{L}\hat{P} = 0$ . Similarly, one has for the *first* three factors of the RHS of (3.45), when applied to any matrix  $X$ :

$$\hat{P}\hat{L}(1 - \hat{P})X \hat{=} \sum_{k(\neq m)} (H_{mk}X_{km} - X_{mk}H_{km}). \quad (4.14)$$

Hence,  $\hat{G}_{\text{ret}}$  is of second and higher orders in the off-diagonal elements of  $H$ . When neglecting higher orders according to Pauli, one has to express the remaining propagator  $\exp[-i(1 - \hat{P})\hat{L}\tau]$  in (3.45) solely in terms of diagonal elements of  $H$ ,  $H_{mm} =: E_m^{(0)}$ . This means

$$e^{-i(1 - \hat{P})\hat{L}\tau} X \hat{=} e^{-i(E_m^{(0)} - E_n^{(0)})\tau} X_{mn}, \quad (4.15)$$

and one obtains

$$\begin{aligned} \hat{P}\hat{L}(1 - \hat{P})e^{-i(1 - \hat{P})\hat{L}\tau}(1 - \hat{P})\hat{L}\hat{P}\rho &\hat{=} \\ \sum_n |H_{mn}|^2 2 \cos[(E_m^{(0)} - E_n^{(0)})\tau] (\rho_{mm} - \rho_{nn}). \end{aligned} \quad (4.16)$$

This result corresponds to a Born approximation in terms of the off-diagonal elements of the Hamiltonian. The time integral required to obtain  $\hat{G}_{\text{ret}}$  according to (3.49) leads to the resonance factor

$$\int_0^T \cos[(E_m^{(0)} - E_n^{(0)})\tau] d\tau = \frac{\sin[(E_m^{(0)} - E_n^{(0)})T]}{(E_m^{(0)} - E_n^{(0)})}, \quad (4.17)$$

familiar from time-dependent perturbation theory. In the limit  $T \rightarrow \infty$ , this quotient becomes a  $\delta$ -function times  $\pi$ , and (3.48) can be written (Pauli 1928)

$$\frac{d\rho_{mm}}{dt} = 2\pi \sum_n |H_{mn}|^2 \delta(E_m^{(0)} - E_n^{(0)}) (\rho_{nn} - \rho_{mm}) =: \sum_n A_{mn} (\rho_{nn} - \rho_{mm}). \quad (4.18)$$

This *Pauli equation* is similar to other master equations, such as (3.51), while the coefficients  $A_{mn}$ , defined on the RHS, are transition rates in analogy to Boltzmann's  $w(\mathbf{p}_1\mathbf{p}_2, \mathbf{p}'_1\mathbf{p}'_2)$  of Sect. 3.1.1. If  $H_1$  contains only two-particle

interactions, the sum over  $n$  may indeed be written as a sum over particle pairs. According to the above definition, the coefficients  $A_{mn}$  conserve energy and satisfy the symmetry under collision inversion,  $A_{mn} = A_{nm}$  [see (3.7)]. Therefore, the Pauli equation conserves total probability,  $\sum_n d\rho_{nn}/dt = 0$ .

The explicit form of the Pauli equation (4.18) may be used to discuss its range of validity, which must be limited by the approximations used when deriving the general master equation (3.48). It depends here on the spectrum of the Hamiltonian, which is often discrete for quantum systems. Nonetheless, Poincaré recurrence times can be neglected *in practice* for macroscopic quantum systems. Their energy spectra are usually so dense that they do *not* lead to any observable differences compared to a continuous spectrum. Quantum systems may even exhibit ‘classical chaos’ (Habib, Shizume and Zurek 1998). On the other hand, even a continuous spectrum would not by itself justify an arrow of time (as is often claimed). The negligibility of recurrences for all times of interest – whether they exist in principle or not – applies in *both* directions of time. The physical importance of the difference between discrete and continuous spectra seems to be grossly overemphasized in mathematical foundations of irreversibility.

However, the energy  $\delta$ -function occurring in (4.18) is meaningful only inside an integral over energy  $E$ , or, as an approximation, under a sum over  $m$ . Therefore, Pauli combined groups of states with almost equal energies to form ‘cells’ (subspaces) representing a coarse-graining in order to apply a random phase approximation in the corresponding sums (see also van Kampen 1954). Erich Joos (1984) was able to show that the off-diagonal elements  $\rho_{mn}$  between states from such macroscopically different subspaces disappear by interaction with the environment (‘decoherence’ – see Sect. 4.3). This *dynamical* argument justifies Pauli’s conceptual cells and his random phase ‘approximation’.

When applied to a single initial state with  $\rho_{00}(0) = 1$ , Pauli’s equation (4.18) assumes the form of *Fermi’s Golden Rule* in the Born approximation. Replacing the sum over initial states  $n$  in (4.18) by an energy integral and a sum over all remaining quantum numbers  $\beta$ , that is,  $\sum_n \cdots \longrightarrow \sum_\beta \int \sigma_\beta(E) \cdots dE$  with a partial density of states  $\sigma_\beta(E)$ , and similarly substituting  $m \longrightarrow E', \alpha$  for the final states, one obtains for the energy-integrated diagonal elements of final states  $\alpha \neq 0$ ,  $\rho_{\alpha\alpha} := \int \rho_{E'\alpha, E'\alpha} \sigma_\alpha(E') dE'$ :

$$\frac{d\rho_{\alpha\alpha}}{dt} = 2\pi |H_{\alpha 0}(E)|^2 \sigma_\alpha(E) \quad . \quad (4.19)$$

Here,  $H_{\alpha 0}(E) := H_{\alpha E, 0E}$ , while  $\alpha$  represents a ‘decay channel’.

Although this Golden Rule (4.19) can thus be derived as an approximation from the unitary dynamics (4.12), it is mainly used to calculate *probabilities* for decay and other non-unitary ‘quantum events’ – conventionally described by a collapse of the wave function – see Sect. 4.6. (*Coherent* exponential decay according to the Schrödinger equation will be discussed in Sect. 4.5.) In contrast, Boltzmann’s probabilistic transition rates  $w(\mathbf{p}_1 \mathbf{p}_2, \mathbf{p}'_1 \mathbf{p}'_2)$  refer to *ensembles of individually deterministic* collision trajectories (distinguished by their

impact parameters). This different interpretation is facilitated by the fact that the formal concept of a density operator is already based on a probability interpretation (see Sect. 4.2). Nobody has ever been able to construct a model that would consistently explain the wave function as representing an ensemble of ‘hidden variables’. (Bohm’s theory, that *presumes* Schrödinger’s wave function, will be discussed in Sect. 4.6.) In particular, the entropy (4.4) does not contain any contribution that might represent the missing information corresponding to such an ensemble (as in Fig. 3.5 for classical measurements).

Pauli’s equation does indeed resemble Born’s original formulation of the probability interpretation (Born 1926). Born used it to describe ‘quantum jumps’ between Schrödinger’s stationary eigenstates of Hamiltonians  $H_0$  that characterize isolated microscopic systems (such as atoms).<sup>1</sup> In quantum field theory, a similar splitting of the Hamiltonian is used to define the *interaction picture*. The special role attributed to the eigenstates of  $H_0$  as representing the ‘real’ physical states, dynamically connected by discrete jumps, was historically motivated by their correspondence with Bohr’s discrete atomic electron orbits. Quantum jumps (or a ‘collapse of the wave function’) are, of course, incompatible with deterministic *trajectories in Hilbert space*, that is, with time-dependent wave functions evolving according to a Schrödinger equation. The system Hamiltonians  $H_0$  are thus assumed not to contain any interaction that would be responsible for stochastic transitions. This early attempt to objectivize the probability interpretation (or the observables used therein) by a dynamical process is therefore based on an essential approximation. (Recall the trivial result obtained for the Pauli equation in the exact energy basis!)

The general structure of the Pauli equation is preserved even when the perturbation expansion in terms of the off-diagonal elements of  $H$  (in a certain basis) is not used. This improved equation is known as Van Hove’s ‘exact’ master equation (Van Hove 1957). It represents the master equation for the Zwanzig projection (4.9) without any *further* approximation. In particular, if the chosen basis of relevance (the eigenbasis of  $H_0$ ) is the independent particle basis, the matrix elements  $H_{mn}$  appearing in the Pauli equation have to be replaced by the elements of a  $T$ -matrix, usually defined as  $T := (S - 1)/2\pi i$ , where  $S$  is the exact two-particle scattering matrix. This procedure presumes the negligibility of simultaneous many-particle collisions (just as Boltzmann’s *Stoßzahlansatz*). However, the adjective ‘exact’ for Van Hove’s equation is misleading even for a dilute gas, as it refers only to the calculation of  $\hat{G}_{\text{ret}}$ ,

<sup>1</sup> While Born may not have been using his concepts quite consistently in these early days of quantum mechanics, in his third (here quoted) paper on the probability interpretation he discussed probabilities for jumps between stationary *wave functions* – not probabilities for the occurrence of classical properties (such as particle positions). In scattering or decay ‘events’ he referred to plane waves as stationary states, which he then *associated* with particle momenta according to de Broglie’s relation. One year before the formulation of the uncertainty relations this was not recognized as being in conflict (in principle) with the position measurement at the detector.



but not to the derivation of the master equation (3.48) in its preferred basis of relevance. Similarly to the choice of subspaces in (4.11), Born's probability interpretation, when applied to *measurements*, depends on the choice of appropriate 'observables'.

In analogy to the classical  $H$ -theorem (3.10), one may again show that the entropy corresponding to the Zwanzig projection  $\hat{P}_{\text{diag}}$  never decreases under the Pauli or Van Hove equation:

$$\frac{dS[\hat{P}_{\text{diag}}\rho]}{dt} = -k \frac{d\left(\sum \rho_{mm} \ln \rho_{mm}\right)}{dt} \geq 0. \quad (4.20)$$

Evidently, this entropy depends crucially on the chosen basis for diagonalization, that is, on the specific concept of relevance used in this master equation.

Because of the formal analogy, the classical canonical distribution,  $\rho_{\text{can}}(p, q) = Z^{-1} \exp[-H(p, q)/kT]$ , now becomes a canonical density operator,  $\rho_{\text{can}} = Z^{-1} \exp(-H/kT)$ . It can be derived precisely as in (3.19) by maximizing the entropy  $S[\rho]$  under the constraint of fixed mean energy and probability norm. The so-called 'new statistics' (Bose or Fermi statistics) in terms of apparent *particles* is obtained when evaluating this canonical density operator in terms of quantum states of free *fields* – conveniently in the *occupation number representation*. Only when expressed in terms of particle states does it appear as a new method for counting them. The success of quantum statistics is indeed one of the strongest arguments against *particles* (in their original sense of pointlike objects in space, distinguishable by their trajectories) as a fundamental kinematical concept.

This conclusion, that fields rather than particles have to be quantized even for fields that never appear classically (such as spinor fields – see Zeh 2003), is also supported by the absence of Gibbs' self-mixing entropy (see footnote 2 of Chap. 3). The empirically correct measure on phase space,  $d^{3N}p d^{3N}q/h^{3N}N!$ , may then be obtained, for example, in the partition function  $Z$  for a grand canonical ensemble,  $p_{E,N}(\mu, T) = \exp[-(E - \mu N)/kT]$ . If this expression is evaluated by means of the familiar textbook approximation in the occupation number representation  $|\{n_{\mathbf{k}}\}\rangle$  for spatial wave modes (often incorrectly regarded as 'single-particle' *wave functions*) with wave numbers  $\mathbf{k} = \mathbf{p}/\hbar$  on a large space volume  $V$ , one obtains for dilute gases – where  $N = \sum_{\mathbf{k}} n_{\mathbf{k}}$  and  $E = \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} n_{\mathbf{k}}$ , with  $\varepsilon_{\mathbf{k}} = p(\mathbf{k})^2/2m$  and  $\varepsilon_{\mathbf{k}} - \mu \gg kT$ :

$$\begin{aligned} Z(\mu, T) &= \sum_{\{n_{\mathbf{k}}\}} \exp\left[-\sum_{\mathbf{k}} \frac{(\varepsilon_{\mathbf{k}} - \mu)n_{\mathbf{k}}}{kT}\right] \\ &\approx \sum_N \frac{V^N}{h^{3N}N!} \exp\left(\frac{N\mu}{kT}\right) \int \exp\left(-\sum_{i=1}^N \frac{\mathbf{p}_i^2}{2mkT}\right) d^{3N}p \\ &\approx \sum_N \left[ \frac{V}{h^3N} \exp\left(\frac{\mu}{kT}\right) \int \exp\left(-\frac{\mathbf{p}^2}{2mkT}\right) d^3p \right]^N. \end{aligned} \quad (4.21)$$

The factorials  $N! \approx N^N$  in the denominator are here *required* (as already known to Planck in 1900) in order to compensate for the sum over all permutations of the  $N$  momenta  $\mathbf{p}_i$  in this  $N$ -fold integral, since they all represent the *same* oscillator quantum states for the various wave modes. The latter are described by wave numbers  $\mathbf{k}$  which formally correspond to momenta  $\mathbf{p}$ . The density matrix, and therefore the partition function, now factorize in terms of wave modes  $\mathbf{k}$  rather than in terms of particle numbers, while the factorials do not have to be introduced ad hoc (as done by Satyendra Nath Bose in order to justify his photon concept).

**General Literature:** Jancel 1963.

## 4.2 Ensembles Versus Entanglement

Quantum wholeness is analyzable.

In the previous section, we derived the von Neumann equation from the Liouville equation by using the formal quantization rules. The dynamics of the density matrix, obtained in this way, is unitary. Therefore, it conserves  $S[\rho]$ , while the Pauli (or Van Hove) equation, albeit apparently derived from the von Neumann equation *as an approximation*, may seem to be superior, as it is able to describe quantum indeterminism and an increase in ensemble entropy, in particular in quantum measurements.

The Liouville equation itself was obtained in Sect. 3.1.2 by applying Hamilton's (that is, Newton's) equations to ensembles that represent incomplete knowledge about classical states. Since quantization of the Hamiltonian dynamics of mechanical systems leads to the Schrödinger equation, one may as well first quantize and then consider ensembles of its solutions  $\psi_\alpha(t)$  with corresponding probabilities  $p_\alpha$ , now describing incomplete knowledge *about the wave function* (see Fig. 4.1). This procedure may offer deeper insight into the meaning of the density matrix than its formal foundation of Sect. 4.1.1.

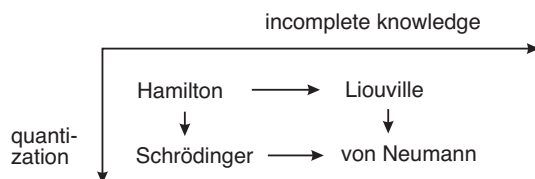
According to this ensemble interpretation, probabilities  $p_\alpha$  rather than the density matrix  $\rho(q, q')$  correspond *conceptually* to the probability distribution  $\rho_I(p, q)$ . The meaning of the density matrix can only be appreciated when considering ensemble expectation values of observables  $A$ , that is, *mean values of expectation values* with respect to different wave functions  $\psi_\alpha$ :

$$\langle A \rangle := \sum_{\alpha} p_{\alpha} \langle \psi_{\alpha} | A | \psi_{\alpha} \rangle = \text{Trace}\{A\rho\} = \sum_n a_n \langle \phi_n | \rho | \phi_n \rangle, \quad (4.22)$$

with

$$A := \sum_n |\phi_n\rangle a_n \langle \phi_n| \quad \text{and} \quad \rho := \sum_{\alpha} |\psi_{\alpha}\rangle p_{\alpha} \langle \psi_{\alpha}|.$$

The symbol  $\langle A \rangle$  denotes here a twofold mean: with respect to the ensemble of quantum states  $\psi_\alpha$  with their probabilities  $p_\alpha$ , *and* with respect to



**Fig. 4.1.** Two routes from classical mechanics to the von Neumann equation

the quantum mechanical indeterminism of measurement results  $a_n$  with their probabilities  $|\langle\phi_n|\psi_\alpha\rangle|^2$ , valid for *given* quantum states  $\psi_\alpha$ . In this way, the concept of a density matrix depends on the probability interpretation of the wave function – though *not* yet on any specific form in terms of ensembles<sup>2</sup> (see Sect. 4.6).

An ensemble interpretation of the density matrix according to  $\rho = \sum_\alpha |\psi_\alpha\rangle p_\alpha \langle\psi_\alpha|$ , used in (4.22), does not require the members  $\psi_\alpha$  of the ensemble of wave functions to be mutually orthogonal; they may even form an overcomplete set. The ensemble can therefore not be recovered from the density matrix. Von Neumann’s entropy (4.4) describes an ensemble entropy of the form  $S[\rho] = -k \sum p_\alpha \ln p_\alpha$  only for the specific ensemble consisting of the orthonormal eigenstates of  $\rho$ .

Just as for classical statistical mechanics, the conservation of entropy reflects dynamical determinism (now for wave functions) – provided the Hilbert state norm is conserved, too. This requires not only determinism, but also the *unitarity* of the Schrödinger equation (not just that of the von Neumann equation). The reason is that the formal density matrix cannot distinguish between the norm and probability  $p_\alpha$  of a wave function.

It should also be emphasized here that this formalism applies as well to wave functionals characterizing quantum field theory (that is, wave functions for a continuum of variables). ‘Backward running’ world lines in Feynman graphs are mere symbols for certain terms which appear in a relativistic perturbation expansion that is used for calculating the unitary propagation of wave functionals (general superpositions) with respect to an arbitrary but given time coordinate. These terms represent integrals over field modes (usually plane waves) – not over particle variables. Feynman’s approach has turned out to be useful even beyond *S*-matrix theory, which is restricted to describing interactions between asymptotically free objects.

The mapping of *general ensembles* of wave functions onto those which diagonalize the density matrix is an information-reducing idempotent operation

<sup>2</sup> If the elements of the probability interpretation are themselves wave functions (as in Born’s original formulation, mentioned in footnote 1, or as in collapse theories), the ensemble consisting of all possible outcomes of all conceivable measurements would be quite different from the initial ensemble (which may consist of one pure state, for example). Nonetheless, probabilities for all these outcomes are implicitly *postulated* by the phenomenological rules used in (4.22).

on these ensembles, similar to a Zwanzig projection. Nonetheless, one may rederive the von Neumann equation (4.1) from the ensemble interpretation under the further assumption that all wave functions defining the ensemble satisfy the same Schrödinger equation  $i\partial\psi_\alpha/\partial t = H\psi_\alpha$ . However, presuming the exact Hamiltonian to be ‘given’ is hardly consistent when regarding *states* as incompletely known. Even in classical physics, the precise Hamiltonian would depend on the (even less known) microscopic state of the environment (see Borel’s argument in Sect. 3.1.2).

Instead of representing an ensemble of wave functions, the density matrix may also describe the local (or ‘reduced’) perspective of *entangled* quantum systems, which are generically of the form

$$\psi(x, y) = \sum_{m,n} d_{mn} \phi_m(x) \Phi_n(y) . \quad (4.23)$$

For spatially separate subsystems, this entanglement defines *quantum nonlocality*. For example, it is responsible for the violation of Bell’s inequality (Bell 1964) or its stronger variants (Greenberger, Horne, Shimony and Zeilinger 1990), and it explains so-called *quantum teleportation* in a way which demonstrates that nothing has to be teleported: it must rather be *prepared* in advance as a component of an entangled state (see Zeh 2005c or Timpson 2005).

All measurements performed on a subsystem – corresponding to the states  $\phi(x)$ , say – of an entangled system can be characterized by the expectation values for all its subsystem observables  $A_\phi$ :

$$\langle A_\phi \rangle := \text{Trace}\{A_\phi \rho_{\text{total}}\} = \text{Trace}_\phi\{A_\phi \rho_\phi\} . \quad (4.24)$$

Here, the ‘reduced density matrix’  $\rho_\phi$  is defined as a partial trace,

$$\rho_\phi := \text{Trace}_\Phi\{\rho_{\text{total}}\} . \quad (4.25)$$

The total density matrix  $\rho_{\text{total}}$  may well be a pure state,  $\rho_{\text{total}} := |\psi\rangle\langle\psi|$ . The *new* density matrix  $\rho_\phi$  would then be explicitly given in terms of the expansion coefficients  $d_{mn}$  of the total state (4.23) as

$$(\rho_\phi)_{mm'} := \langle \phi_m | \rho_\phi | \phi_{m'} \rangle = \sum_n d_{mn} d_{m'n}^* , \quad (4.26)$$

rather than in terms of probabilities  $p_\alpha$ , which would instead lead to (4.8).

Both types of density matrices are Hermitean and positive by construction. They can therefore be diagonalized in the form  $\rho_\phi = \sum_n |\tilde{\phi}_n\rangle p_n \langle\tilde{\phi}_n|$ , with non-negative eigenvalues  $p_\alpha$ , in their eigenbasis  $\{\tilde{\phi}_n\}$ . This diagonal form defines a formal (or apparent) ensemble of orthonormal states. Although the LHS of (4.26) is thus identical with a density matrix describing an ensemble of (orthogonal or other) states, it is evident from the RHS that it does *not* represent one. Therefore, the ‘apparent ensemble’ or ‘improper mixture’ (d’Espagnat 1966) must not be used in an attempt to *explain* the probability

interpretation (4.24) on which it is based. The density matrix formalism is blind to the measurement problem (see below and Sect. 4.6).

For an entangled state such as (4.23), the eigenbases of the subsystem density matrices define the *Schmidt canonical form*,

$$\psi(x, y) = \sum_k \sqrt{p_k} \tilde{\phi}_k(x) \tilde{\Phi}_k(y) . \quad (4.27)$$

In contrast to the general representation (4.23) this is a single sum (Schmidt 1907, Schrödinger 1935). Phase factors for the coefficients  $\sqrt{p_k}$  have here been absorbed into the phase-ambiguity in the definition of the orthonormal states  $\tilde{\phi}_k$  or  $\tilde{\Phi}_k$ . For given subsystems, this representation (and hence its time dependence – see Kübler and Zeh 1973) is determined by the state  $\psi(t)$  of the total system – except for accidental degeneracy of the  $p_k$ 's.

The *neglect* of all correlations between two subsystems describes a specific loss of information, and so defines a new (nonlinear) Zwanzig projection,

$$\hat{P}_{\text{sep}} \rho := \rho_\phi \otimes \rho_\Phi . \quad (4.28)$$

A stronger Zwanzig projection of locality,  $\hat{P}_{\text{local}} \rho = \prod_k \rho_{\Delta V_k}$ , where the volume elements  $\Delta V_k$  form a complete set of local subsystems, would lead to a density matrix that factorizes, as in (3.38), in terms of these volume elements. It is again required in order to obtain the approximate concept of an *entropy density*  $s(\mathbf{r})$ . In contrast to this local picture, indistinguishable particles *cannot* be used to define subsystems that might give rise to a ‘substantial picture’. Therefore, the formal correlations between particles which describe symmetrization or antisymmetrization of the wave function does *not* represent any entanglement. These pseudo-correlations are merely an artifact from the use of classical particle concepts – see (4.21).

As a consequence of the nonlocality of quantum states, and in fundamental contrast to classical physics, the entropies  $S[\hat{P}_{\text{sep}} \rho]$  or  $S[\hat{P}_{\text{local}} \rho]$  of a completely defined (pure) quantum state are nontrivial: generically they do not vanish, since states of subsystems are *not defined* (rather than merely being unknown). *Apparent* ensembles, which are defined for them, may even be regarded as the *representative ensembles* used in statistical thermodynamics (see Chap. 3). However, one may now wonder (1) why microscopic systems are often found in pure states (such as eigenstates of their Hamiltonians  $H_0$ ), and (2) why the macroscopic world is successfully described by means of *given* classical concepts rather than in terms of their superpositions.

A local concept of relevance that, in contrast to  $\hat{P}_{\text{sep}}$ , preserves all ‘statistical’ correlations (those based on incomplete information), while removing all quantum correlations (entanglement), may be defined by using the Schmidt canonical representation in the form

$$\hat{P}_{\text{classical}}(|\psi\rangle\langle\psi|) := \sum_k p_k |\tilde{\phi}_k\rangle\langle\tilde{\phi}_k| \otimes |\tilde{\Phi}_k\rangle\langle\tilde{\Phi}_k| . \quad (4.29)$$

Quantum correlations would here require a double sum over  $k$  and  $k'$  (in a non-Schmidt basis a sum over two *pairs* of indices). The RHS can be regarded as describing incomplete information about a presumed product state.

Quantum entanglement in bipartite systems has also been studied for ‘mixed states’ of the total system (Werner 1989, Peres 1996). Its consequences may be suppressed in such a mixture either by the presumed *averaging* over the ensemble of pure states that defines this mixed state, or by *tracing out the entangled environment* that gave rise to a reduced state for the total system. However, it would be quite inappropriate to define physical states as ‘not’ or ‘less entangled’ (as implicitly done by means of effective *measures of entanglement*) just because this entanglement cannot be confirmed by measurements in this situation.

A random *pure* state (4.23) would not lead to the ‘statistical’ result  $\rho_{\text{irrel}} \approx 0$  – as used in (3.46) – for the relevance concept (4.28), since only the improbable factorizing states do not contain any correlations. (The reader may wish to skip the rest of this somewhat technical paragraph.) For example, the linear entropy according to (4.7), given by  $S_{\text{lin}} = 1 - \sum_{kk'} |\rho_{kk'}|^2$  if normalized to vanish for a pure state, assumes its maximum in a Hilbert space of finite dimension  $D$ ,  $S_{\text{lin}}^{\text{max}} = 1 - 1/D$ , for the maximal mixture  $\rho_{kk'} = \delta_{kk'}/D$ . For *pure states* in the tensor product of two Hilbert spaces with dimensions  $M$  and  $N$  (hence with  $D = MN$ ), one obtains for the *mean* linear subsystem entropy, defined below (4.7), in either subsystem (Lubkin 1978):

$$\bar{S}_{\text{lin}}^{\text{sub}} = 1 - \frac{M + N}{D + 1} < 1 - \frac{1}{M} \quad \text{and} \quad < 1 - \frac{1}{N} . \quad (4.30)$$

This entanglement entropy vanishes only for  $M = 1$  or  $N = 1$ . Since the *linear information*  $I_{\text{lin}} := 1 - S_{\text{lin}}$  factorizes for products of density matrices (rather than being additive as would its logarithmic measure), its value for the total system is given by  $I_{\text{lin}}[\hat{P}_{\text{sep}}\rho_{\text{pure}}] = (I_{\text{lin}}^{\text{sub}})^2$ , as the entropies of the subsystems must be equal for a pure total state – see (4.27). For  $D \gg 1$  one has, according to (4.30),

$$\bar{I}_{\text{lin}}[\hat{P}_{\text{sep}}\rho_{\text{pure}}] \approx \frac{(M + N)^2}{MN} I_{\text{lin}}^{\text{min}} ,$$

so that  $\bar{I}_{\text{lin}}[\hat{P}_{\text{sep}}\rho_{\text{pure}}] \approx (M/N)I_{\text{lin}}^{\text{min}}$  for  $M \gg N$ , with  $I_{\text{lin}}^{\text{min}} = 1 - S_{\text{lin}}^{\text{max}} = 1/MN$  (characterizing maximal mixtures). While  $\rho_{\text{irrel}} = (1 - \hat{P}_{\text{sep}})\rho$  vanishes in a random mixture  $\sum_{\alpha} |\psi_{\alpha}\rangle p_{\alpha} \langle \psi_{\alpha}|$ , the mean information for pure total states,  $\sum_{\alpha} p_{\alpha} I_{\text{lin}}[(1 - \hat{P}_{\text{sep}})|\psi_{\alpha}\rangle \langle \psi_{\alpha}|]$ , does not. Since the irrelevant information is measured by  $I[\rho] - I[\hat{P}_{\text{sep}}\rho]$ , rather than by  $I[(1 - \hat{P}_{\text{sep}})\rho]$ , the random pure state (with  $S_{\text{lin}}[\rho] = S_{\text{lin}}^{\text{min}} = 0$ ) must possess large *local* entropy  $S_{\text{lin}}[\hat{P}\rho]$ , and therefore has to carry its information predominantly in  $\rho_{\text{irrel}}$ .

Similar relations hold for the logarithmic entropy (Page 1993, Foong and Kanno 1994). It is impossible to reach  $S^{\text{max}}$  for  $M \neq N$  for a pure total state. Because of  $I_{\phi} = I_{\bar{\phi}}$ , the *local information* about the larger system cannot be

entirely transformed into correlations. However, every (small) subsystem of the completely described quantum universe would essentially possess maximum entropy for a random global state.

If the total wave function  $\psi$  evolves according to a Schrödinger equation, the reduced density matrix does *not* in general obey a von Neumann equation. While its exact dynamics can still be explicitly formulated (Kübler and Zeh 1973, Pearle 1979), it remains entangled with the rest of the total system. Indeed, the reduced density matrix  $\rho_\phi$ , multiplied by the unit operator in  $\Phi$ -subspace, defines a further (linear) Zwanzig projection,

$$\hat{P}_{\text{sub}}\rho_{\text{total}} := \rho_\phi \otimes 1_\Phi . \quad (4.31)$$

Phenomenological irreversible master equations for  $\rho_\phi$  (instead of a von Neumann equation) are known as ‘open systems’ quantum dynamics (see Sect. 4.4). They are often derived by presuming an uncorrelated environmental heat bath (Favre and Martin 1968, Davies 1976). This assumption of lacking initial quantum correlations is similar to Boltzmann’s chaos assumption. Master equations should therefore *explain* the canonical ensembles describing heat baths rather than presuming their existence. Open systems have also been described by means of path integrals (Feynman and Vernon 1963).

As already mentioned, both expectation values, (4.22) and (4.24), which were used to derive the concept of a density matrix, depend on the probability interpretation. Von Neumann (1932) introduced a model interaction in an attempt to describe ideal measurements (or ‘measurements of the first kind’) *dynamically*. It is defined by the unitary transformation  $\phi_n\Phi_0 \xrightarrow{t} \phi_n\Phi_n$ , where  $\phi_n$  is an eigenstate of an observable  $A = \sum_n |\phi_n\rangle a_n \langle \phi_n|$ , while  $\Phi_0$  and  $\Phi_n$  are the initial state of the apparatus and its final ‘pointer positions’, respectively. This dynamics describes a fork of causality in the classical *configuration space* on which the wave function is defined (see footnote 1 of Chap. 2). The observable  $A$  is thus *defined* by this interaction, whereby its eigenvalues  $a_n$ , characterize the ‘pointer scale’. If the microscopic system is initially in one of the eigenstates of  $A$ , it does not change during such an ideal measurement, while the apparatus evolves into the corresponding pointer state  $\Phi_n$ . In quantum optics, such measurements are also called ‘quantum non-demolition measurements’, since photons are usually absorbed when being measured. In the case  $n = 1, 2$  and  $\Phi_0 = \Phi_1$  they are identical to ‘controlled-not gates’ – much discussed in the theory of quantum computing.

However, for a general initial state,  $\sum_n c_n \phi_n$ , one now obtains for the same interaction, and for the same initial state  $\Phi_0$  of the apparatus,

$$\left( \sum_n c_n \phi_n \right) \Phi_0 \xrightarrow{t} \sum_n c_n \phi_n \Phi_n =: \psi_{\text{final}} . \quad (4.32)$$

If the pointer states  $\Phi_n$  are mutually orthogonal, too, both sides of (4.32) are Schmidt-canonical. The RHS is now an *entangled* state, while an *ensemble* of

different measurement results (that is, of states  $\phi_n \Phi_n$  with probabilities  $|c_n|^2$ ), would require the fork of causality to be replaced by a fork of indeterminism. (The formal ‘plus’ characterizing the superposition would have to become an ‘or’.) This discrepancy represents the *quantum measurement problem*, that would be obscured in a phenomenological description by means of reduced density matrices for the subsystems only. The density matrices resulting from these two types of forks are identical, since there is no way of distinguishing these different situations operationally by a *local* measurement. As emphasized before, this argument does not *explain* the fork of indeterminism that lies at the heart of the probability interpretation.

This measurement problem prevails regardless of the complexity of the measurement device (that might give rise to thermodynamically irreversible behavior), and regardless of any perturbations caused by, and in the environment, since the states  $\Phi$  may be assumed to describe this complexity completely, and even to include the whole ‘rest of the Universe’. The popular argument that quantum mechanical indeterminism might, in analogy to the classical situation, be caused by thermal fluctuations that occur during a measurement process (see Sect. 3.3 or Peierls 1985, for example) is incompatible with universal unitarity. It would instead require the existence of an initial ensemble of microscopic states which in principle had to *determine* the outcome. However, the ensemble entropy of the RHS of (4.32) does not represent an ensemble that would allow the measurement to be interpreted as in Fig. 3.5.

If both systems in (4.32) are microscopic, the dynamics representing the fork of causality can even be reversed in practice (the measurement could be ‘undone’) in order to demonstrate that all components still exist. This reversal leads to observable consequences that may depend on all existing components, including their relative phases. This excludes the interpretation of (4.32) as a dynamical fork of indeterminism (a fork between mere possibilities), even though von Neumann’s fork of causality is defined in terms of wave packets on a *classical configuration space*. Therefore, the transition from quantum to classical (Sect. 4.3) can be understood only if it explains why the fundamental arena for wave functions often *appears* as a space of classical configurations.

The interaction (4.32) is an example for the generic case of quantum mechanical subsystems which are not individually obeying unitary dynamics. Similar arguments would apply to the *ensemble* dynamics of systems with classical correlations (that is, if  $\rho = \hat{P}_{\text{classical}}\rho$ ). In this case, the effective subsystem Hamiltonian  $H_\phi$ , say, would depend on the state  $\Phi_k$  of the complementary system by means of a partial expectation value,

$$H_\phi^{(k)}(t) := \langle \Phi_k(t) | H | \Phi_k(t) \rangle_{\text{Op}}, \quad (4.33)$$

where  $H$  was defined to act on the tensor product of  $\phi$  and  $\Phi$ . There can be no Hamiltonian  $H_\phi$  valid for the whole ensemble any more. This is equivalent to the induced Hamiltonians of interacting classical mechanical systems, which are given by



$$H_\phi(p_1 \dots q_n, t) := H(p_1 \dots q_n; p_{n+1}(t) \dots q_N(t)) . \quad (4.34)$$

Here, particle numbers  $1, \dots, n$  are meant to characterize the considered subsystem ‘ $\phi$ ’, while all others  $(n+1, \dots, N)$  represent the ‘environment’. Each element of the ensemble would then satisfy another Hamiltonian or Schrödinger equation – in contrast to the assumptions leading to the Liouville or von Neumann equation. Nonetheless, for each element of an ensemble representing incomplete knowledge, the subsystem evolution would be *determined* in this classical case. Neglecting the statistical correlations dynamically by using  $\hat{P}_{\text{sep}}$  in a master equation would amount to applying the whole resulting ensemble of sub-Hamiltonians (in the forward direction of time) to each individual element of the ensembles of states of the subsystems. However, only under the unstable assumption  $\rho = \hat{P}_{\text{classical}}\rho$  (that is, without any entanglement) would the quantum mechanical situation simply be equivalent to the classical one of (4.34), or as in Sect. 3.1.2.

It should be kept in mind, therefore, that the local concepts of relevance,  $\hat{P}_{\text{sep}}$ ,  $\hat{P}_{\text{local}}$  and  $\hat{P}_{\text{classical}}$ , appear ‘natural’ only to our classical prejudice. In the unusual situation of controllable entanglement (as in EPR/Bell type experiments), quantum correlations may become relevant by means of the re-localization of superpositions even for local observers. *Dynamical* locality, as described by means of point interactions in field theory, merely warrants the dynamical consistency of these concepts of relevance, or gives rise to the approximate validity of autonomous master equations for  $\hat{P}_{\text{local}}\rho$ .

**General Literature:** d’Espagnat 1976, 1983.

## 4.3 Decoherence

Novel ideas in science are at first completely neglected, then fiercely attacked, and finally regarded as well known.

Konrad Lorenz

In Sect. 3.1 we saw how molecular collisions produce statistical correlations, which describe ‘irrelevant’ information. Although other relevance concepts may also be appropriate for describing irreversible phenomena, the formation of statistical correlations seems to be the most important one in classical description. In a gas, these correlations arise by means of a momentum transfer between molecules, eventually leading to a Maxwell distribution – the distribution of highest entropy for given mean energy if correlations are neglected.

If one specific ‘molecule’ happens to possess macroscopic mass (such as a bullet flying through the gas), its recoil may approximately be neglected in collisions with molecules – except for the resulting friction, whose importance depends on the density of the gas. The bullet may then remain in a non-equilibrium state of almost free motion for some time. On the other hand,

collisions drastically affect the microscopic molecules. Although their states after scattering must strongly depend on the bullet's position at collision time, this dependence cannot be regarded as representing *information* about it if the molecular motions are already chaotic. In contrast, light scattered off the bullet *does* carry information, as we may easily confirm by using our eyes. The reason is that light interacts weakly or coherently with matter, and remains in a state far from equilibrium if absorption can be neglected (see Chap. 2).

The effect of an individual molecule or photon on a macroscopic object may thus be neglected in classical description, but this conclusion has to be radically revised in quantum mechanics. The quantum interaction can be described as an ideal (though uncontrollable) 'measurement' of the bullet's position and shape by the molecule in the sense of von Neumann. If the bullet were initially in a superposition of different positions, as one would have to expect for an object in a generic quantum state, this would lead to an entangled state as in (4.32). In this case, the initial superposition becomes *dislocalized* (it is at *no* place any more). This is called 'decoherence' if the dislocalization is irreversible in practice.<sup>3</sup> (Reversible dislocalization of a superposition – such as in a Stern-Gerlach device – may be regarded as 'virtual decoherence'.) It turns out that real decoherence is not only unavoidable for all macroscopic objects, but even the most abundant and most important irreversible process in Nature (Zeh 1970, 1971, 1973, Leggett 1980, Zurek 1981, 1982a, Joos and Zeh 1985).

In general, decoherence is not pure, but accompanied by a distortion of the system under consideration (recoil). For an environmental heat bath this would be required by the *fluctuation-dissipation theorem*, which leads to 'quantum Brownian motion' – a combination of decoherence, dissipation and fluctuation. However, the quantitative relation between these phenomena depends on actual parameters, such as temperature and mass ratios. Since fluctuation and dissipation may so become arbitrarily small, 'ideal' measurements by the environment are appropriate for studying 'pure decoherence' as a genuine quantum phenomenon. Chaotic molecules then contribute to decoherence just as ordered light. Evidently it is the physical effect on the environment that is essential – not any transfer of information. 'Quantum information' is here no more than a misleading renaming of entanglement.

Decoherence is also important for strongly interacting *microscopic* systems, such as individual molecules in a gas, although it is then far from being pure (recoil is essential). Instead of quasi-classical behavior, one now obtains quasi-stochastic dynamics – as successfully *used* in the *Stoßzahlansatz*. Interacting microsystems constituting solids can often be approximated by coupled harmonic oscillators (Caldeira and Leggett 1985). While solutions are then analytically available, they are also known to possess certain pathological properties. In particular, they are non-ergodic.

---

<sup>3</sup> The term decoherence has often been misused in the literature. See Sect. 3.4.3 of Joos et al. (2003) on 'True, False and Fake Decoherence'.

Applying the terminology used in the previous section, decoherence may be understood as the justification of a *specific*  $\hat{P}_{\text{semidiag}}$  for a given subsystem by presuming the relevance of locality, as described by the corresponding  $\hat{P}_{\text{sub}}$  – see (4.31). If this  $\hat{P}_{\text{semidiag}}$  turns out to be dynamically valid under all normal circumstances, its eigenspaces characterize ‘quasi-classical’ properties or superselection rules (Zeh 1970, Zurek 1982a). Classical concepts *emerge* approximately in the form of apparent ensembles of narrow wave packets through unavoidable and practically irreversible interaction with the environment. They do not have to be presumed as an independent fundamental ingredient, required for an interpretation of the formalism (as done in the Copenhagen interpretation). From a pragmatic point of view, which does not distinguish between proper and improper mixtures, this would already be sufficient to solve the measurement problem. In a consistent description of reality in terms of wave functions, however, one must assume either a genuine collapse to be *triggered* by decoherence in some way, or appropriately redefine conscious observers within an Everett interpretation (see Sect. 4.6).

The interaction (4.32) was introduced by von Neumann to describe the controllable measurement of a microscopic system  $\phi$  by an appropriate device (with ‘pointer’ states  $\Phi_n$ ). Its fact-like time asymmetry, leading from factorizing to entangled states, could be reversed with sufficient effort if both subsystems were microscopic (‘recoherence’ or ‘erasure of measurement results’). For genuine quantum measurements, the pointer states  $\Phi_n$  must be macroscopic. They are then ‘measured’ in turn by their uncontrollable environment, and thus become irreversibly quasi-classical. This explains why measurements which lead to macroscopic pointer positions cannot be undone.

It is this universality and unavoidability of entanglement with the natural environment that seems to have been overlooked for the first 50 years of quantum theory. All attempts to describe macroscopic objects quantum mechanically as being isolated, and therefore by means of a Schrödinger equation, were thus doomed to failure – even when including environment-induced dynamical terms that might describe a distortion. Decoherence is different, and extremely efficient, since it does not require an environment that *disturbs* the system. The distortion *of the environment by the system* affects the density matrix of the system, too, because of quantum nonlocality, but on a much shorter time scale than thermal relaxation or dissipation (Joos and Zeh 1985, Zurek 1986).

Some examples of decoherence will now be discussed in detail.

**General Literature:** Joos et al. 2003, Zeh 2005c, Schloßhauer 2006.

#### 4.3.1 Trajectories

Imagine a two-slit interference experiment with bullets or small dust particles, described by quantum mechanics. Then not only their passage through the slits, but their whole path would be ‘measured’ by scattered molecules or

photons. No interference fringes could ever be observed for such macroscopic objects – even if the resolution of the registration device were fine enough.

In this respect, macroscopic objects are similar to alpha ‘particles’ in a Wilson chamber, which interact strongly with gas molecules by means of their electric charge. For all these objects, their unavoidably arising entanglement with their environment leads to a reduced density matrix that can be represented by an ever-increasing ensemble of narrow wave packets following slightly stochastic trajectories (see also Mott 1929). This result is not restricted to the quantum description of motion in space: propagating wave packets in the configuration space of macroscopic variables may similarly explain their apparent ‘histories’. For spatial motion the argument also demonstrates that the concept of an  $S$ -matrix does not apply to macroscopic objects, since it presumes asymptotically free states.

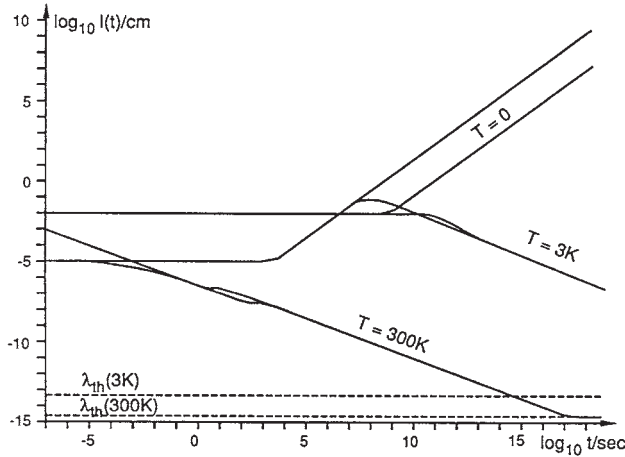
Several very instructive interference experiments have recently been performed with *mesoscopic molecules* that are in the transition region between isolated quantum and classical behavior. Various mechanisms of decoherence, including the emission of thermal radiation from internal molecular degrees of freedom, can be studied for them in detail (Arndt et al. 1999, Hornberger, Hackermüller and Arndt 2005).

For a continuous variable, such as position, decoherence competes with the dispersion of the wave packet that is reversibly described by the Schrödinger equation. Even the scattering rate of photons, atoms, or molecules off small dust particles in intergalactic space suffices to destroy any coherence that would define spreading wave packets for their centers of mass (see Fig. 4.2). If the wavelengths of the abundant scatterers are larger than the width of the wave packet, an otherwise free ‘particle’ is dynamically described by the master equation (Joos and Zeh 1985)

$$i\frac{\partial\rho(x,x',t)}{\partial t} = \frac{1}{2m}\left(\frac{\partial^2}{\partial x'^2} - \frac{\partial^2}{\partial x^2}\right)\rho - i\lambda(x-x')^2\rho. \quad (4.35)$$

It can be *derived* from a universal Schrödinger equation by assuming the dynamical irrelevance of all correlations with the environment after they have formed, and by neglecting recoil (see also Chap. 3 and Appendix 1 of Joos et al. 2003). The coefficient  $\lambda$  is here determined by the rate of scattering and its efficiency in orthogonalizing states of the environment. In the small wavelength limit, a *single* collision is usually sufficient to destroy any coherence beyond the wavelength. The decoherence rate is then simply given by the scattering rate (that is, the product of the flux of environmental particles and the total cross-section). Even the interpretation of the wave mechanical scattering process as consisting of individual collision *events* can be explained by further decoherence of superpositions of different ‘collision times’ in a process that is actually smooth (see Sect. 4.3.6).

So one does not have to *postulate* a fundamental semigroup in order to describe open quantum systems (Sect. 4.4). If the environment forms a heat bath, (4.35) describes the infinite-mass limit of *quantum Brownian motion*



**Fig. 4.2.** Time dependence of the coherence length  $l(t)$  for the center of mass of a small dust grain of  $10^{-14}$  g with radius  $10^{-5}$  cm under continuous measurement by thermal radiation. The six curves represent two initially pure Gaussian wave packets, differing by their initial widths  $l(0)$ , and three different temperatures  $T$  of the radiation.  $T = 0$  describes the free dispersion of the wave packet according to the Schrödinger equation, which holds otherwise as an approximation for a limited time only. Scattering of atoms and molecules is in general far more efficient than that of thermal photons – even in intergalactic space. Brownian motion becomes relevant only when the coherence length approaches the de Broglie wavelength  $\lambda_{th}$ . From Joos and Zeh (1985)

(see Caldeira and Leggett 1983, Zurek 1991, Hu, Paz and Zhang 1992, Omnès 1997). This demonstrates that, even for entirely negligible recoil (which would be responsible for noise and friction), there remains an important effect that is based on quantum nonlocality.

Apparent classical properties thus *emerge* from the wave function, and are maintained, by a process that cannot be reversed. In particular, particle aspects (such as tracks in a bubble chamber) arise in the form of macroscopic phenomena (bubbles) which are observed at certain positions in space because of their decoherence. Similarly, the disappearance of interference between *partial waves* in a *Welcher Weg* measurement (Scully, Englert and Walther 1991) does not require any wave–particle ‘complementarity’. Furthermore, no superluminal tunnelling (see Chiao 1998) may occur according to a consistent quantum description, since *all parts* of a wave packet propagate (sub-)luminally, while its group velocity does not represent the propagation of any physical objects.

Master equations for open systems can also be derived by means of the *decoherence functional* (Feynman and Vernon 1963, Mensky 1979). Feynman’s path integral is thereby used as a *tool* for calculating the propagation of a global density matrix, while the environment is again continuously traced out

when getting entangled with the considered system. The intuitive picture of an *ensemble of paths* (representing different *possible trajectories*) is justified only if this superposition of paths decoheres into narrow wave packets. A ‘restriction’ of the path integral by the presence of absorbers (Mensky 2000) would be equivalent to a corresponding reduction of the (total) wave function.

All quasi-classical phenomena, including those representing apparently reversible (friction-free) mechanics, rely conceptually on irreversible decoherence. This requires the continuous production of objective physical entropy (increasing entanglement), which may be macroscopically negligible, but is large in terms of bits. If the quasi-classical trajectories are chaotic, this entropy production may be controlled by the classical Lyapunov exponent (Zurek and Paz 1994, Monteoliva and Paz 2000), even though the entanglement entropy does not require any (initial) uncertainties that would grow *in the direction of calculation*, as assumed in the classical theory of chaos (see Sect. 3.1.2).

**General Literature:** Joos’s Sect. 3.2 of Joos et al. (2003), Hornberger, Hackermüller, and Arndt (2005).

#### 4.3.2 Molecular Configurations as Robust States

Chirality of molecules, such as right- or left-handed sugar, represents a discrete elementary variable controlled by decoherence. Although a chiral state is described by a certain wave function, it is not an energy eigenstate, which would have to be a parity eigenstate, that is, a symmetric or antisymmetric superposition of both chiralities (see Zeh 1970, Primas 1983, Woolley 1986). The reason is that it is chirality (not parity) that is continuously ‘measured’, for example by scattered air molecules – in analogy to position rather than momentum being measured for a macroscopic ‘mass point’. For sugar molecules under normal conditions, the decoherence time scale is of the order of  $10^{-9}$  s (Joos and Zeh 1985), while the tunneling time between different chirality states is extremely long.

As a consequence, each individual molecule in a bag of sugar retains its chirality, while a parity state – if it had come into existence in a mysterious or expensive way – would almost immediately ‘collapse’ into an apparent ensemble of two chirality states (with equal probabilities). Parity would thus *not be conserved* for sugar molecules, while chirality is always confirmed when measured twice (although it is not a constant of the motion).

This *robustness* against decoherence seems to characterize properties that we usually regard as ‘elements of classical reality’, such as spots on the photographic plate or other ‘pointer states’ of a measurement device. Discrete states may even be protected against otherwise possible transitions (tunneling) by the *quantum Zeno effect*. For continuous variables, the concept of robustness is compatible with a (regular) time dependence according to a master equation, as described in the previous section for the quasi-classical center of mass motion of macroscopic objects. Since entropy production by interaction with

the environment is lowest for a density matrix that is already diagonal in terms of robust states, this property has been called a ‘predictability sieve’, and proposed as a definition of classical states (Zurek, Habib and Paz 1993).

Dynamical robustness is also essential for the physical concept of memory or information storage, such as in DNA, brains or computers. Even ‘states-of-being-conscious’ (see Chap. 1) seem to be quasi-classical in this sense (Tegmark 2000) – at least inasmuch as they are able to communicate. In contrast to such robust properties, which can be assumed to exist regardless of their actual measurement, *potentially* measurable quantities have been called ‘counterfactuals’. Their superpositions, which would themselves describe individual physical states, must not be assumed to describe ensembles of definite (really existing though unknown) properties. Such different concepts of reality (operational or phenomenological versus hypothetical though consistent and economically chosen) can thus be analyzed and understood in terms of decoherence, which is thereby assumed to represent a physical process in a consistent (nonlocal) quantum reality, while elements of phenomenological (classical) reality ‘emerge’ (or become ‘factual’) only under certain environmental conditions. If these conditions may change, such as for microscopic systems under different measurements, the emerging concepts naturally vary between ‘complementary’ modes of description.

Chemists know furthermore that atomic nuclei or strongly bound ions as constituents of *large* molecules have to be described classically (for example as quasi-rigid configurations, which may vibrate or rotate in a time-dependent manner), while the electrons have to be described by stationary or adiabatically comoving wave functions. This asymmetric behavior is often attributed, by means of a Born–Oppenheimer approximation, to their large mass ratio. However, this argument is insufficient, since this approximation applies as well to small molecules that are found in discrete energy eigenstates, which are completely described by stationary wave functions, giving rise to discrete rotational and vibrational energy bands rather than quasi-classical states.

The formation of time-dependent (particle-like) wave packets for the atomic nuclei in large molecules can instead be understood once again by means of decoherence (Joos and Zeh 1985). For example, the positions of nuclei are usually permanently monitored by scattering of lighter molecules that form the environment. But why only the nuclei (or ions), and why not very small molecules? The answer requires a quantitative investigation in each individual case, and the result depends on a delicate balance between internal dynamics and interaction with the environment, whereby the density of states plays a crucial role (Joos 1984). This may then lead approximately to either (a) unitary evolution (including stationary states), (b) a master equation, or (c) freezing of the motion (quantum Zeno effect). Much numerical work remains to be done for such complex systems, while simple ones may be described by an effective master equation, such as (4.35), for example.

**General Literature:** Joos’s Sect. 3.2.4 of Joos et al. 2003.



### 4.3.3 Quantum Computers

Digital computers are based on robust binary states, carrying ‘bits’ of information. Even neural networks can be described to some extent by states of cells having ‘fired’ or not, while DNA is based on four different ‘letters’, each one therefore representing two bits. Chiral molecules also represent bits, although they would not be very convenient for information handling.

Just as chiral states may be robust because of their decoherence, so are all macroscopic constituents that are used in classical computers. However, on a microscopic scale there also exist quantum bits (or ‘qubits’), which may occur in all conceivable superpositions of their two basic states. In some cases, such as photon polarizations or spinors that may form spin lattices, they may even be assumed to be isolated from the environment to a good approximation. Such isolated qubits form the essential constituents of quantum computers. Because of their greater variety of possible states (for example spin-up and spin-down in any direction of space), and the possibility of getting entangled, they offer quite novel possibilities for computing (see Shor 1994).

The problem here is that completely isolated systems, required for a unitary evolution, could hardly be manipulated or read as wished for a usable computer. On the other hand, any uncontrollable effect of the collective state of an  $n$ -qubit system on the environment would immediately destroy (that is, irreversibly dislocalize) the crucial superposition that forms the state of this system as a whole. This vulnerability of quantum computers against decoherence grows exponentially with their size, so that macroscopic quantum computers may have to be excluded by superselection rules, similarly to macroscopic superpositions in general. Superpositions containing a large number of entangled electrons that have been prepared and observed in the laboratory (Mooij et al. 1999, Friedman et al. 2000) are facilitated by ‘freezing out’ most of the degrees of freedom in a degenerate state – in stark contrast to what would be required for the complexity of a quantum computer.

In an attempt to overcome this problem, various *correction codes* have been proposed (see Bouwmeester, Ekert, and Zeilinger 2000). They are conventionally based on some concept of multiple redundancy (an internal kind of back-up), that would have to further enlarge the number of qubits. However, while redundancy may be used as a protection against distortions of the computer by the environment, decoherence is a distortion of the environment by the computer. It can only be corrected for inasmuch as the environment remains controllable – certainly not a very realistic assumption. Usable quantum computers may therefore be excluded in practice for some time to come (see also Haroche and Raimond 1996). It would be quite inconsistent, though, to study the possibility of quantum computers even in principle, while at the same time denying the *reality* of all components of a quantum superposition or wave function – as appropriately emphasized by David Deutsch (1997). Decoherence, too, is the consequence of such an assumption.



In order to give rise to a *classical* computer, each bit would have to be decohered after each calculational step. This would produce precisely the minimum amount of entropy of  $k \ln 2$  that was conjectured to be required by Landauer for other reasons (see the end of Sect. 3.3), but then refuted by Bennett in a classical deterministic setting. This entropy production would thus again (have to) be avoided in quantum computers according to the deterministic Schrödinger equation, which is valid only for isolated systems.

**General Literature:** Bouwmeester, Ekert, and Zeilinger (2000).

#### 4.3.4 Charge Superselection

Gauss' law,  $q = (1/4\pi) \int \mathbf{E} \cdot d\mathbf{S}$ , tells us that every local electric charge requires a certain flux of electric field lines through a sphere surrounding it at any distance. For a superposition of different charges, one would therefore obtain an entangled quantum state of charges and fields,

$$\begin{aligned} \sum_q c_q \psi_q^{\text{total}} &= \sum_q c_q \chi_q \Psi_q^{\text{field}} = \sum_q c_q \chi_q \Psi_q^{\text{near}} \Psi_q^{\text{far}} \\ &=: \sum_q c_q \chi_q^{\text{dressed}} \Psi_q^{\text{far}} , \end{aligned} \quad (4.36)$$

where  $\chi_q$  represents the bare charge, while  $\Psi_q^{\text{field}} = \Psi_q^{\text{near}} \Psi_q^{\text{far}}$  is the state vector of its correlated electrostatic field, symbolically written as a tensor product of a near field and a far field (see Sect. 2.3). The dressed (physical) charged particle would then be described by a density operator of the form

$$\rho_{\text{local}} = \sum_q |\chi_q^{\text{dressed}}\rangle |c_q|^2 \langle \chi_q^{\text{dressed}}| , \quad (4.37)$$

provided that the states of the far field for different charge  $q$  are mutually orthogonal (uniquely distinguishable). The charge is thus decohered by its own Coulomb field, and no charge superselection rule has to be *postulated* (see Giulini, Kiefer and Zeh 1995). The formal decoherence of the bare charge by its near field remains unobservable, since experiments can only be performed with dressed charges.

While this result explains the observed charge superselection rule, one may ask what it means locally. What if an electric charge is accompanied by a negative one at a different place? Or at what distance and on what time scale would the superposition of two different locations of a point charge (such as those of an electron during an interference experiment) be decohered by the quantum state of the corresponding dipole field. A *classical* retarded Coulomb field would contain causal information about the precise path of its source particle. However, interference between different paths of an electron has been demonstrated to exist at least over distances of the order of millimeters (Nicklaus and Hasselbach 1993). This indicates that the Coulomb field contributes

to decoherence only by its monopole component, sufficient to explain charge superselection.

This conclusion can indeed be understood in terms of quantum theory, since photons with diverging wavelength (which may be regarded as representing static fields) cannot distinguish different charge positions – even though the number of such virtual photons would diverge in a Coulomb field. Static dipole (or higher) multipole moments do not possess any far fields. Therefore, only the ‘topological’ Gauss constraint  $\partial_\mu F^{\mu 0} = 4\pi j^0$  contributes to the decoherence of the physical particle by the Coulomb field. Any time-dependence (including a retardation) must then be described in terms of transverse photons, represented by the vector potential  $\mathbf{A}$  (with  $\text{div}\mathbf{A} = 0$  in the Coulomb gauge). In this picture, only the spatial distribution of electric field lines – not their total flux – forms dynamical degrees of freedom that have to be quantized. Charge decoherence has therefore been regarded as ‘kinematical’, although it might as well be assumed to be dynamically *caused* by the retarded field of the (conserved) charge in its past – or equivalently by the advanced field resulting from its future. Note, however, that a kinematical Coulomb constraint is in conflict with the concept of a physical Hilbert space that is spanned by direct products of *local* states.

Dipoles and higher moments (which can define position *differences* for a point charge), can thus be measured by the environment either through emission (or scattering) of transverse (‘real’) photons, or by the irreversible polarization of nearby matter (Kübler and Zeh 1973, Anglin and Zurek 1996). The latter effect has now been experimentally confirmed (Sonnentag and Hasselbach 2005). In general, this decoherence is not ‘pure’, but related to energy transfer, although the recoil caused by emission of soft photons may be negligible. The (often virtual) decoherence of individual charged particles *within* solid bodies is discussed in Imry (1997).

The emission of photons would require the charge to be *accelerated*. For example, a *transient dipole* of charge  $e$  and maximum distance  $d$ , caused by spatially separating opposite charges for a time interval  $t$ , requires accelerations  $a$  of the order  $d/t^2$ . According to Larmor’s classical formula (see Sect. 2.3), the intensity of radiation is then at least  $2e^2a^2/3$ . In order to resolve the position difference, the emitted radiation has to consist of photons with energy greater than  $\hbar c/d$  (that is, wavelengths smaller than  $d$ ). The probability that information about the dipole is radiated away by at least one photon is then very small: of order  $\alpha Z^2(d/ct)^3$ , where  $\alpha$  is the fine structure constant and  $Z$  the charge number. In more realistic cases, such as interference experiments with electrons, stronger accelerations may occur, but they would in general still cause negligible decoherence.<sup>4</sup> Decoherence of the position of a charged

---

<sup>4</sup> This limitation of the information capacity of an electromagnetic field by its quantum nature must also give rise to an upper bound for the validity of Borel’s argument of Sect. 3.1.2.

particle is therefore dominated by *scattering* of photons, and by interaction with charged or polarizable matter.

The *gravitational field* of a point mass is similar to the Coulomb field of a point charge. Superpositions of different mass should therefore be decohered by the quantum state of the monopole contribution of spatial curvature, and thus give rise to a mass superselection rule. However, superpositions of different energies (hence masses) evidently exist, since they form the time-dependent states of local systems. This situation may not yet be sufficiently understood.

The Coulomb field would vanish globally if the total charge of the Universe were zero (see Giulini, Kiefer and Zeh 1995). This would eliminate the need for a Gauss constraint for the Universe. The gravitational counterpart of this global consequence is the absence of time from a closed Universe in quantized general relativity (the Hamiltonian constraint – see Sect. 6.2).

**General Literature:** Kiefer’s Sect. 4.1.1 and Giulini’s Chap. 6 of Joos et al. 2003.

#### 4.3.5 Quasi-Classical Fields and Gravity

Not only are the quantum states of charged particles decohered by their fields – quantum states of fields may in turn be decohered by the currents on which they act. In this case, ‘coherent states’, that is, Schrödinger’s time-dependent but dispersion-free Gaussian wave packets for the amplitudes of classical wave modes (eigenmodes of coupled oscillators), have been shown to be robust for similar reasons as electric charges, chiral molecules or the wave packets describing the center of mass motion of quasi-classical objects (Kübler and Zeh 1973, Kiefer 1992, Zurek, Habib and Paz 1993, Habib et al. 1996). This explains why macroscopic states of neutral boson fields appear as *classical fields*, and why superpositions of macroscopically different ‘mean fields’ or different vacua (Sect. 6.1) are never observed.

Coherent harmonic oscillator states, which form states of minimum Heisenberg uncertainty, can be defined (for each wave mode  $k$ ) as eigenstates  $|\alpha_k\rangle$  of the non-Hermitean annihilation (or energy-lowering) operators  $a_k$  with their complex eigenvalues  $\alpha_k$  (that is,  $a_k|\alpha_k\rangle = \alpha_k|\alpha_k\rangle$ ). These Gaussian wave packets are centered at a time-dependent classical field amplitude  $\alpha_k(t) = \alpha_k^0 e^{i\omega t}$ , where  $\text{Re}(\alpha_k)$  and  $\text{Im}(\alpha_k)$  represent the electric and magnetic field strengths, formally equivalent to the position and momentum of a mechanical oscillator. Since the interaction between the field and its charged sources is usually linear in the field operators  $a_k$  or  $a_k^\dagger$ , these coherent states form an (overcomplete) robust ‘pointer basis’: they create minimal entanglement with their ‘environment’ (that consists here of charged sources that happen to be present).

In contrast to these superpositions of many different photon numbers (or oscillator quantum numbers), *single-photon* states resulting from the decay of different individual atoms (or even the  $n$ -photon states resulting from the

decay of a *different number*  $n$  of atoms) are unable to interfere with one another, since they are entangled with mutually orthogonal final states of the sources. Two incoherent components of a one-photon state may then appear as ‘different’ photons (using Dirac’s language), although the photons themselves are indistinguishable. A quasi-classical collective state of the source, however, would hardly change (judged in terms of the Hilbert space inner product) when emitting a photon. It is thus able to *produce* the coherent superpositions of different photon numbers discussed above (see also Kiefer 1998).

Although the coherent states behave macroscopically, superpositions of *different* ones,  $c_1|\alpha_1\rangle + c_2|\alpha_2\rangle$  (called ‘Schrödinger cat states’), have been produced and maintained for a short time as one-mode laser fields in a cavity (Monroe et al. 1996). These mesoscopic superpositions must decohere, similarly to a Schrödinger cat, although on a time scale that is slow enough to allow this decay of coherence to be monitored as a function of time. In this way, decoherence was for the first time confirmed experimentally as a smooth process in accord with the Schrödinger equation (Davidovich et al. 1996, Brune et al. 1996).

Arguments similar to those used in quantum electrodynamics (QED) apply to *quantum gravity* (Joos 1986, Kiefer 1999 – for applications to quantum cosmology see Chap. 6). Quantum states of matter and geometry must be entangled, and give rise to mutual decoherence. The classical appearance of spacetime geometry is thus no reason *not* to quantize gravity. The beauty of Einstein’s theory can hardly be ranked so much higher than that of Maxwell’s to justify its exemption from quantization. An exactly classical gravitational field interacting with a quantum particle would be incompatible with the uncertainty relations – as has been known since the early Bohr–Einstein debate. The reduced density matrix for the metric must therefore be expected to represent an apparent mixture of different quasi-classical curvature states. Since the observer cannot avoid being correlated to them, spacetime curvature always appears to be classically given – see Sects. 4.6 and 6.2.

Moreover, the entropy and thermal radiation (of all fields) characterizing a black hole or an accelerated Unruh detector (Sects. 5.1 and 5.2) are consequences of the entanglement between relativistic vacua on two half-spaces separated by a horizon (each one forming the environment of the other). This entanglement entropy measures the same type of ‘apparent’ ensemble as the entropy produced according to the master equation (4.35) for a macroscopic mass point. The disappearance of coherence behind a horizon has nonetheless occasionally been regarded as a *fundamental* violation of unitarity, and even as the ultimate source of irreversibility (see Sects. 4.4, 5.1 and 6.2). This appears neither justified nor required (see Kiefer, Müller and Singh 1994, Kiefer 2007, Zeh 2005a).

**General Literature:** Kiefer’s Chap. 4 of Joos et al. 2003, Kiefer 2004.

#### 4.3.6 Quantum Jumps

Quantum objects are often observed by means of flashes on a scintillation screen or ‘clicks’ of a counter. These macroscopic phenomena are then interpreted as caused by pointlike objects, passing through the observing instrument during a short time interval, while this is in turn understood as evidence for a discontinuous ‘decay event’ (for example, of an atomic nucleus). A *rate equation* for such events is equivalent to a master equation, while a *constant* relative rate would describe exponential decay of the source. Discrete quantum jumps between two energy eigenstates have even been observed for single atoms in a cavity by permanently monitoring their energy, thus enforcing decoherence between energy eigenstates (Nagourney, Sandberg and Dehmelt 1986, Sauter et al. 1986, Pegg, Loudon and Knight 1986, Gleyzes et al. 2006). Therefore, formal creation and annihilation operators are often misunderstood as defining discrete events, even though they occur in a Hamiltonian that constitutes a Schrödinger equation.

This Schrödinger equation would describe a state vector that smoothly develops components with different particle numbers, or a wave function that leaks out of an unstable system (such as a quantum ‘particle’ in a potential well). This contrast between discrete events and the Schrödinger equation is clearly the empirical root of the probability interpretation of the wave function in terms of events and particles. A wave function can exponentially decay only in a limited region of space (for example within an expanding sphere for a limited time – see Sect. 4.5). This wave function is a *superposition* rather than an ensemble of different decay times. Their interference and the dispersion of the corresponding outgoing wave lead to deviations from an exponential decay law. Although these deviations are too small to be observed for decay into infinite space, interference between different decay times has often been confirmed in other situations, not least as ‘coherent state vector revival’ for photons emitted into cavities with reflecting walls (Rempe, Walther and Klein 1987).

In Sect. 4.3.1, the appearance of particles following tracks in a cloud chamber has been explained in terms of an apparent ensemble of narrow wave packets arising by means of decoherence. Similar arguments may as well explain apparently discrete events. Even if quantum objects remain isolated before being detected, they would be decohered in the detector – usually on a very short time scale. Therefore, the same decoherence that describes localization in space also explains localization in time. Jumps between discrete energy levels, observed under continuous measurement, represent apparently discrete ‘decay histories’, which can be explained by Mott-type quantum correlations between successive measurements of short but finite individual duration (including the decoherence of their outcomes). Neither particles nor genuine quantum jumps are required as *fundamental* concepts in quantum theory (Zeh 1993, Paz and Zurek 1999). Whenever decay fragments (or the decaying object) interact appropriately with their environment, interference between two partial waves

describing a decayed state and a not yet decayed state disappears on a very short (though finite) decoherence time scale, thus giving rise to an apparent *ensemble* of decay times. This time scale is in general much shorter than the time resolution of measurements.

If the decay status is thus permanently ‘monitored’ by the environment, a set of identical decaying objects is thus more appropriately described by a rate equation than by a Schrödinger equation (Sect. 4.5). This rate equation leads to an exact exponential law, since it excludes any interference between different decay times. Similarly, decay products emitted in superpositions of sufficiently different energies are absorbed into mutually orthogonal final states of the environment. Microscopic systems with their discrete energy levels must therefore decohere into eigenstates of their own Hamiltonians. This explains why the atomic world is characterized by stationary states, and von Neumann spoke of an *Eingriff* (intervention) required for their change.

So it seems that this situation of continuously monitored decay has led to the myth of quantum theory as a stochastic theory for fundamental *quantum events* (see Jadczyk 1995). Bohr (1928) remarked that “the essence” (of quantum theory) “may be expressed in the so-called quantum postulate, which attributes to *any atomic process* an essential discontinuity, or rather individuality ...” (my italics). This statement is in conflict with many microscopic and mesoscopic quantum phenomena that have since then been observed. Heisenberg and Pauli similarly emphasized their preference for matrix mechanics because of its (evidently misleading) superiority in describing discontinuities. Ole Ulfbeck and Aage Bohr (2001) recently emphasized the unpredictable occurrence of ‘clicks in the counter’, while denying the existence of any quantum events in the source that would precede them. This comes close to the consequences of decoherence, but rather than taking into account entanglement with the environment the authors conclude that “the wave function then loses its meaning”. According to the decoherence theory, the underlying entanglement processes are always smooth, and described by a Schrödinger equation. The short decoherence time scales lead to the impression of quantum *jumps* between energy eigenstates, for example, while narrow wave packets are interpreted as particles or classical variables (even though the certainty of classical properties has to be restricted by the uncertainty relations in order to comply with the Fourier theorem).

While the description of all physical phenomena in terms of time-dependent entangled wave functions now appears as a consistent picture, an important question remains: how should the probabilities, which were required to justify the concept of a density matrix in Sect. 4.2, be understood if they are *not* probabilities for quantum jumps or for the occurrence of measurement results in the form of fundamental ‘events’. This discussion will be resumed in Sect. 4.6.

**General Literature:** Joos’s Sect. 3.4.1 of Joos et al. 2003.

## 4.4 Quantum Dynamical Maps

Open systems can be phenomenologically described by means of semigroups – thus *postulating* an arrow of time. In quantum theory, they possess some novel aspects in comparison to their classical counterparts (Sect. 3.4). For example, these ‘quantum dynamical maps’ have been used to formalize von Neumann’s ‘first intervention’ (the reduction of the wave function) as part of the dynamics (Kraus 1971). This is possible, since semigroups can not only describe the transition of pure states into ensembles, but also the ‘selection’ of an *individual* element from them (see below). Otherwise they are equivalent to an entropy-enlarging Zwanzig-type master equation with respect to the corresponding  $\hat{P}_{\text{sub}}$ . Although ‘irrelevant’ correlations with the environment, which would arise according to the unitary global dynamics, now represent quantum entanglement, they are usually not distinguished from classical statistical correlations when it comes to applications.

This confusion of concepts is equivalent to a popular but insufficient ‘naive’ interpretation of decoherence, which pretends to derive genuine ensembles. Quantum dynamics is occasionally even *defined* in terms of semigroups, assumed to act on the density matrix as a fundamental kinematical object characterizing quantum systems. (Hence the term ‘statistical operator’ for the density operator.) However, this ‘minimal statistical interpretation’ entirely neglects the difference between genuine and apparent ensembles, and thus all consequences of entanglement beyond the considered systems (quantum non-locality). Even the superposition principle has been claimed to be derivable in this formalism (Ludwig 1990), although it is then simply reintroduced in a hidden form (for example by changing the laws of statistics in an unjustified way).

Semigroups are certainly mathematically elegant and powerful. Therefore, they would form candidates for *new* theories if conventional (Hamiltonian) quantum theory should prove wrong empirically as a universal theory. The question is whether mathematical elegance already warrants physical relevance or is merely convenient within a certain approximation. To quote Lindblad (1976): “It is difficult, however, to give physically plausible conditions ... which rigorously imply a semigroup law of motion for the subsystem. ... Applications ... have led some authors to introduce the semigroup law as the fundamental dynamical postulate for open (non-Hamiltonian) systems.” Such a law would *fundamentally* introduce an arrow of time (see also Sect. 4.6), but it would depend on the choice of ‘systems’ – and may be at variance with certain experiments which confirm quantum nonlocality.

The simplest quantum systems (such as spinors) are described by a two-dimensional Hilbert space. Their density matrix may be written by means of the Pauli matrices  $\sigma_i$  ( $i = 1, 2, 3$ ) in the form

$$\rho = \frac{1}{2}(1 + \boldsymbol{\sigma} \cdot \boldsymbol{\pi}) , \quad (4.38)$$



where the (mathematically) real *polarization vector*  $\boldsymbol{\pi} = \text{Trace}\{\boldsymbol{\sigma}\rho\}$  – that is, the expectation value of *all* spin components – completely defines  $\rho$  as a general Hermitean  $2 \times 2$  matrix of trace 1. The latter is in turn equivalent to a (genuine or apparent) *ensemble* of spinors. The length of  $\boldsymbol{\pi}$  is a measure of the purity of the ‘mixed state’  $\rho$ , since  $\text{Trace}\{\rho^2\} = (1 + \boldsymbol{\pi}^2)/2$ , with  $\boldsymbol{\pi}^2 \leq 1$ . A pure state corresponds to a unit polarization vector, while an arbitrary density matrix (a general ‘state’ in the language of mathematical physics) is characterized by the mean value  $\boldsymbol{\pi} = \sum_{\alpha} p_{\alpha} \boldsymbol{\pi}_{\alpha}$  of all unit vectors  $\boldsymbol{\pi}_{\alpha}$  in an ensemble of spinors that may represent this density matrix.

A general trace-preserving linear superoperator  $\hat{P}$  acting on  $\rho$  must be defined on 1 and  $\boldsymbol{\sigma}$  in order to be completely defined:

$$\hat{P}1 := 1 + \boldsymbol{\pi}_0 \cdot \boldsymbol{\sigma} , \quad \hat{P}\boldsymbol{\sigma} := \mathbf{A} \cdot \boldsymbol{\sigma} , \quad (4.39)$$

with a real vector  $\boldsymbol{\pi}_0$  and a linear vector transformation  $\mathbf{A}$ .  $\hat{P}$  is idempotent (a Zwanzig ‘projector’) if  $\mathbf{A}^2 = \mathbf{A}$  and  $\boldsymbol{\pi}_0 \cdot \mathbf{A} = 0$  ( $\mathbf{A} = 0$ , for example). If  $\boldsymbol{\pi}_0 \neq 0$ ,  $\hat{P}$  creates new information – even from the unit matrix (see Sect. 3.2).

Dynamical combination of the projection  $\hat{P}$  with a Hamiltonian evolution (which would describe a rotation of  $\boldsymbol{\pi}$ ) in the form of a master equation leads to the *Bloch equation* for the vector  $\boldsymbol{\pi}(t)$ ,

$$\frac{d\boldsymbol{\pi}}{dt} = \boldsymbol{\omega} \times (\boldsymbol{\pi} - \boldsymbol{\pi}_0) - \sum_i \gamma_i (\pi^i - \pi_0^i) \mathbf{e}_i , \quad (4.40)$$

in a certain vector basis  $\{\mathbf{e}_i\}$  (see Gorini, Kossakowski and Sudarshan 1976). Values of  $\gamma_i < 0$  or  $|\boldsymbol{\pi}_0| > 1$  would violate the positivity of the density matrix at some  $t > 0$ , and thus have to be excluded.<sup>5</sup> The second term on the RHS describes anisotropic ‘damping’ towards  $\boldsymbol{\pi}_0$ . This formal creation of information (which is contained in the vector  $\boldsymbol{\pi}_0$ ) may represent very different physical situations, such as equilibration with an external heat bath of given (possibly lower) temperature, or evolution towards a certain measurement result. The density matrix defined by the polarization vector  $\boldsymbol{\pi}_0$  is often called a *reference state*, while the *relative entropy* with respect to it [see (3.55)] never decreases under the Bloch equation – even when the physical entropy of the local spinor system does. Hermiticity of  $\hat{P}$  (corresponding to a genuine, entropy-raising projection operator) would require  $\boldsymbol{\pi}_0 = 0$  and  $\mathbf{A} = \mathbf{A}^\dagger$ , that is, a projection of the vector  $\boldsymbol{\pi}$  onto a specific component.

If the two-dimensional Hilbert space describes something other than spin, such as isotopic spin, a  $K, \bar{K}$  system, or fermion occupation numbers, the

<sup>5</sup> As mentioned in Sect. 4.2, *all* subsystem density matrices remain positive under global Hamiltonian dynamics, and even under a collapse of the global state vector. This property of ‘complete positivity’ has to be separately *postulated* for phenomenological quantum dynamical maps (see Kraus 1971), thus further illustrating the fact that these maps *cannot* be regarded as representing fundamental physics.



polarization vector lives in an *abstract* three-dimensional space that usually cannot simply be ‘rotated’ in practice. The abstract formalism can also be generalized to  $n$ -dimensional Hilbert spaces. For this purpose the Pauli matrices have to be replaced by the  $(n^2 - 1)$  Hermitean generators of the corresponding group  $SU(n)$ , while the real ‘coherence vectors’ (the generalizations of the polarization vector  $\boldsymbol{\pi}$ ) now live in the vector space spanned by them. For example,  $SU(3)$  gives rise to the ‘eight-fold way’. The most important *new* property then is that there are more than one (in fact,  $n - 1$ ) *commuting* Hermitean generators. They may contain a nontrivial subset that is decohered under all realistic environmental conditions, and thus may form the center of a phenomenological set of observables (the set of ‘classical observables’ – see Sect. 4.3). For example, maps of density matrices of dimension  $n = 4$  which happen to completely decohere with respect to a certain basis may reproduce the classical maps of Figs. 3.8a, c and d.

In the infinite-dimensional Hilbert space of quantum mechanics, the *Wigner function*

$$\begin{aligned} W(p, q) &:= \frac{1}{\pi} \int e^{2ipx} \rho(q + x, q - x) dx \\ &\equiv \frac{1}{2\pi} \iint \delta\left(q - \frac{z + z'}{2}\right) e^{ip(z - z')} \rho(z, z') dz dz' \\ &= \text{Trace}\{\Sigma_{p,q} \rho\}, \end{aligned} \quad (4.41)$$

where the third line is written in analogy to  $\boldsymbol{\pi} = \text{Trace}\{\boldsymbol{\sigma} \rho\}$ , assumes the role of the coherence vector  $\boldsymbol{\pi}$ . Evidently,

$$\Sigma_{p,q}(z, z') := \frac{1}{2\pi} e^{ip(z - z')} \delta\left(q - \frac{z + z'}{2}\right) \quad (4.42)$$

is a generalization of the Pauli matrices (with the vector index replaced by  $p, q$ ). On a finite  $q$ -interval of length  $L$ ,  $\Sigma_{p,q}$  would require an additional term  $-(1/2\pi L)e^{ip(z - z')}$  in order to warrant tracelessness.

The Wigner function is thus a continuous set of expectation values of these generalized Pauli matrices. They form the components (one for each phase space point) of a generalized coherence vector. This ‘vector’ characterizes the density matrix  $\rho$  completely – just as in (4.38) and regardless of its interpretation. Although it does *not* represent a probability density on phase space, as illustrated by its possibly negative values, one may calculate all expectation values in the *form* of a mean value for a classical function of state  $f(p, q)$ , viz.,  $\langle F \rangle = \int f(p, q) W(p, q) dp dq$ . However, a ‘quasi-classical’ Gaussian wave packet, for example, is a coherent quantum mechanical superposition of position or momentum eigenstates in spite of its (in this case) non-negative Wigner function.

Lindblad (1976) was indeed able to generalize the Bloch equation to infinite-dimensional Hilbert spaces. He wrote it (in a form that applies to the density matrix) as

$$i\frac{\partial\rho}{\partial t} = [H, \rho] - \frac{i}{2} \sum_k \left( L_k^\dagger L_k \rho + \rho L_k^\dagger L_k - 2L_k \rho L_k^\dagger \right), \quad (4.43)$$

with arbitrary generators  $L_k$  in Hilbert space. It describes a *creation* of information, that is, a local decrease of the corresponding von Neumann entropy, if and only if some generators do not commute with their Hermitean conjugates  $L_k^\dagger$ . This can be shown by applying the non-Hamiltonian terms of (4.43) to the unit matrix  $\rho = 1$ , which describes *no* information. Otherwise this equation describes information loss (a genuine Zwanzig projection).

This can also be seen by means of the general representation of a Zwanzig projector on density operators in quantum mechanical Hilbert space,  $\hat{P}\rho = \sum_k V_k \rho V_k^\dagger$ . It is similar to the square root of a positive operator, written in its eigenbasis. If  $L_k^\dagger = L_k$ , the non-Hamiltonian Lindblad terms assume the form of a double commutator,  $L^2\rho + \rho L^2 - 2L\rho L = [L, [L, \rho]]$ . For  $L = \sqrt{2\lambda}x$  one recovers (4.35), that is, decoherence in the  $x$ -basis, precisely as *derived* from unitary interaction with the environment.

In this way one may, in particular, describe transitions of pure states into formal ensembles of measurement results with their corresponding Born probabilities. However, ‘damping’ towards a definite *pure* state (a semigroup proper in the sense of Fig. 3.8c), would require the second term of (4.40) with a unit vector  $\pi_0$ . It allows one to describe the evolution into a (freely chosen) definite measurement outcome. This dynamics can then readily be combined with a stochastic formalism that is defined to select the possible final states of a measurement in accordance with the Born rules (Bohm and Bub 1966, Pearle 1976, Gisin 1984, Belavkin 1988, Diósi 1988). If applied continuously, such as by means of the Itô process, this formalism describes a genuine collapse as a smooth but indeterministic process (Pearle 1989, Ghirardi, Pearle and Rimini 1990). If this modification of the Schrödinger equation were correct, it should in principle be observable, although it would usually be camouflaged by environmental decoherence (Joos 1986, Tegmark 1993).

Many explicit collapse models of this kind have been proposed in the literature. Some remain ambiguous about their true intentions (that is, whether they are meant fundamental or phenomenological), or simply disregard the difference between genuine and apparent ensembles (proper and improper mixtures). In particular, the *quantum state diffusion model* (Gisin and Percival 1992) presumes that reduced density matrices can be ‘untangled’ into genuine ensembles (with only one of their members assumed to represent reality). However, this would again be equivalent to a modification of the global Schrödinger equation (Diósi and Kiefer 2000).

**General Literature:** Alicki and Lendi 1987, Diósi and Lukács 1994, Stamatescu’s Chap. 8 of Joos et al. 2003.

## 4.5 Exponential Decay and ‘Causality’ in Scattering

There are only a few absolutely stable ‘particles’ (elementary quantum objects), while all others are known as decaying on vastly different time scales. In quantum theory, they may be described formally by means of *complex energies*. For a Schrödinger type time dependence  $e^{-iEt}$ , a negative imaginary part,  $E = E_0 - i\gamma$  with  $\gamma > 0$ , would lead to an exponentially decreasing wave function. This does not just describe probabilities for different decay times, since all parts of the wave function form *one* coherent superposition (see below and Sect. 4.3.6). Even though microscopic, these objects have to be regarded as *open* quantum systems. For example, an excited atom is coupled to an initial vacuum (or a photon heat bath of zero temperature). Unbounded space represents an ‘absorber’ of infinite capacity for the decay fragments.

The decaying system may also be described by means of an  $S$ -matrix for the decay fragments, where unstable states show up as poles in the complex energy or momentum plane. This  $S$ -matrix must represent the fundamental (time-symmetric) dynamics. Exponential decay then seems to characterize a fundamental direction in time (see Prigogine 1980, for example), similar to Ritz’s retarded electrodynamics (Chap. 2). Since there are no energy eigenstates with complex eigenvalues (sometimes called ‘Gamow vectors’) in Hilbert space, this situation has even led to the proposal of ‘rigged Hilbert spaces’ (Böhm 1978). However, decaying systems may well be described in conventional quantum mechanical terms, where the exponential time dependence applies only approximately in a limited spacetime region.

Exponential decay of an arbitrary quantity  $A$  would be the consequence of a constant loss rate, described by

$$\frac{dA}{dt} = -\lambda A, \quad (4.44)$$

with  $\lambda > 0$ . The absolute rate of change,  $dA/dt$ , is then completely determined by  $A$  itself. This asymmetry under time reversal may be the consequence of a special *initial* condition, similar to that characterizing irreversible master equations. In particular, if  $A$  is a conserved quantity, any back-flow, must be negligible. This condition represents a fact-like T-asymmetry that may be explained by assuming a sufficiently large and initially empty reservoir (comparable to the ‘irrelevant channel’ used in Sect. 3.2). If recurrence times are sufficiently large, the exponential law (4.44) may remain an excellent approximation, describing the decaying object for a very long time.

This disappearance of a ‘substance’  $A$  from a given subsystem or region in space is an entirely classical model. However, the time dependence (4.44) is best known from radioactive decay in quantum theory, where  $A$  represents the non-decay *probability*. It is then regarded as the standard example of *quantum indeterminism* – usually understood as fundamental and law-like. This interpretation of (4.44) would mean that decay *events* occur at unpredictable though definite instants in time.

The decay law (4.44) defines an elementary master equation (3.48) with a Green's function  $\hat{G}_{\text{ret}}$  simply given by the decay rate  $\lambda$  (see Sect. 4.1.2). Its foundation on time-symmetric fundamental dynamics (such as a universal Schrödinger equation) requires quite analogous assumptions, for example the negligibility of any back-flow into 'doorway states' that are directly coupled to  $A$  (see Fig. 3.4). Therefore, a conserved quantity has to disappear fast enough from such doorway states into 'deeper' (dynamically more distant) states, which must form a large reservoir.

A simple model is provided by the  $T$ -symmetric finite reaction chain

$$\frac{dA_n}{dt} = -(\lambda_n + \lambda_{n-1})A_n + \lambda_n A_{n+1} + \lambda_{n-1} A_{n-1}, \quad (4.45)$$

with  $n = 0, \dots, N$ ,  $\lambda_{-1} = \lambda_N = 0$ , and the (improbable) initial condition  $A_{n \neq 0} \approx 0$ .  $n = 1$  represents here the doorway channel of Sect. 3.2. For  $\lambda_0 \ll \lambda_{n \neq 0}$ , one obtains

$$\frac{dA_0}{dt} \approx -\lambda_0 A_0, \quad (4.46)$$

as long as  $A_1 \ll A_0$ . This requires only  $\lambda t \ll N$ , rather than  $\lambda t \ll 1$ , since all  $A_{n \neq 0}$  will relax into partial equilibrium  $A_{n \neq 0} \approx A_1$  on a short time scale (or just propagate away for  $N \rightarrow \infty$ ).

Exponential decay can similarly be described by a deterministic wave equation on a continuum, where the small transition rate  $\lambda_0$  is replaced by a potential barrier. It is irrelevant that the Schrödinger equation does here not describe the conserved quantity ('probability') itself. An *overall* time dependence according to a complex energy eigenvalue,  $\psi(t) \propto \exp[-i(E_0 - i\gamma)t]$ , would not be compatible with unitarity, but it may well represent an approximation that is valid in a bounded though growing spacetime region (Khalfin 1958, Petzold 1959, Peres 1980a) – similar to the reaction chain (4.45). Distant regions in space form a large reservoir.

In scattering theory, unstable states correspond to poles of the  $S$ -matrix  $S_{nn'}(k)$ , analytically continued into the complex plane, at points  $k = k_1 - ik_2$  in the lower right half-plane ( $k_1 > 0$  and  $k_2 > 0$ ), where  $k$  is the wave number,  $k^2 = k_1^2 - k_2^2 - 2ik_1k_2 = 2mE$ . In the restricted spacetime region, where exponential behavior is observed after the incoming waves producing the decaying system have ceased, the wave function is dominated by the *Breit-Wigner* part (i.e., the pole contribution). This requires a (positive) *time delay* during the scattering process, which must be described by the relevant partial wave  $\psi_l(r, t) Y_{lm}(\theta, \phi)$ . Its radial factor  $\psi_l(r, t)$  may be expanded in terms of energy eigenstates,  $\psi_l^{(k)}(r, t) := \phi_{k,l}(r) e^{-i\omega(k)t}$ , in the form

$$\begin{aligned} \psi_l(r, t) &= \int_0^\infty f_l(k) \psi_l^{(k)}(r, t) dk \\ &\xrightarrow{r \rightarrow \infty} \int_0^\infty f_l(k) \frac{e^{-ikr} - (-1)^l S_l(k) e^{ikr}}{r} e^{-i\omega(k)t} dk, \end{aligned} \quad (4.47)$$

where  $S_l(k) = e^{2i\delta_l(k)}$  is the corresponding diagonal element of the  $S$ -matrix.

For sufficiently large values of  $t$ , the factor  $e^{-i\omega(k)t}$  oscillates rapidly with  $k$ . This leads to destructive interference under the integral, except in regions of  $r$  and  $t$  where the phase  $kr + \omega(k)t$  (for the incoming wave), or  $kr - \omega(k)t + 2\delta_l(k)$  (for the outgoing one), is almost independent of  $k$  over the width of the wave packet  $f_l(k)$  (which may be centered at  $k_0$ , say). For the outgoing wave, for example, this requirement means

$$\left. \frac{d}{dk} [kr - \omega(k)t + 2\delta_l(k)] \right|_{k_0} \approx 0 \quad \Rightarrow \quad r \approx \frac{d\omega(k_0)}{dk_0} t - 2 \frac{d\delta_l(k_0)}{dk_0}. \quad (4.48)$$

A noticeable delay compared to propagation with the group velocity  $d\omega/dk$  requires a large value of  $d\delta_l/dk$ , such as in the vicinity of a complex pole of  $\delta_l$ . For sufficiently large times  $t$ , but not too large distances  $r$  from the scattering center, and for initial momentum packets much wider than the size of the imaginary part  $k_2$ , only the pole contribution remains. For this one may write

$$S_l(k) = e^{2i\delta_l(k)} \approx \frac{k - k_1 - ik_2}{k - k_1 + ik_2}, \quad (4.49)$$

and hence  $k_0 = k_1$  for the surviving wave packet that represents the decaying state. In this spacetime region, the contribution of the pole to (4.49) is given by its residue, whence

$$\begin{aligned} \psi_l(r, t) &\xrightarrow{t \rightarrow \infty} -(-1)^l f_l(k_1) \int_0^\infty \frac{k - k_1 - ik_2}{k - k_1 + ik_2} \frac{e^{i[kr - \omega(k)t]}}{r} dk \\ &\approx (-1)^l 2\pi k_2 f_l(k_1) \frac{e^{i[k_1 r - \omega(k_1)t]}}{r} \exp \left[ k_2 \left( r - \frac{d\omega(k_1)}{dk_1} t \right) \right] \end{aligned} \quad (4.50)$$

(assuming  $k_2 \ll |k_1|$ ). In the last factor one recognizes the ‘imaginary part of the energy’,  $\gamma = k_2 d\omega(k_1)/dk_1$ .

A positive delay (a ‘retardation’) of the scattered wave at the resonance requires

$$\frac{d\delta_l}{dk} \approx -\frac{d}{dk} \left( -\arctan \frac{k_2}{k - k_1} \right) = \frac{k_2}{(k - k_1)^2 + k_2^2} > 0. \quad (4.51)$$

The pole must therefore reside in the lower half-plane. This condition is often referred to as *causality in scattering*, since the retardation specifies a direction in time related to intuitive causality (Chap. 2). This position of the poles is also used for deriving *dispersion relations* in  $T$ - or  $TCP$ -symmetric quantum field theory. However, no time direction can be specified by the structure of the  $S$ -matrix, since the latter is a consequence of the time-reversal-invariant Hamiltonian. Exponential decay is a fact-like asymmetry that would be reversed, using the *same*  $S$ -matrix, for scattering states with a time-reversed boundary condition. So one would have to force the outgoing wave rather

than the incoming one to form a wave packet of limited duration. This would require the former decay products to arrive in the form of *advanced* coherent and exponentially growing waves over a very long span of time. A similar fact-like asymmetry, also caused by boundary conditions, occurred for the retarded radiation reaction of extended charges (2.31). The time arrow of exponential decay is thus determined by the boundary condition used in (4.47), rather than by the position of poles in the  $S$ -matrix.

The investigation of wave packets for non-relativistic particles decaying into free space beyond the pure pole contribution (Petzold 1959, Winter 1961, Peres 1980a) shows that deviations from the exponential law are essential not only at and just after formation of the decaying object, but also for very large times. In the extreme limit (when the decaying wave function has become unobservably small), exponential decay would be replaced by a power law. This consequence is in conflict with the interpretation of decay as a stochastic emission of particles (see Sect. 4.3.6). It must instead be understood as a *coherent back-flow* according to the dispersion of the outgoing wave. Although representing a very small effect in absolute terms, this deviation from exponential decay is even further reduced whenever the decay products interact with surrounding matter. In the usual case of strong coupling to the environment (such as absorption or decoherence), the exponential law remains an excellent approximation as long as the thermodynamical arrow characterizing absorbers remains valid. On the other hand, decay by emission of weakly interacting photons inside reflecting walls of a cavity has been confirmed to lead to the predicted deviations from exponential decay. It may even cause a ‘coherent revival’ of the decaying state (see Rempe, Walther and Klein 1987, Haroche and Kleppner 1989), but has also been observed in other situations (Wilkinson et al. 1997). Deviations from exponential decay depend crucially on the density of available final states. For example, coherent decay into a single final state is well known to lead to harmonic oscillation (complete periodic back-flow).

The Breit–Wigner contribution (4.50) describes a non-normalizable (even exponentially increasing) wave function. This result is an artifact of the pure pole approximation. The normalized state is correctly described by the wave packet  $f_l(k)$  in (4.47), which has been replaced by a constant in (4.50). In an exact treatment, its square-integrable tails warrant normalizability even for large  $t$  by correcting the pure Breit–Wigner contribution.

In the general case, the scattering process would have to be described by a normalized time-dependent density matrix,

$$\begin{aligned} \rho_{lm,l'm'}(r, r'; t) &\xrightarrow{r \rightarrow \infty} \int_0^\infty \int_0^\infty \rho_{lm,l'm'}(k, k') \frac{e^{-ikr} - (-1)^l S_l(k) e^{ikr}}{r} \\ &\times \frac{e^{ik'r'} - (-1)^{l'} S_{l'}^*(k') e^{-ik'r'}}{r'} e^{-i[\omega(k) - \omega(k')]t} dk dk', \end{aligned} \quad (4.52)$$

where  $\rho_{lm,l'm'}(k, k')$  is determined by the preparation procedure. After completion of the direct scattering process, and in the case of a resonance in the

$l_0$ -wave, this leads approximately to

$$\begin{aligned} \rho_{lm,l'm'}(r, r'; t) &\xrightarrow[t \rightarrow \infty]{} \delta_{ll_0} \delta_{l'l_0} \rho_{l_0 m, l_0 m'}(k_1, k_1) \\ &\times \int_0^\infty \int_0^\infty \frac{k - k_1 - ik_2}{k - k_1 + ik_2} \frac{k' - k_1 + ik_2}{k' - k_1 - ik_2} \frac{e^{i(kr - k'r')}}{rr'} e^{-i[\omega(k) - \omega(k')]t} dk dk', \end{aligned} \quad (4.53)$$

in the relevant spacetime region. This approximation describes once again a *pure* Breit–Wigner wave packet (or at most a mixture of magnetic quantum numbers in the case of rotational symmetry).

Hence, there are no exactly exponential states (Gamow vectors) which would require or justify the ‘rigging’ of the Hilbert space of quantum mechanics. Similarly, there are no exact energy eigenstates in reality, since their infinite exponential tails according to  $\exp(-\sqrt{-2mE}r)$  can never form completely within finite time. If exact energy eigenstates did occur by means of instantaneous quantum jumps, they would lead to superluminal effects (as has even been found surprising – see Hegerfeldt 1994). Even non-relativistically, stable and decaying states must form dynamically, that is, in accordance with a time-dependent Schrödinger equation, and hence with the time–energy uncertainty relation. Relativistically, an exactly bounded spatial support for a quantum state requires small (usually unobservable) uncertainties in energy and particle number, related to the Casimir effect or the Unruh radiation (see Sect. 5.2). Neglecting this consequence of quantum field theory (by arguing solely in terms of single-particle states, for example) leads to inconsistencies which illustrate the danger of remaining mathematically exact while not distinguishing between hypothetically fundamental and phenomenological (approximately valid) concepts – see also Sect. 4.3.2.

Radioactive decay is investigated in practice by means of *ensembles of many objects* in identical internal states (unstable nuclei, say). If these objects are distinguishable, for example by their position, their total state may be described as a direct product. According to the Schrödinger equation, each factor state will then evolve into a superposition of its initial state (with an exponentially decreasing amplitude) and a direct product of its final state and outgoing waves for the emitted objects. The *precise* time-dependent form of all components depends on what happens to the decay fragments (whether they are reflected, absorbed, or propagate into infinite space).

If the emitted particles (now assumed *not* to be reflected somewhere) are regarded as part of the environment of the decaying system, and thus traced out, the remaining  $N$ -atom density matrix describes a time-dependent *apparent ensemble* of different direct products. An ensemble variable  $n(t) < N$  may here count the number of undecayed nuclei at each value of  $t$  (that is, in an ensemble of objects rather than potential states). The formal probabilities of these states in their apparent ensemble,  $p_n(t)$ , must then reflect the time-dependence of the Schrödinger amplitude, for example  $p_0(t) = e^{-2N\gamma t}$ . (In general, components with the same number  $n$ , but *different individual* decayed nuclei, also decohere from one another – see Sects. 4.3.5 and 4.3.6.)

This apparent ensemble of discrete numbers is dynamically approximately described by a master equation. Therefore, it is formally equivalent to an ensemble of solutions of a stochastic (Langevin-type) equation that essentially describes individual discrete ‘histories’  $n(t)$  – here in the form of ‘descending staircase functions’. In terms of a universal Schrödinger equation, the number of undecayed nuclei  $n$  is a ‘robust’ property in the sense of Sect. 4.3.2 if decay can be assumed to be irreversible (in particular when monitored by detectors). The various dynamically robust branches of the wave function, arising by the fast but smooth action of decoherence, describe individual histories for integer numbers  $n(t)$ , which represent successions of almost discrete quantum jumps at certain times  $t_1, t_2, \dots$  (as discussed in Sect. 4.3.6). Similar staircase functions have now also been observed for decaying photons in a cavity (Gleyzes et al. 2006) – thus directly confirming Fig. 3.30 of Joos et al. (2003). However, deviations from exact steps can always be calculated if the interaction with the environment is known (Joos 1984): quantum theory is *not* a stochastic theory for quantum jumps.

## 4.6 The Time Arrow in Various Interpretations of Quantum Theory

The truth could not be worth much  
if everybody was a bit right.

Physicists who completely agree about all applications of quantum mechanics often differ entirely about its interpretation, and even on the question of whether there remain *any* meaningful problems beyond the mere formalism (see Fuchs and Peres 2000). Although most of them would agree that quantum theory allows no more than probabilistic predictions, they often *derive* irreversible master equations, which describe an increase in entropy, from the deterministic and time-symmetric Schrödinger equation, using special initial conditions as in classical statistical physics (see Sect. 4.1.2). However, a dynamical probability interpretation must be relevant for the arrow of time – regardless of whether it is based on a fundamental stochastic (time-asymmetric) law or on an incompleteness of the theory (hidden variables) that refers to an unknown future. Its consequences cannot be avoided just by adding new words. For example, quantum theory is often called ‘deterministic but acausal’ – while this statement is then justified by the ‘uncertainty’ of classical properties (such as particle positions or momenta), which just *do not apply* to quantum states. Most physicists seem to disregard this consistency problem in an act of *Verdrängung*.

The deepest roots of these conceptual inconsistencies seem to arise from the fundamental difference between Heisenberg’s and Schrödinger’s ‘pictures’ (see Zeh 2004). While Heisenberg maintained classical concepts in principle (suggesting only a limitation of the ‘certainty’ of their values), Schrödinger



described microscopic physical states by wave functions, which can be regarded as certain. The classical configuration space on which they are usually defined would thereby replace three-dimensional space as a new ‘arena of dynamics’ rather than describing *potential states*.<sup>6</sup> Whether wave functions or ‘observables’ (which *formally* replace the classical variables in the Heisenberg picture) carry the dynamical time-dependence is merely a *consequence* of the chosen picture.<sup>7</sup>

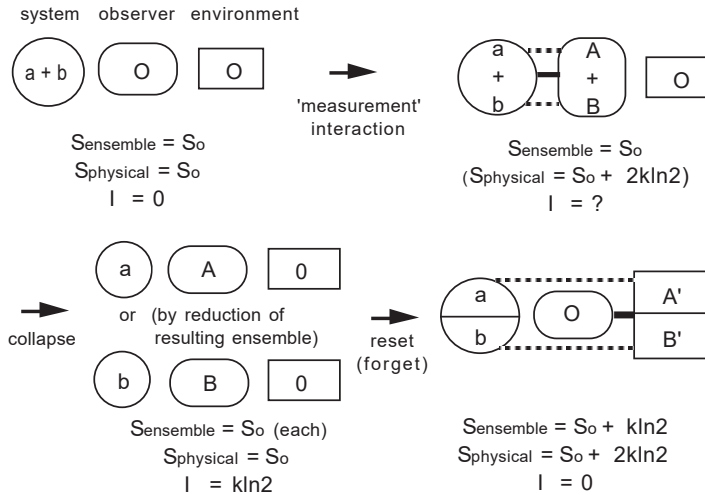
Although both pictures are equivalent when used to calculate formal expectation values for isolated systems, they describe the time arrow of quantum measurements in different ways. Most physicists seem to subscribe to one or the other picture (or perhaps a variant thereof) when it comes to interpretations (‘probabilities for *what?*’). Typically, in the Schrödinger picture one regards the collapse of the wave function as a dynamical process, while in the Heisenberg picture it is viewed as an (extraphysical) increase of ‘human knowledge’. I hope that keeping this difference in mind for the rest of this section may help to avoid some misunderstandings that often lead to emotional debate. One should therefore concentrate on what is actually *done* when the theory is successfully applied – though not in a merely pragmatic way. Which concepts are *fundamentally required*, rather than being approximately justified, or even mere tradition and prejudice?

Any meaningful concept of incomplete information or knowledge has to refer to an *ensemble* of possible states. For example, physical entropy, which quantifies irreversibility, is in quantum statistical mechanics defined by means of von Neumann’s functional of the density matrix (4.4). According to Sect. 4.2, it measures the size of (genuine or apparent) ensembles of mutually orthogonal (hence operationally distinguishable) *wave functions*. While only *genuine* ensembles represent incomplete information, the time-dependence of the density matrix determines that of local entropy in general. Conservation of *global* von Neumann entropy reflects the unitarity of the von Neumann equation (when applicable) – equivalent to the unitarity *and* determinism of the Schrödinger equation. No ensemble of classical or any other (unknown)

---

<sup>6</sup> The identity of configuration space and space in single particle quantum mechanics is a consequence of the exceptional kinematics of mass points. This has led to a popular confusion of single-particle wave functions with spatial fields, and to the misnomer of a ‘second quantization’ in quantum field theory – see Zeh (2003).

<sup>7</sup> This contrast between the Heisenberg and the Schrödinger pictures has to be distinguished from the ‘dualism’ between two competing classical concepts (particles and fields) that is part of *one* (the Copenhagen) interpretation. In classical theory, particle positions and field strengths characterize *different physical objects*, which are both constituents of general physical systems. A dualism (or ‘complementarity’), apparently required to characterize quantum objects, should more correctly be understood as a conceptual inconsistency, often attributed to a ‘lacking microscopic reality’. However, this conceptual dualism applies only to the ‘phenomenological reality’ (see Sect. 4.3.2). A critical account of the origin of these conceptual problems can be found in Beller (1996, 1999).



**Fig. 4.3.** Quantum measurement of a *superposition*  $(|a\rangle + |b\rangle)/\sqrt{2}$  by means of a *collapse* process, here assumed to be triggered by the macroscopic pointer position. The initial entropy  $S_0$  is smaller by one bit than in Fig. 3.5 (and may in principle vanish), since there is no initial ensemble ‘*a or b*’ for the property to be measured. Dashed lines before the collapse now represent quantum entanglement. (Compare the ensemble entropies with those of Fig. 3.5!) The collapse itself is often divided into two steps – see (4.54) below: first increasing the ensemble entropy by replacing the superposition by an ensemble, and then lowering it by reducing the ensemble (applying the ‘or’ – for macroscopic pointers only). The total increase of ensemble entropy, evident in the final diagram, is a consequence of the first step of the collapse. It brings the entropy up to its classical initial value of Fig. 3.5. The reset here illustrates also why decoherence is usually irreversible even when a measurement result is ‘erased’ (and even without a collapse – in which case the final ensemble entropy would again be  $S_0$ ). From Chap. 2 of Joos et al. (2003)

variables representing the potential values of observables is ‘counted’ by von Neumann’s entropy. Figure 3.5, characterizing classical measurements, *cannot* therefore be applied to quantum measurements. In terms of quantum states it has to be replaced by Fig. 4.3, which includes a collapse of the wave function. The transition from a superposition to an ensemble (depicted by the second step) affects the final value of von Neumann’s ‘ensemble’ entropy (that would be reduced by a mere increase of information, as in the first step of Fig. 3.5). For similar reasons there can be no ‘postselection’ (no retarded increase of information about the past) by a quantum measurement, as suggested by Aharonov and Vaidman (1991): there is nothing to ‘select’ from in the absence of an ensemble of hidden variables.

A wave function *and* a set of classical configurations are kinematically used in Bohm’s quantum theory (Bohm 1952, Bohm and Hiley 1993). This theory is often praised for exactly reproducing all predictions of conventional

quantum theory in a deterministic way. However, this is not surprising, since it leaves Schrödinger's wave function entirely unchanged, while the assumed trajectories for classical states, which would *determine* all observed quantities according to this model, have to remain unobservable and in drastic conflict with classical intuition ('surrealistic') in order to reproduce the empirically confirmed quantum probabilities by means of their *postulated* statistical distribution. Because of the 'phenomenological' wave-particle dualism (see Sect. 4.3.2), it also remains controversial in this theory whether the classical configurations must contain photon positions or electromagnetic fields (Holland 1993).

Although wave functions and trajectories in configuration space are equally assumed to be real in this theory,<sup>8</sup> they are *treated* quite differently. While the former are usually regarded as 'given', the latter are always represented by an ensemble (without thereby contributing to the entropy). Their initial probability distribution in this ensemble, which has to be regarded as incomplete information, is *postulated* to comply with the Born rule. Since the Bohm trajectories themselves remain unobservable, they can be said to serve as no more than artificial and empirically unfounded selectors for the 'active' branch of the global wave function, to which the actual trajectory would be confined according to its dynamics. For example, entropy is calculated, in the form  $S[\hat{P}|\psi\rangle\langle\psi|]$  with an appropriate Zwanzig projection  $\hat{P}$ , from such a *component*  $\psi$  – as though the wave function had been reduced by a real collapse (see Sect. 5 of Dürr, Goldstein and Zanghi 1993). While this description requires the same fact-like time asymmetry of the global wave function as decoherence, the selection of subsets of trajectories defines an *external* arrow of time. A justification of this different treatment of wave functions and Bohm trajectories is not at all obvious (see Zeh 1999b).

Similarly to Bohm's theory, collapse theories (Pearle 1976, Ghirardi, Rimini, and Weber 1986) and the Everett interpretation (Everett, 1957) also assume the wave function to represent a real physical object. This is in contrast to genuine hidden variables theories, which intend to *derive* or *explain* the wave function from some (hoped-for) more fundamental level of description. These latter theories are affected by various no-go theorems (such as Bell's theorem) if they are assumed to be local. Otherwise, however, it is hard to see what could be gained from them in comparison to the global wave function itself as a nonlocal object.

While collapse theories propose stochastic modifications of the Schrödinger equation, the Everett interpretation is based on the concept of 'splitting ob-

---

<sup>8</sup> "No one can understand this theory until he is willing to think of  $\psi$  as a real objective field rather than just a 'probability amplitude' " (Bell 1981). This statement may apply to quantum theory in general – quite in contrast to an analogous statement about Bohm's trajectories, which are empirically unfounded and thus have no more than 'religious' status. A stochastic theory can *always* be deterministically completed by means of unobservable variables: any by definition unobservable (pseudo-)random number generator will do.

servers’ – a quite natural consequence in a kinematically nonlocal theory if observers have to remain local for dynamical reasons (see also Chap. 6). This splitting is facilitated by means of decoherence, defined as the dislocalization of superpositions (Sect. 4.3). Because of the locality of interactions, this process describes an effective dynamical decoupling (called a ‘branching’) of the global wave function into components that are characterized by different quasi-classical (robust and quasi-local) properties – including those of systems that can be regarded as observers. This decoherence has turned out to be the most efficient and most ubiquitous irreversible process in Nature.

While decoherence eliminates the basic motivation for the Heisenberg–Bohr interpretation (that *presumes* genuine classical concepts, and their values as ‘coming into being’ during fundamental irreversible events outside the laws of physics – see Pauli’s remark towards the end of the Introduction), this aspect of decoherence may not have been duly appreciated even by some authors who significantly contributed to it (Omnès 1988, 1992, Gell-Mann and Hartle 1990, 1993 – see also Omnès 1998). Apparently guided by the Heisenberg picture, they investigate consequences of decoherence for certain ‘consistency conditions’ (originally proposed by Griffiths 1984), which are assumed to regulate the applicability (or different kinds of ‘truth’) of varying classical concepts within variable presumed uncertainties (‘coarse graining’) at selected discrete times. However, environmental decoherence allows one to *derive* quasi-classical concepts in terms of *wave packets* (all that is needed) – close to what Schrödinger had originally in mind. Their apparent ensembles, formally described by the density matrix, obey master equations – such as (4.35), and *in this way consistently define* quasi-classical ‘histories’.

In contrast to Bohm’s or Everett’s theories, consistent histories usually presume an *absolute* quantum arrow (see Hartle 1998 and footnote 2 of Chap. 5) – just as collapse models do. Their selection by a formal ‘consistency’ requirement may indeed be inconsistent itself (Kent 1997), while master equations derived by means of decoherence are never assumed to hold exactly. Consistent histories have also been claimed to be equivalent to stochastic trajectories *of wave functions*, in the ‘quantum state diffusion model’ assumed to exist for *all* systems (Diósi et al. 1995, Brun 2000 – see Sect. 4.4). However, this would be in conflict with the empirically confirmed quantum nonlocality.

On the other hand, the concept of apparent ensembles of wave functions in decoherence theory, or the density matrix in general, are based on a probability interpretation (as explained in Sect. 4.2), while the question ‘probabilities for (or information about) *what?*’ has not yet been answered on a fundamental level. Many quantum phenomena seem to favor the answer: new wave functions (such as narrow wave packets). For example, a spot on the photographic plate (regarded as a measurement ‘pointer’) has to be described in terms of local molecular *wave functions* – not in terms of any classical variables. The strongest support yet for such an interpretation may come from an analysis of the decoherence of neuronal states in the brain (Tegmark 2000, see also Zeh 2000). These quasi-classical neuronal states may form the final link (or

the ultimate pointer basis) in a chain of observational interactions that are all describable in terms of a global wave function.

The superposition of such quasi-classical observer states, which would result like Schrödinger cats from the unitary dynamics, was evidently the reason for postulating a collapse of the wave function as a *real* dynamical process (von Neumann's 'first intervention').<sup>9</sup> It may be formulated in two steps (see Fig. 4.3):

$$\begin{array}{ccccccc}
 |\psi\rangle\langle\psi| = \underbrace{\sum_{mn} |\phi_m\rangle c_m c_n^* \langle\phi_n|}_{S=0} & \longrightarrow & \underbrace{\sum_n |\phi_n\rangle |c_n|^2 \langle\phi_n|}_{S \geq 0} & \longrightarrow & \underbrace{|\phi_{n_0}\rangle \langle\phi_{n_0}|}_{S=0} , & (4.54) \\
 \text{with} & & & & & \\
 & & S=0 & \longrightarrow & S \geq 0 & \longrightarrow & S=0 .
 \end{array}$$

They represent (1) the transition from a pure state into an ensemble (characterized by an increase in von Neumann's ensemble entropy), and (2) the selection of a specific state from this ensemble (thus lowering the ensemble entropy as depicted by the first step of Fig. 3.5). The first step can be represented by a master equation that describes a loss of information, while the complete stochastic process (4.54) corresponds to a quantum Langevin equation (a Langevin equation for wave functions – see Sect. 4.4). Since the complete process describes an individual physical evolution, it has to be used when calculating the changing *physical* entropy according to (3.58). The master equation describes the dynamics of an ensemble that represents entanglement *and* lacking information about the stochastically evolving wave function.

If macroscopic properties  $\alpha$ , say, are again regarded as 'always given' (as in Sect. 3.3.1), *physical entropy* can be characterized by a function  $S(\alpha) = k \ln N_\alpha$ , similar to the last term of (3.58), where  $N_\alpha$  is now the dimension of the subspace representing a fixed value (or small interval) of  $\alpha$ . The time dependence of this entropy,

$$S(t) = S(\alpha(t)) , \quad (4.55)$$

is then determined by the macroscopic dynamics  $\alpha(t)$ , which in general includes a succession of collapse 'events', each one a dynamical projection onto a subspace corresponding to definite macroscopic properties. This means that *the stochastic collapse is part of the macroscopic dynamics*  $\alpha(t)$ . It is dynamically objectivized by the process of decoherence that describes an *apparent* collapse. In contrast to classical Hamiltonian dynamics, which determines the time dependence of any macroscopic quantity,  $\alpha(t) := \alpha(p(t), q(t))$  – see Sect. 3.3.1, the Schrödinger equation does *not* determine  $\alpha(t)$ . Therefore, the

<sup>9</sup> Despite frequent claims to the contrary, the dynamical collapse was never part of the Copenhagen interpretation, although quantum jumps have traditionally been used as an argument against a wave function representing reality. The collapse is then claimed to represent 'just a normal increase of information', even though an ensemble representing such incomplete information is excluded.

dynamics (4.32) cannot describe the transformation of entropy into lacking information in accordance with the negentropy principle (3.62). In other words, not even *macroscopically* different states have to possess different predecessors ('sufficient reasons') in quantum theory. Such causal predecessors, if they existed, would have to be counted by the initial ensemble entropy as in Fig. 3.5.

Would the collapse, if *used* in this way as part of the dynamics of wave functions, now specify an arrow of time that could perhaps even be responsible for irreversible thermodynamics? The ensemble entropy (4.4) would increase only during the auxiliary first step of the collapse (describing ignorance about the outcome). For entangled systems, for example those occurring after an interaction of type (4.32), an individual ('real') collapse,

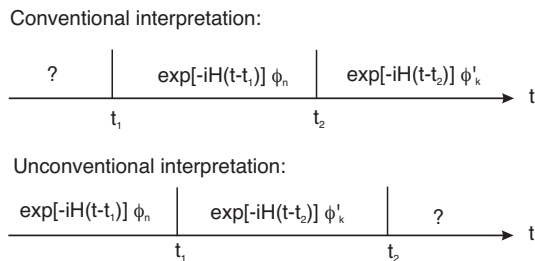
$$\psi = \sum_n c_n \phi_n \Phi_n \xrightarrow{t} \phi_{n_0} \Phi_{n_0} , \quad (4.56)$$

does *not* alter the ensemble entropy. However, it specifies an arrow as it transforms the entangled state into a factorizing one. Therefore, the additive ('physical') entropy *decreases* in this process after it may have correspondingly *increased* during the interaction (4.32), since

$$S[\hat{P}_{\text{sep}}|\psi\rangle\langle\psi|] \geq S[\hat{P}_{\text{sep}}|\phi_{n_0}\Phi_{n_0}\rangle\langle\phi_{n_0}\Phi_{n_0}|] = 0 . \quad (4.57)$$

A collapse has never been confirmed empirically as a dynamical process. It has nonetheless to be *taken into account* regardless of its interpretation before (or, at least, when) the observer becomes aware of the macroscopic pointer position. As he thereby becomes himself quantum correlated with the pointer state  $\Phi_{n_0}$ , the corresponding 'state-of-being-conscious' in his brain also becomes a pure state as a consequence (if not the origin) of this collapse. This 'observer state' (whatever it may be in detail) can thus be used for postulating a psycho-physical parallelism in accordance with von Neumann's intentions (see also London and Bauer 1939). There is thus no need for genuine classical (or other) variables *anywhere in-between* the observed microscopic system and the brain of the observer. Spontaneous localization of the pointer position in an apparatus would *not* lead to a pure state-of-being-conscious. Instead, one could regard the neuronal system as the true measurement device, while treating the whole outside world as one quantum system. Although this description of observations by conscious beings remains vague in detail, it is all that is *in principle* required for a physical formulation in terms of quantum states. While a 'real' collapse is usually assumed to occur as soon as certain phase relations have become irreversibly dislocalized, this assumption is merely convenient, as it comes close to reproducing a classical description.

Does the entropy-reducing collapse, wherever it may occur (or simply be *applied*), then have to be regarded as a quantum mechanical revival of Maxwell's demon, which could classically be exorcized in Sect. 3.3.2? Lubkin (1987) demonstrated that, just as in the classical case, this entropy decrease according to the collapse cannot be utilized in a cyclic process that would allow the construction of a perpetual mobile of the second kind. However, the



**Fig. 4.4.** Dynamics of the wave function in the case of retarded (conventional) and advanced (acausal) collapse. In contrast to classical waves, the choice of the usual ('retarded') interpretation of the collapse is a matter of pure convention

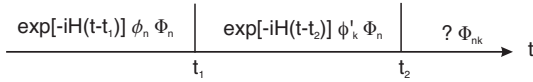
collapse may have important entropy-reducing consequences in a non-cyclic cosmic evolution (see Sect. 6.1).

It appears furthermore noteworthy that the usual form of the collapse is based on an extension of *intuitive causality* (Chap. 2) to a region where by assumption it cannot be confirmed. For example, the expression  $|\langle \phi'_k | \phi_n(\Delta t) \rangle|^2$ , with  $\phi_n(\Delta t) := \exp(-iH\Delta t)\phi_n$ , defines the probability of finding the state  $\phi'_k$  in an appropriate measurement at time  $t_2 = t_1 + \Delta t$ , provided the system was found in the state  $\phi_n$  in a previous measurement (of the first kind) at time  $t_1$ . This interpretation assumes the wave function to collapse into the state  $\phi_n$  during the first measurement, and then to evolve unitarily according to the Hamiltonian  $H$  until it is measured again. An equivalent (though unconventional) time-reversed interpretation may be obtained from the identity of the above matrix element with  $\langle \phi'_k(-\Delta t) | \phi_n \rangle$ . This means that the wave function may as well be assumed to collapse during the first measurement in an 'acausal' manner from an 'advanced' state  $\phi_n$  into  $\phi'_k(-\Delta t)$ , which will then unitarily evolve into the state  $\phi'_k = \phi'_k(0)$  just *before* the second measurement starts (Penrose 1979). These two versions of the collapse are indicated in Fig. 4.4. The second one is counterintuitive, since the observed system would have to 'know in advance' what kind of measurement will be performed, and when. Its exclusion is therefore an application of *intuitive causality*, similar to the exclusion of advanced fields, while the equivalence of the two descriptions is based on the T-symmetry of the formal quantum probabilities (Aharonov, Bergmann and Lebowitz 1964).

In contrast to the advanced electromagnetic fields (Sect. 2.4), which can be excluded empirically by means of small test charges, our preference for the causal version of the collapse is purely conventional. However, macroscopic registration devices, described by states  $\Phi$ , that are continuously monitored by the environment in accordance with the time arrow of increasing entanglement or quantum causality, have to be assumed to possess quasi-classical states at all times. In this case, their states during two successive measurements must be described by  $\Phi_0(t)$  for  $t < t_1$ , by  $\Phi_n(t)$  for  $t_1 < t < t_2$ , and by  $\Phi_{nk}(t)$  for  $t > t_2$ , where the number of indices corresponds to the increasing number of



Unconventional interpretation:



**Fig. 4.5.** Behavior of the total state (including that of the measurement device) under the unconventional assumption of advanced collapse and ‘continuous measurement’ of the pointer position, described by the state  $\Phi$ . Only the wave function of the microscopic system  $\phi$ , but not that of the macroscopic pointer, depends on the convention

registered (robust) measurement results (see also Bläsi and Hardy 1995). This holds even in the acausal collapse version – as indicated in Fig. 4.5.

All these ambiguities can be avoided if the Schrödinger equation is not modified at all. For a wave function that is assumed to describe reality completely, this leads to the Everett interpretation. There is then no law-like quantum arrow of time – but how can measurements *lead to* any definite and irreversible results?

Everett’s interpretation is based on the Schrödinger picture: “This paper proposes to regard pure wave mechanics as a complete theory” (Everett 1957, see also Zeh 1970, 1973). So *there are no longer any observables* to be introduced as further fundamental ingredients of the theory, as required in the Heisenberg picture. (The pragmatic equivalence of both pictures is in any case questionable for open, that is, locally non-Hamiltonian systems.) In particular, the ‘many-worlds interpretation’, which assumes many simultaneous histories formed by *classical* properties (DeWitt 1971, Deutsch 1997), or an ensemble of Feynman paths in classical configuration space (Sokolovsky 1998), misinterprets Everett’s proposal by presuming classical concepts. A *superposition* of Feynman paths would be identical to a time-dependent wave function as an individual dynamical state.

The ultimate observer states ( $\chi_n^{\text{obs}}$ , say), which differ in general in separate ‘branches’ of the wave function, need *not* necessarily represent quasi-classical states. Everett concluded from the Schrödinger equation that all components of (4.32) continue to exist in *one* superposition,  $\sum \psi^{(n)} \chi_n^{\text{obs}}$  (“All components are actual”). His point is that they can be *experienced* only separately because of their separate observer states  $\chi_n^{\text{obs}}$ , which are here *postulated* on empirical grounds to represent subjective awareness. Evidently, no local observer state is defined in the *global* Everett wave function. This interpretation answers von Neumann’s quest for a psycho-physical parallelism in quantum mechanical terms without introducing a collapse (see Zeh 1970, 1979, 2000, Squires 1990, Lockwood 1996).

Since configuration space assumes the role of space as the arena of reality in all versions of the Schrödinger picture, the *fork of indeterminism* that seems to characterize probabilistic quantum theory is reinterpreted in the Everett interpretation as a *fork of causality* (see footnote 1 of Chap. 2). All observer



states  $\chi_n^{\text{obs}}$  (with different  $n$ ) must be remembering the same *pre*-measurement history. To all of them, the rest of the world is thereafter described by their different ‘relative states’  $\psi^{(n)}$  (their co-factor states, which these observers may renormalize for convenience). The formal ‘plus’ of a superposition is objectively reduced to an effective ‘and’ by the irreversible dislocalization of superpositions, while the ‘or’ is observer-related (‘subjective’) – though objectivized with respect to correlated components of different observers.

Only if both factors of all components of the Everett wave function were defined to be mutually orthogonal would the sum of their products,  $\sum \psi^{(n)} \chi_n^{\text{obs}}$ , resemble the Schmidt representation (4.27) (Kübler and Zeh 1973, Zeh 1973, Albrecht 1992, 1993). This representation depends furthermore on the precise borderline between observer and the rest of the world. However, the essential (and objective) aspect of decoherence is the irreversible dynamical spreading of a superposition over *many* local subsystems  $\Phi^{(k)}$  (possibly including observers) in the form  $\sum_n c_n \phi_n \prod_k \Phi_n^{(k)}$  – thereby propagating in accordance with the relativistic spacetime structure.

As ‘the other’ robust components which must form according to the Schrödinger equation are not observable to ‘us’, the assumption of their existence is *operationally* meaningless (see Sect. 4.3.2 about the concept of ‘operational reality’). The Everett interpretation is therefore indistinguishable from an appropriately chosen collapse interpretation. In the Everett interpretation this equivalence requires the additional assumption that components only *branch* with increasing time, but in practice never (re)combine according to

$$\sum_n c_n \phi_n \Phi_n \xrightarrow{t} \left( \sum_n c_n \phi_n \right) \Phi_0 . \quad (4.58)$$

Because of the  $T$ -symmetric Schrödinger dynamics, the absence of this process requires quite generally that no suitable (conspiratorial) components  $n \neq n_0$  *exist* on the LHS according to the initial condition for the global wave function. This demonstrates also that the wave function cannot merely describe *possibilities* (in contrast to wave functions that *could* have arisen in a stochastic process, for example). The law-like arrow of the collapse is thus replaced by a fact-like arrow. This ‘quantum causality’ is formally analogous to Boltzmann’s *Stoßzahlansatz*, which requires that correlations are irrelevant in practice after being formed in collisions.

One may postulate an appropriate cosmic initial condition for this purpose by requiring that *all* existing nonlocal quantum correlations, such as those on the LHS of (4.58), are retarded, that is, have been *caused* somewhere during the past history of the quantum Universe. At the big bang ( $t = 0$ , say) one would thus have to assume a completely uncorrelated state, that is, a wave function of a form like

$$\psi(t \rightarrow +0) = \lim_{\Delta V_k \rightarrow 0} \prod_k \psi_{\Delta V_k}^0 , \quad (4.59)$$

where  $\psi_{\Delta V_k}$  are local states on appropriate small volume elements  $\Delta V_k$ . This kind of state is invariant under the corresponding Zwanzig projection  $\hat{P}_{\text{local}}$  by definition. However, the assumption (4.59) appears ‘natural’ only to our causal prejudice (thus applying the *double standard* in Price’s terminology). In contrast to its classical counterpart of initially absent (or future-irrelevant) correlations, this condition is not of a *statistical* nature (meaningful only for an ensemble that describes incomplete information), but a condition on the objective state of the quantum Universe. It would explain both the absence of quantum recoherence *and* the applicability of thermodynamical master equations.

Even though the Everett branches are sufficiently defined by means of decoherence (that is, by the dislocalization of superpositions – Sect. 4.3), this does not yet define any probabilities for them, as they are required by the Born rules. The branches are *all* assumed to exist once (with different albeit as yet physically meaningless norms). Everett’s original claim that the probability interpretation is a consequence of the formalism was based on the density matrix, and so must be circular (see Sect. 4.1.2). Graham (1970) therefore considered *series* of  $N$  equivalent measurements ( $N$  subsequent branchings – similar to series of decay events discussed at the end of Sect. 4.5). He was then able to demonstrate that the total norm of all those of the resulting Everett branches which represent series of outcomes that differ significantly from the Born rule must become extremely small, and vanish in the limit  $N \rightarrow \infty$  (see also Jammer 1974). While this result permits an elegant formulation of the probability postulate (by assuming merely that we happen to live in an Everett branch of not extremely small norm), this assumption is still completely equivalent to what is to be derived.<sup>10</sup>

Therefore, the probability measure has to be regarded as an *empirical* input to the theory (just like the Schrödinger dynamics, for example). The norm is a plausible candidate for this measure, because it is dynamically conserved under the Schrödinger equation. While stochastic collapse models postulate quantum probabilities as part of their dynamics, the Everett interpretation must *equivalently* assume that ‘we’ are living in a branch that has been selected by chance in accordance with the Hilbert space norm.

**General Literature:** Jammer 1974, Busch, Lahti and Mittelstaedt 1991, d’Espagnat 1995, Schloßhauer 2004, Zeh 2005.

<sup>10</sup> This may be illustrated by the example of results obtained for  $N$  subsequent measurements distinguishing between spin states  $|+\rangle$  and  $|-\rangle$  in  $N$  identical initial superpositions  $a|+\rangle + b|-\rangle$ . Since the *number* of branches which contain  $n$  ‘spin-ups’, say, is then statistically given by the binomial coefficient  $\binom{N}{n}$  regardless of the values of the coefficients  $a$  and  $b$ , the distribution of measurement outcomes  $n$  over many such series of  $N$  measurements,  $p_N(n)$ , would form a Poisson distribution centered at the required value if and only if the probability *for each branch* is assumed to be given by its norm  $|a|^{2n}|b|^{2(N-n)}$ . This is precisely Born’s probability rule. (Example provided by Erich Joos.)