

Foundation of Statistical Mechanics under Experimentally Realistic Conditions

Peter Reimann

Fakultät für Physik, Universität Bielefeld, 33615 Bielefeld, Germany

(Received 19 September 2008; published 7 November 2008)

We demonstrate the equilibration of isolated macroscopic quantum systems, prepared in nonequilibrium mixed states with a significant population of many energy levels, and observed by instruments with a reasonably bound working range compared to the resolution limit. Both properties are satisfied under many, if not all, experimentally realistic conditions. At equilibrium, the predictions and limitations of statistical mechanics are recovered.

DOI: 10.1103/PhysRevLett.101.190403

PACS numbers: 05.30.-d, 03.65.-w

Fundamental aspects of equilibrium statistical mechanics (ESM) are intensely reconsidered at present in the context of (almost) integrable many-body quantum systems [1–4], bringing back to our attention that very basic issues are still not satisfactorily understood [5–13]. As in every theory, we are faced with the three subproblems to realistically model preparation, time evolution, and measurement of a given system. It is well known and will be worked out in detail below that the question of experimentally realistic initial conditions and observables is much more urgent in the “derivation” of ESM than in most other fields [5,7–9,11,14]. Regarding time evolution, we take the widely (yet not unanimously) accepted viewpoint that standard quantum mechanics without any additional “postulate” or “hypothesis” must do [15,16]. The two key questions are then: How far does a nonequilibrium seed evolve to a stationary long-time behavior (“equilibration”)? How much is this steady state in agreement with the corresponding ESM ensemble (“thermalization”)?

Since open systems (interacting and entangled with an environment) are not directly tractable by standard quantum mechanics, the starting point must be a closed (autonomous) system (microcanonical framework), incorporating all relevant thermal baths, reservoirs, etc. [15,16]. Accordingly, the system “lives” in some Hilbert space \mathcal{H} and is at any time instant $t \geq 0$ in a mixed state (including pure states as special case) $\rho(t) = U_t \rho(0) U_t^\dagger$ with propagator $U_t := \exp[-iHt/\hbar]$, seed $\rho(0)$, and time-independent Hamiltonian H . Denoting its eigenfunctions and eigenvalues by $|n\rangle$ and E_n ($n = 0, 1, 2, \dots$) and the matrix elements $\langle m|\rho(t)|n\rangle$ by $\rho_{mn}(t)$ we thus obtain

$$\rho(t) = \sum \rho_{mn}(0) e^{i[E_n - E_m]t/\hbar} |m\rangle\langle n|, \quad (1)$$

where the sum runs over all $m, n \geq 0$. As usual, observables are represented by Hermitian operators A with expectation values $\text{Tr}\{\rho(t)A\}$ and, without loss of generality, are assumed not to depend explicitly on time.

Generically, the ensemble $\rho(t)$ is not stationary right from the beginning, in particular, for an initial condition $\rho(0)$ out of equilibrium. But if the right-hand side of (1)

depends on t initially, it cannot approach for large t any time-independent “equilibrium ensemble” whatsoever. In fact, any mixed state $\rho(t)$ returns arbitrarily “near” to its seed $\rho(0)$ for certain, sufficiently large time-points t , and similarly for the expectation values $\text{Tr}\{\rho(t)A\}$, see Appendix D in Ref. [17]. More specifically, consider any $\rho(0)$ with at least one $\rho_{mn}(0) \neq 0$ and $\omega := [E_n - E_m]/\hbar \neq 0$. Choosing

$$A = B + B^\dagger, \quad B := |m\rangle\langle n|/\rho_{mn}(0) \quad (2)$$

it follows that $\text{Tr}\{\rho(t)A\} = 2 \cos(\omega t)$. It is thus clearly impossible to “derive” (since it is not correct) ESM for arbitrary initial conditions and observables.

Our first basic assumption concerns the quantity [18]

$$R := \left[\sum \rho_{nn}^2(0) \right]^{1/3} \leq [\max_n \rho_{nn}(0)]^{1/3}. \quad (3)$$

According to (1), the $\rho_{nn}(t)$ represent the “occupation probabilities” of the energy eigenstates and are independent of t . For a system with f degrees of freedom, there are roughly $10^{O(f)}$ energy levels in every interval of 1 J beyond the ground state energy E_0 [15,19]. For a macroscopic system with $f = O(10^{23})$, the levels are thus unimaginably dense on any decent energy scale and even the most careful experimentalist will not be able to prepare the system such that the resulting ensemble $\rho(0)$ populates only a few energy eigenstates with significant probabilities [15]. For example, assume that there are exactly $10^{(10^{23})}$ energy levels per J. Even if the system preparation defines the energy up to an experimental uncertainty of $10^{-(10^{22})}$ J, there still remain $N := 10^{0.9 \times 10^{23}}$ energy levels which may be occupied with significant probabilities. If all of them are populated equally, we obtain $\rho_{nn}(0) = 1/N$ for N indices n and $\rho_{nn}(0) = 0$ for all other n , yielding $R \leq 10^{-0.3 \times 10^{23}}$ according to (3). If not all N levels are populated equally, but rather any $\rho_{nn}(0)$ may assume arbitrary values between zero and $10^{(10^{22})}$ times the average population $1/N$, Eq. (3) still yields $R \leq 10^{-0.26 \times 10^{23}}$. Returning to the general case, we can conclude that even if the system energy is fixed up to an extremely small experimental uncertainty and even if

the energy levels are populated extremely unequally, we still expect that R will be extremely small, typically

$$R = 10^{-\mathcal{O}(f)}. \quad (4)$$

Physical reasons for such a broadly spread energy level population include: the time-energy uncertainty relation, imperfect (as opposed to ideal) measurements during the system preparation ($t < 0$), entanglement processes and a time dependence of the Hamiltonian, both caused by that part of the environment from which the system is isolated for $t \geq 0$ but not for $t < 0$.

Given $\rho(0)$, let $\mathcal{H}_+ \subset \mathcal{H}$ be the Hilbert space spanned by those basis vectors $|n\rangle$ for which $\rho_{nn}(0) \neq 0$,

$$\mathcal{H}_+ := \text{span}\{|n\rangle | \rho_{nn}(0) > 0\}. \quad (5)$$

Exploiting Cauchy-Schwarz's inequality [20]

$$|\rho_{mn}(t)|^2 \leq \rho_{mm}(t)\rho_{nn}(t) \quad (6)$$

it follows that $\rho_{nm}(t) = 0$ whenever $\rho_{nn}(t) = 0$ or $\rho_{mm}(t) = 0$; i.e., "all nontrivial things are expected to happen within \mathcal{H}_+ ".

Our second basic assumption is that A represents an experimental device with a *finite range* of possible outcomes of a measurement within \mathcal{H}_+ [21],

$$\Delta_A := \max_{\mathcal{H}_+} \langle \psi | A | \psi \rangle - \min_{\mathcal{H}_+} \langle \psi | A | \psi \rangle = a_{\max} - a_{\min}, \quad (7)$$

where the maximization and minimization is over all normalized vectors $|\psi\rangle \in \mathcal{H}_+$ and where a_{\max} and a_{\min} are the largest and smallest eigenvalues of the restriction (or projection) of A onto \mathcal{H}_+ . Moreover we require that this working range Δ_A of the device A is limited to experimentally reasonable values compared to its resolution limit δA , for instance $\Delta_A < 10^{1000} \delta A$.

In the worst case, $\mathcal{H}_+ = \mathcal{H}$. However, in many cases the populations $\rho_{nn}(0)$ may be safely negligible, e.g., beyond some finite upper energy threshold, yielding a finite-dimensional \mathcal{H}_+ , while \mathcal{H} is typically infinite dimensional. Hence, (7) will be finite even for operators A with an unbound spectrum on \mathcal{H} . In any case, our above specified class of admissible observables A clearly includes any realistic measurement apparatus.

For the example (2) we can infer from (3), (6), and (7) that $\Delta_A \geq 2/R^{3/2}$. Hence, the oscillations $\text{Tr}\{\rho(t)A\} = 2\cos(\omega t)$ are below the resolution limit under our two basic assumptions. These (or similar) assumptions seem thus indispensable for taming the oscillations in (1).

Given any $\rho(t)$, we define the auxiliary operator [22]

$$\rho_{\text{eq}} := \sum \rho_{nn}(0) |n\rangle\langle n| \quad (8)$$

and focus on the mean square deviation

$$\sigma_A^2 := \overline{[\text{Tr}\{\rho(t)A\} - \text{Tr}\{\rho_{\text{eq}}A\}]^2}, \quad (9)$$

where the overbar indicates an average over all times

$t \geq 0$. The two trace-terms in (9) can be unified into $\text{Tr}\{\tilde{\rho}(t)A\}$ with $\tilde{\rho}(t) := \rho(t) - \rho_{\text{eq}}$. Introducing $\tilde{A} := A - \min_{\mathcal{H}_+} \langle \psi | A | \psi \rangle$ we can infer from (7) that

$$0 \leq \langle \psi | \tilde{A} | \psi \rangle \leq \Delta_A \quad \text{for all normalized } |\psi\rangle \in \mathcal{H}_+. \quad (10)$$

Since $\langle n | \tilde{\rho}(t) | n \rangle = 0$ it follows that $\text{Tr}\{\tilde{\rho}(t)\} = 0$ and that the variance (9) can be rewritten as $[\overline{\text{Tr}\{\tilde{\rho}(t)\tilde{A}\}}]^2$. With the help of (1) and $\rho_{nm} := \rho_{nm}(0)$ we finally obtain

$$\sigma_A^2 = \sum' \tilde{A}_{jk} \rho_{kj} \tilde{A}_{mn} \rho_{nm} \overline{e^{i[E_j - E_k + E_m - E_n]t/\hbar}}, \quad (11)$$

where the sum \sum' runs over all $j, k, m, n = 0, 1, 2, \dots$ with $j \neq k$ and $m \neq n$. Next we exploit the fact that $E_j - E_k + E_m - E_n$ vanishes for generic [16] Hamiltonians H only for $j = n$ and $k = m$, given $j \neq k$ and $m \neq n$ [23]. Since the time averaged exponentials in (11) vanish if $E_j - E_k + E_m - E_n \neq 0$ we can conclude that

$$\sigma_A^2 = \sum' |\tilde{A}_{mn}|^2 |\rho_{mn}|^2 \leq \sum |\tilde{A}_{mn}|^2 |\rho_{mn}|^2, \quad (12)$$

where the first sum runs over all $m \neq n$ and the second over all m, n . With (6) and (8) we thus obtain

$$\sigma_A^2 \leq \sum \tilde{A}_{mn} \rho_{nn} \tilde{A}_{nm} \rho_{mm} = \sum \langle m | \tilde{A} \rho_{\text{eq}} | n \rangle \langle n | \tilde{A} \rho_{\text{eq}} | m \rangle.$$

The sum over n amounts to an identity operator and that over m yields $\text{Tr}\{[\tilde{A} \rho_{\text{eq}}]^2\}$. This trace over the entire space \mathcal{H} can be restricted to \mathcal{H}_+ without changing its value, as follows from (5) and (8). Again, this trace remains unchanged if we now replace \tilde{A} by $\tilde{A}_p := P \tilde{A} P$, where P is the projector onto \mathcal{H}_+ . Next, we evaluate this trace with the help of the eigenvectors $|\chi_n\rangle$ of \tilde{A}_p (restricted to \mathcal{H}_+), yielding

$$\sigma_A^2 \leq \sum \langle \chi_m | \rho_{\text{eq}} \tilde{A}_p | \chi_n \rangle \langle \chi_n | \rho_{\text{eq}} \tilde{A}_p | \chi_m \rangle.$$

Observing that $\tilde{A}_p |\chi_n\rangle = |\chi_n\rangle \langle \chi_n | \tilde{A}_p | \chi_n \rangle$ (since $|\chi_n\rangle$ is eigenvector of \tilde{A}_p) and that $\langle \chi_n | \tilde{A}_p | \chi_n \rangle = \langle \chi_n | \tilde{A} | \chi_n \rangle$ (since $|\chi_n\rangle \in \mathcal{H}_+$ and thus $P|\chi_n\rangle = |\chi_n\rangle$) we can exploit (10) to obtain

$$\sigma_A^2 \leq \Delta_A^2 \sum \langle \chi_m | \rho_{\text{eq}} | \chi_n \rangle \langle \chi_n | \rho_{\text{eq}} | \chi_m \rangle.$$

The sum over n yields the identity operator (on \mathcal{H}_+) and that over m amounts to $\text{Tr}\{\rho_{\text{eq}}^2\}$. With (3) and (8) we finally arrive at

$$\sigma_A^2 \leq \Delta_A^2 R^3. \quad (13)$$

Considering $\text{Tr}\{\rho(t)A\}$ as a random variable with uniformly distributed $t \geq 0$, a similar (but simpler) calculation as before yields for its mean value the result $\overline{\text{Tr}\{\rho(t)A\}} = \text{Tr}\{\rho_{\text{eq}}A\}$. Hence, (9) is its variance and by combining (13)

with Chebyshev's inequality [24], we can conclude that

$$\text{Prob}(|\text{Tr}\{\rho(t)A\} - \text{Tr}\{\rho_{\text{eq}}A\}| \geq R\Delta_A) \leq R. \quad (14)$$

In view of (4) it follows that for the overwhelming majority of times $t \geq 0$ the difference between $\text{Tr}\{\rho(t)A\}$ and $\text{Tr}\{\rho_{\text{eq}}A\}$ is way below the instrumental resolution limit δA for any experimentally realistic observable, see below (7). In other words, the system looks exactly as if it were in the mixed state ρ_{eq} for the overwhelming majority of times $t \geq 0$ [22], though the “true” $\rho(t)$ is actually quite different, see above (2). This is our first main result.

Note that (14) is still compatible with the recurrence property of $\text{Tr}\{\rho(t)A\}$ mentioned above (2) but implies that such excursions from the “apparent equilibrium state” ρ_{eq} must be exceedingly rare events.

Exactly the same apparent equilibration towards ρ_{eq} emerges if one propagates $\rho(0)$ backward in time (keeping the system isolated). Along the entire real t -axis, an initial condition $\rho(0)$ far from equilibrium thus closely resembles one of the above mentioned rare excursions, just that the location of this excursion is on purpose chosen as the time-origin. In other words, the quantum mechanical time inversion invariance is maintained, but when starting out of equilibrium, an “apparent time arrow” emerges with extremely high fidelity.

While (14) provides a bound for the *relative* amount of time the system exhibits notable deviations from equilibrium, the typical duration of one given excursion, or equivalently, the characteristic relaxation time of an out of equilibrium initial condition $\rho(0)$ remains unspecified. Since one can easily find examples with arbitrarily large or small relaxation times, any further quantification of the relaxation process inevitably would require a considerably more detailed specification of the Hamiltonian H , the initial state $\rho(0)$, and the observable A .

Considering and estimating quantities like (9) is very natural and has a long tradition: Merits and shortcomings of the early works are reviewed, e.g., in [7], most notably Ludwig's approach [5]. In particular, many of them [6,7] involve an extra average over initial conditions with the effect that any specific nonequilibrium seed must be excluded as “potentially untypical” from the general conclusions. Turning to the more recent precursors, Peres' approach [8] is comparable to ours up to Eq. (12) but then proceeds with the conjecture that the \tilde{A}_{mn} are pseudorandom matrix elements, statistically independent of the ρ_{nm} , for which there are general arguments [8] and numerical evidence [25] (and counter-evidence [1]) but no proof. For *pure* states, Srednicki obtained similar results [11] by exploiting a common conjecture about the semiclassical behavior of classically smooth observables A in systems with a fully chaotic classical limit. Again, this conjecture is based on good arguments [26] but no proof. Moreover, typical classical many-body systems are not expected to behave fully chaotic [27]. Somewhat similar

conclusion have been reached even earlier by Deutsch [10] via additional hand waving arguments. Finally, rigorous results comparable to (14) are due to [4,12], but only for rather special Hamiltonians H and initial conditions.

According to (8) and the discussion below (14), expectation values become practically indistinguishable from

$$\text{Tr}\{\rho_{\text{eq}}A\} = \sum \rho_{nn}(0)A_{nn} \quad (15)$$

after initial transients have died out (equilibration). In how far is this in agreement with ESM, predicting [15] thermalization, i.e., the appearance of the microcanonical ensemble ρ^{mic} instead of ρ_{eq} in (15)? In case ρ^{mic} and ρ_{eq} yield measurable differences for experimentally realistic $\rho(0)$ and A , the “purely quantum mechanical” prediction (15) is commonly considered as “more fundamental” [1–3]. In other words, our derivation of ESM is complete, provided the latter is valid itself. In the opposite case, there is nothing to derive, but the quantity in (15) still governs the (time-) typical behavior. This the second main result of our Letter.

A first well-known validity condition for ESM is a “sharp energy E ”; i.e., all the $\rho_{nn}(0)$ with $E_n \in I := [E, E + \Delta E]$ sum up to almost unity, ΔE being small but still experimentally realistic [15]. In particular, the ρ_{nn}^{mic} are constant for all $E_n \in I$ and zero otherwise [15], and as a second (often tacit) validity condition for ESM, the resulting expectation values $\text{Tr}\{\rho^{\text{mic}}A\}$ are assumed to be (practically) independent of the exact choice of ΔE and E . Basically, this means that the details of $\rho_{nn}(0)$ do not matter in (15), henceforth called property (P). The same conclusion (P) follows from the equivalence of the microcanonical and canonical ensembles (for all energies E), considered as a self-consistency condition for ESM. Clearly, property (P) is tantamount to replacing ρ_{eq} in (15) by ρ^{mic} . Our first remark regarding (P) itself, is that no experimentalist can control the populations $\rho_{nn}(0)$ of the unimaginably dense energy levels E_n , apart from the very gross fact that they are “mainly concentrated within I ”. If the details would matter, not only ESM would break down but also reproducing measurements, in particular, in different labs, would be largely impossible. Second, one can readily construct observables and initial conditions, being experimentally realistic according to our definitions but still violating (P). The fact that ESM is known to have an extremely wide experimental applicability implies that our so far notion of “experimentally realistic” is still too general. The simplest option seems to require/assume that the expectation values $A_{nn} = \langle n|A|n \rangle$ hardly vary within the energy interval I . This is similar in spirit to classical coarse graining, and, in fact, is part of the already mentioned common conjecture about the semiclassical behavior of fully chaotic classical systems [26]. Next, even if the A_{nn} notably vary, following Peres [8] the immense number of relevant summands in (15) may—for “typical” A and

$\rho(0)$ —lead to a kind of statistical averaging effect and thus a largely $\rho(0)$ -independent final result. All these conjectures about “truly realistic experimental conditions” become even more compelling by considering that, “canonically”, A only affects a small subsystem, weakly coupled to a “big” rest, which can be readily traced out in (15), with the effect of an extra averaging step [6,7,12,13]. Yet, the apparent universality of property (P) and its relation to “more basic” system properties like “ergodicity” and “(non-)integrability” are still not very well understood [1,7,8,26,28].

Numerically, the validity and limits of such conjectures and of ESM itself have been exemplified, e.g., in [1,3,9,25]. While the details are not yet settled, equilibration in agreement with (15) was seen in all cases. Also the numerical observation that already quite small particle numbers often work surprisingly well is explained by (14) in view of (4).

In the classical case, proving the counterpart of the relation $\overline{\text{Tr}\{\rho(t)A\}} = \text{Tr}\{\rho_{\text{eq}}A\}$ (see above Eq. (14)) is tantamount to the notorious ergodicity problem [7]. The next step, namely, evaluating the classical counterpart of (9) remains as an even more difficult open problem.

-
- [1] M. Rigol, V. Dunjko, and M. Olshanii, *Nature (London)* **452**, 854 (2008).
- [2] T. Kinoshita, T. Wenger, and D. S. Weiss, *Nature (London)* **440**, 900 (2006); S. Hofferberth *et al.*, *ibid.* **449**, 324 (2007).
- [3] M. Rigol, V. Dunjko, V. Yurovsky, and M. Olshanii, *Phys. Rev. Lett.* **98**, 050405 (2007); C. Kollath, A. M. Läuchli, and E. Altman, *ibid.* **98**, 180601 (2007); S. R. Manmana, S. Wessel, R. M. Noack, and A. Muramatsu, *ibid.* **98**, 210405 (2007).
- [4] M. Cramer, C. M. Dawson, J. Eisert, and T. J. Osborne, *Phys. Rev. Lett.* **100**, 030602 (2008); M. A. Cazalilla, *ibid.* **97**, 156403 (2006); T. Barthel and U. Schollwöck, *ibid.* **100**, 100601 (2008).
- [5] G. Ludwig, *Z. Phys.* **150**, 346 (1958); **152**, 98 (1958).
- [6] P. Bocchieri and A. Loinger, *Phys. Rev.* **114**, 948 (1959).
- [7] R. Jancel, *Foundations of Classical and Quantum Statistical Mechanics* (Pergamon, London, 1969); I. E. Farquhar, *Ergodic Theory in Statistical Mechanics* (Interscience, New York, 1964).
- [8] A. Peres, *Phys. Rev. A* **30**, 504 (1984).
- [9] R. V. Jensen and R. Shankar, *Phys. Rev. Lett.* **54**, 1879 (1985).
- [10] J. M. Deutsch, *Phys. Rev. A* **43**, 2046 (1991).
- [11] M. Srednicki, *J. Phys. A* **29**, L75 (1996); **32**, 1163 (1999).
- [12] H. Tasaki, *Phys. Rev. Lett.* **80**, 1373 (1998).
- [13] N. Dass, S. Rama, and B. Sathiapalan, *Int. J. Mod. Phys. A* **18**, 2947 (2003).
- [14] A. Y. Khinchin, *Mathematical Foundations of Quantum Statistics* (Graylock Press, New York, 1960), Chap. 3; E. P. Wigner, *Am. J. Phys.* **31**, 6 (1963); N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 1992), Chap. XVII.7; A. Sugita, *Nonlinear Phenom. Complex Syst.* **10**, 192 (2007); O. Penrose, *Foundations of Statistical Mechanics* (Pergamon, Oxford, 1970), Chap. 1; S. Popescu, A. J. Short, and A. Winter, *Nature Phys.* **2**, 754 (2006).
- [15] Landau and Lifshitz, *Statistical Physics* (Pergamon, Oxford, 1980); B. Diu, C. Guthmann, D. Lederer, and B. Roulet, *Elements de Physique Statistique* (Hermann, Paris, 1996).
- [16] Though real systems are never strictly closed, our present standpoint (also supported by numerical experience) implies that this should not be the pivotal point: Otherwise, standard quantum mechanics alone would not be sufficient to model the (entangled) time evolution, and, on the other hand, ESM would not be applicable to strictly closed systems. After including the most relevant perturbations “from outside” into the considered system, it must be possible to theoretically model it as strictly isolated from the rest of the world, with a well-defined Hamiltonian H , possibly not known in detail but with generic properties.
- [17] A. Hobson, *Concepts in Statistical Mechanics* (Gordon and Breach, New York, 1971).
- [18] The inequality in (3) follows from $\sum \rho_{nn}^2(t) \leq \sum \rho_{nn}(t) \max_m \rho_{mm}(t)$ and $\text{Tr}\{\rho(t)\} = \sum \rho_{nn}(t) = 1$.
- [19] This is a result of “elementary level counting” and as such not one of the unresolved fundamental issues of ESM.
- [20] Since $\rho(t)$ is non-negative and Hermitian, $(\psi, \phi) := \langle \psi | \rho(t) + \epsilon | \phi \rangle$ is a well-defined scalar product for any $\epsilon > 0$. Hence, Cauchy-Schwarz’s inequality applies and by letting $\epsilon \rightarrow 0$ one obtains (6).
- [21] P. Reimann, *J. Stat. Phys.* **132**, 921 (2008).
- [22] One readily sees that (8) is indeed a well-defined density operator and satisfies Liouville-von-Neumann’s equation.
- [23] See, e.g., [8,11,12], and, in particular, Sect. 3.2.1 in S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghi, *J. Stat. Phys.* **125**, 1197 (2006), and references therein.
- [24] $\text{Prob}(|x - \mu| > \kappa \sigma) \leq \kappa^{-2}$ for any $\kappa > 0$ and any random variable x with average μ and variance σ^2 [17]. With $\kappa := \Delta_A R / \sigma_A$ and (13) one recovers (14).
- [25] M. Feingold, N. Moiseyev, and A. Peres, *Phys. Rev. A* **30**, 509 (1984).
- [26] See Ref. [11], M. Feingold and A. Peres, *Phys. Rev. A* **34**, 591 (1986), and references therein.
- [27] T. Prosen, *Ann. Phys. (Leipzig)* **235**, 115 (1994).
- [28] S. Weigert, *Physica (Amsterdam)* **56D**, 107 (1992); B. Sutherland, *Beautiful Models* (World Scientific, Singapore, 2004), Chap. 2.1; A. Enciso and D. Peralta-Salas, *Theor. Math. Phys.* **148**, 1086 (2006).