Renormalization group and quantum information

2006 J. Phys. A: Math. Gen. 39 7993

(http://iopscience.iop.org/0305-4470/39/25/S13)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **39** (2006) 7993–8006 doi:10.1088/0305-4470/39/25/S13

Renormalization group and quantum information

José Gaite

Instituto de Matemáticas y Física Fundamental, CSIC, Serrano 113 bis, 28006 Madrid, Spain

Received 1 December 2005, in final form 23 March 2006
Published 7 June 2006
Online at stacks.iop.org/JPhysA/39/7993

Abstract

The renormalization group is a tool that allows one to obtain a reduced description of systems with many degrees of freedom while preserving the relevant features. In the case of quantum systems, in particular, one-dimensional systems defined on a chain, an optimal formulation is given by White's 'density matrix renormalization group'. This formulation can be shown to rely on concepts of the developing theory of quantum information. Furthermore, White's algorithm can be connected with a peculiar type of quantization, namely, *angular quantization*. This type of quantization arose in connection with quantum gravity problems, in particular, the *Unruh effect* in the problem of black-hole entropy and Hawking radiation. This connection highlights the importance of quantum system boundaries, regarding the concentration of quantum states on them, and helps us to understand the optimal nature of White's algorithm.

PACS numbers: 05.10.Cc, 03.67.−a

1. Introduction

The renormalization group arose in quantum field theory as a transformation of the coupling constant(s) equivalent to a resummation of perturbation theory. It was later generalized by Wilson to statistical systems as a transformation of the full probability distribution, which is defined by an unbounded set of parameters. One of the most interesting aspects of this general renormalization group is that it can be understood as a transformation that removes small-scale degrees of freedom but preserves the set of degrees of freedom relevant to describe overall features, such as they are needed in the description of phase transitions, for example. In that sense, a renormalization group transformation is not reversible, so we should speak of a *semigroup* rather than of a group.

Regarding quantum many-body systems, one is most interested in calculating relevant features of the ground state (and maybe some excited states). This is a problem suitable for a renormalization group treatment, and indeed related to the problem of calculating the properties of classical statistical systems. There are many particular renormalization group algorithms that remove small-scale degrees of freedom in various ways. In principle, any algorithm

that approaches a fixed point is valid. However, the rate of convergence can vary broadly. It has been a problem to find efficient algorithms. In this regard, a deeper understanding of the process of removal of small-scale information surely helps. Here we shall focus on White's 'density matrix renormalization group' [1, 2], which has proved to be very powerful and which is indeed based on a deep analysis of the process of removal of small-scale information.

The analysis of the process of removal of small-scale information is actually a part of the analysis of information processing and, therefore, belongs to information theory. Furthermore, an analogy can be established between renormalization group transformations and the *irreversible evolution* of statistical systems, as implied by Boltzmann's H-theorem [3]. Since the density matrix renormalization group is an optimal tool for quantum systems (albeit mainly one-dimensional chains), we can expect a strong link with the theory of quantum information. This theory has its origins in work done early in the past century, shortly after the discovery of quantum theory itself. However, it has only undergone a period of rapid development during the past years, in relation with the prospects of quantum computation. The semigroup character of the quantum renormalization group is best understood by using concepts of quantum information theory, as we shall do.

White's algorithm involves a doubling of the system at each renormalization group iteration. As we will see, this makes sense from the quantum information standpoint and, in addition, has an interesting interpretation, since there is a relation with angular quantization [4]. Indeed, White's algorithm can be connected with another way of solving quantum systems; namely, quantum chains can be solved by relating them with classical two-dimensional systems on a lattice and using the *corner transfer matrix* method. The continuum limit gives rise to a peculiar type of quantization, namely, *angular quantization*, valid for *relativistic* quantum field theories, but different from the standard canonical quantization. This type of quantization was introduced in connection with quantum gravity problems, namely, the problems of black-hole entropy and Hawking radiation [5]. Angular quantization yields the relevant states in the calculation of the density matrix, showing precisely how the full spectrum is truncated to remove small-scale degrees of freedom [4]. In particular, it renders transparent the importance of quantum system boundaries, where quantum states concentrate.

Some new developments in the area of renormalization group and quantum information have appeared recently [6].

2. The density matrix renormalization group

Strongly correlated electron systems have become an important subject in condensed matter physics. This has led to the development of suitable approximation methods for quantum lattice models. Chiefly among them is the renormalization group, which has the philosophy of truncating the multitude of states to the *relevant* ones to describe the physical properties of the system in certain domain. There are many ways to implement this idea (not all of which can be properly called renormalization group methods). The various formulations of the quantum renormalization group have different efficacy, depending on the particular model to which they are applied. Two successful classical approaches are Wilson's treatment of the Kondo impurity problem and Kadanoff's blocking technique. These methods belong to the the class of numerical renormalization groups in *real space* (as opposed to Fourier space). Wilson's treatment of the Kondo problem is accurate, but it relies on the special nature of the interaction in it. The blocking technique is universal for lattice models but rather inaccurate, as we analyse next.

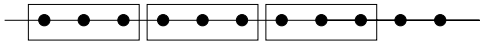


Figure 1. Blocking of a one-dimensional chain, with three sites per block.

2.1. Problems with the real space renormalization group

Consider a one-dimensional lattice model with N sites. If N is sufficiently large to represent a realistic system, the Hamiltonian involves a huge number of states, because the total number of states grows exponentially with N. There is no possibility of diagonalizing such a large Hamiltonian. Instead, we may break the chain into a number of equal size blocks, and we treat a block as a small size system (see figure 1). We can then diagonalize the block Hamiltonian and discard the higher energy states. Each block, with the states kept on it, can now be considered as a site of a new system with a fraction of the initial N sites. In so doing, we get an effective or *renormalized* interaction between sites. This procedure can be iterated until the initial size N is reduced to a small number.

However, the blocking RG converges slowly. White and Noack [7] realized that the problem lies in the choice of block eigenstates as the states to be kept: these states belong to a small system, namely, the block, in which the boundary conditions are very important. In other words, isolating a block from its neighbours destroys the quantum correlations between them, which are very important for the low-energy spectrum of the total system (with N sites). These correlations are somehow recovered by the renormalization of the couplings, but in a small amount. A partial remedy is the 'combination of boundary conditions' approach [7], namely, to consider block states corresponding to various boundary conditions. This approach is effective in some cases only.

So White and Noack [7] proposed to diagonalize a larger block, the 'superblock', which includes the basic block. Then the problem is to project the 'superblock' state (or states) onto block states: we need a criterium to select which ones to keep. White's intuition led him to appeal to Feynman's philosophy on the density matrix formalism: a density matrix simply represents the correlation of a quantum system with the rest of the universe. This correlation is usually called *entanglement*. The conclusion White drew is that the block states to be neglected, among the density matrix eigenstates, are those with small eigenvalues, because they hardly contribute to physical observables [1]. Let us recall White's procedure [1] more precisely.

2.2. Density matrix renormalization group algorithm

Let us have a one-dimensional quantum system on a chain (finite or infinite). We select a relatively large block (the 'superblock') but such that it can be exactly diagonalized. We obtain its ground state $|\psi\rangle$, which we take as *environment* of a smaller block included in the superblock. Let $|i\rangle, i = 1, \ldots, \ell$, be a complete set of block states and $|j\rangle, j = 1, \ldots, J$, be the states of the rest of the 'superblock' (Feynman's rest of the universe). Then $|\psi\rangle = \sum_{i,j} \psi_{ij}|i\rangle|j\rangle$. We want to find a subset of block states $|a\rangle, a = 1, \ldots, m < \ell$, such that they provide an optimal reduced representation of the block in the environment (boundary conditions) given by the superblock state. In other words, we want $|\tilde{\psi}\rangle = \sum_{a,j} \tilde{\psi}_{aj}|a\rangle|j\rangle$ to be as close to $|\psi\rangle$ as possible. White's prescription is to minimize the 'distance'

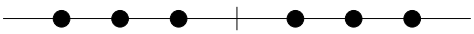
$$S = ||\psi\rangle - |\tilde{\psi}\rangle|^2. \tag{1}$$

Since both $|\psi\rangle$ and $|\tilde{\psi}\rangle$ are actually matrices, this distance is in fact the standard distance between matrices. He shows that this minimization problem amounts to the *singular value*

decomposition of the rectangular matrix ψ_{ij}. One writes $\psi = UDV^T$, where U and D are $\ell \times \ell$ matrices, V is an $\ell \times J$ matrix ($J \geqslant \ell$), U is orthogonal and V column orthogonal, and D is a diagonal matrix containing the singular values. The arbitrary integer $m < \ell$ defines the number $\ell - m$ of singular values to be neglected. Actually, U is the matrix formed by the eigenvectors of the block density matrix and $\rho = UD^2U^T$. Removing $\ell - m$ singular values is equivalent to keeping the m most important eigenvectors $|a\rangle$ of the block density matrix.

The construction of an iterative algorithm that implements the previous result is not difficult. A convenient algorithm [1], inspired by Wilson's treatment of the Kondo problem, can be schematically expressed as follows:

(1) Select a sufficiently small, soluble block $[0, L]$: 

(2) Reflect the block on the origin: 

(3) Compute the ground state.
(4) Compute the density matrix of the block $[0, L]$.
(5) Discard eigenstates with smallest eigenvalues.
(6) Add one site next to the origin.
(7) Go to 2.

(This algorithm is to be iterated indefinitely and represents the 'infinite-system algorithm', but there is also a 'finite-system algorithm'. We refer the reader to White's papers [1] for more details.) One has to adjust this procedure in such a way that the iteration keeps the Hilbert space size approximately constant. The procedure can be performed algebraically for a chain of coupled *harmonic* oscillators [4]. Otherwise, it has to be performed numerically.

2.2.1. Density matrix renormalization group for mixed states. We have assumed so far that the system is in a pure state (the ground state), which we want to calculate. One can also consider mixed states [1]. In particular, it is useful to consider the properties of thermal states. If we represent the mixed state by means of a set of Boltzmann weights w_k, then we have to minimize

$$S = \sum_k w_k ||\psi^k\rangle - |\tilde{\psi}^k\rangle|^2, \tag{2}$$

where ${\tilde{\psi}^k}_{ij} = \sum_{\alpha,j} a^k_\alpha u^\alpha_i v^{k,\alpha}_j$. The optimal solution is again to neglect the smallest (most singular) eigenvalues of the density matrix

$$\rho_{ii'} = \sum_k w_k \sum_j \psi^k_{ij} \psi^k_{i'j},$$

namely, the smallest values of $\sum_k w_k |a^k_\alpha|^2$. The same type of iterative algorithm can still be used.

3. Entanglement entropy and quantum information

Entanglement or non-separability refers to the existence of quantum correlations between two sets of degrees of freedom of a physical system that can be considered as subsystems [8, 9]. Two (sub)systems in interaction are entangled and their entanglement continues after their interaction has ceased. This fact gives rise to the Einstein–Podolsky–Rosen (EPR) paradox. So, while there can be entanglement without interaction, interaction always produces entanglement.

It is clear that entanglement plays a role in the density matrix renormalization group and, in general, in quantum phase transitions. This has been realized recently, by researchers in quantum information theory [10, 11]. The subject linking quantum information theory to traditional problems in condensed matter theory surely deserves further study. Here we review the concept of entanglement and other relevant concepts of information theory, regarding their role in the density matrix renormalization group.

3.1. Entanglement entropy of a bipartite system

The entanglement of two parts of a quantum system can be measured by the von Neumann entropy. This entropy is defined in terms of the density matrices of each part. We may consider, for later convenience, one part as 'left' and another as 'right' or one part as 'exterior' and another as 'interior'. Then, let us represent states belonging to the left part with small letters and states belonging to the right part with capital letters. A basis for the global states (left plus right) is $\{|a\rangle\}\otimes\{|A\rangle\}$. Let us take a global state, say the ground state for definiteness, and represent it in this basis as

$$|0\rangle = \sum_{aA} \psi_{aA}|a\rangle \otimes |A\rangle, \tag{3}$$

defining a coefficient matrix ψ_{aA}. Then we have two *different* density matrices, for each part:

$$\rho_L = \frac{\psi^* \psi^T}{\mathrm{Tr}\,\psi^* \psi^T}, \quad \rho_R = \frac{\psi^\dagger \psi}{\mathrm{Tr}\,\psi^\dagger \psi}. \tag{4}$$

Correspondingly, we have two von Neumann entropies:

$$S_L = -\mathrm{Tr}_a(\rho_L \ln \rho_L), \quad S_R = -\mathrm{Tr}_A(\rho_R \ln \rho_R). \tag{5}$$

Now it is important to recall the 'symmetry theorem', which states that both entropies are equal, $S_L = S_R$. This can be proved in several ways; for example, by using the *Schmidt decomposition* of the entangled state: both ρ_L and ρ_R have the same non-zero eigenvalues [9]. Let us remark that the Schmidt decomposition embodies entanglement and, actually, the singular value decomposition (as used by White) is its finite-dimensional version. The equality of entropies may seem somewhat paradoxical, since there can be many more degrees of freedom in one part (the exterior or rest of the universe) than in the other. But the entanglement entropies are associated with properties shared by both parts, that is, with (quantum) correlations.

Let us see how interaction produces entanglement and increases the entropy. Consider two non-interacting parts of a quantum system that are originally in respective mixed states. After their interaction, which we describe as an arbitrary unitary evolution of the composite system, the initial density matrix $\rho_L \otimes \rho_R$ has evolved to ρ'_{LR}. As a consequence of *subadditivity* of the entropy, it is easy to see that the partial traces ρ'_L and ρ'_R have in general von Neumann entropies S'_L and S'_R such that $S'_L + S'_R \geqslant S_L + S_R$ [8, 9]. Of course, if the initial state is a product of pure states, $S_L = S_R = 0$. In essence, this increase of entropy after interaction is an abstract form of the second law of thermodynamics.

3.2. Information theory and maximum entropy principle

The entropy concept arose in thermodynamics but only took a truly fundamental meaning with the advent of information theory. In this theory, entropy is just uncertainty or missing information, while information itself is often called *negentropy*. We recall basic definitions: the information attached to an event that occurs with probability p_n is $I_n = -\log_2 p_n$ (measured

in *bits*); therefore, the average information (per event) of a source of events is

$$S(\{p_n\}) = \sum_n p_n I_n = -\sum_n p_n \log_2 p_n.$$

This average information is called the entropy of the source. Note that improbable events convey more information but contribute less to the entropy: $pI(p) = -p\log_2 p$ is concave and has its maximum at $p = e^{-1}$.

The previous definitions, given by Shannon in his theory of communication, may seem unrelated to thermodynamic entropy as a property of a physical system. However, according to the foundations of statistical mechanics on probability theory (the Gibbs concept of ensembles), a clear relation can be established. This was done by Jaynes [12], by appealing to the Bayesian philosophy of probability theory. In this philosophy, the concept of *a priori* knowledge is crucial. Indeed, although the exact microscopic state of a system with many degrees of freedom may be unknown, one has some *a priori* knowledge given by the known macroscopic variables. Jaynes postulates, according to Bayesian philosophy, that the best probability distribution to be attributed to a stochastic event is such that it incorporates only the *a priori* knowledge about the event and nothing else. This postulate amounts to Jaynes' *maximum entropy* principle: given some constraints, one must find the maximum entropy probability distribution (density matrix, in the quantum case) compatible with those constraints, usually, by implementing them via Lagrange multipliers. In particular, more constraints mean less missing information and so less entropy.

3.3. Information geometry

Distinguishability of probability distributions is an important concept in information theory. The question is when two probability distributions are sufficiently distinguishable for some purpose and what measures are necessary to distinguish them. This in an important problem in statistics, in particular, in estimation theory. It has led to endow spaces of probability distributions with a metric geometry.

Let us briefly review the fundamental concepts of information geometry [13]. Let $p(x,\xi)$ be an n-parameter family of probability distributions ($\xi \in \mathbb{R}^n$). The primordial concept is the existence of a metric, namely, the *Fisher information matrix*:

$$g_{ij}(\xi) = 4\int \partial_i\sqrt{p(x,\xi)}\,\partial_j\sqrt{p(x,\xi)}\,\mathrm{d}x \tag{6}$$

(the derivatives are taken with respect to the parameters). This metric provides any space of probability distributions with a Riemannian structure. Hence, one can introduce the α*-connections*, the case $\alpha = 0$ being the standard Riemannian connection with respect to the Fisher metric. The next important concept is the notion of *divergence function*. It is a real positive function of a pair of probability distributions that vanishes if both distributions in the pair coincide. So divergences are distance-like measures, but they do not satisfy in general the remaining axioms of distance, in particular, they are not in general symmetric. However, a divergence's differential form is in fact symmetric and constitutes a metric. Most important are the α-divergences, given (in the discrete case) by

$$D^{(\alpha)}(p_i, q_j) = \sum_i p_i f(q_i/p_i), \tag{7}$$

$$f(x) = \begin{cases} \frac{4}{1-\alpha^2}(1 - x^{(1+\alpha)/2}) & \alpha \neq \pm 1 \\ -\ln x & \alpha = -1 \\ x\ln x & \alpha = 1. \end{cases} \tag{8}$$

They are related to generalized entropies: Rényi's α-entropies, Tsallis' entropy, etc.

In general, opposite sign divergences satisfy

$$D^{(-\alpha)}(p_i, q_j) = D^{(\alpha)}(q_i, p_j), \tag{9}$$

and are called *dual*. In particular, the case $\alpha = 0$ is symmetric and actually provides a distance, namely, $\sqrt{D^{(0)}(p_i, q_j)}$, with

$$D^{(0)}(p_i, q_j) = 2\sum_i (\sqrt{p_i} - \sqrt{q_i})^2. \tag{10}$$

The ± 1-divergence is called the Kullback–Leibler divergence or *relative entropy* and is particularly important. Its differential form yields the Fisher metric in the continuous case.

Note that the distance defined by equation (10) has a simple interpretation. To see it, let us associate with a probability distribution the vector $\{\sqrt{p_i}\}$, so that the probability normalization becomes a vector normalization. Therefore, probability distributions become rays in a real vector space [14]. The distance $\sqrt{D^{(0)}(p_i, q_j)}$ is just the standard distance in this vector space, or rather the induced distance in the corresponding projective space. In the continuous case, the standard Euclidean metric induces a metric in the parameter manifold that is precisely the Fisher metric (6).

The above-defined concepts are classical but they all admit quantum generalizations [13]. The square root of the $\alpha = 0$ quantum divergence coincides with the *Bures distance* between density matrices. This distance, restricted to the subspace of diagonal density matrices (for a fixed basis), does indeed become simply the square root of $2\,\mathrm{Tr}\big(\rho_1^{1/2} - \rho_2^{1/2}\big)^2$ [15]. In general, we can write $\rho_1 = W_1 W_1^\dagger$, $\rho_2 = W_2 W_2^\dagger$ for different pairs of operators $\{W_1, W_2\}$, and the Bures distance can be defined by the infimum

$$D^{(0)}(\rho_1, \rho_2) = 2\inf \mathrm{Tr}(W_1 - W_2)(W_1 - W_2)^\dagger. \tag{11}$$

For pure states, the Bures distance is just the natural distance in the complex projective Hilbert space, such that its infinitesimal form is the Fubini–Study metric [16–18]. Notably, the distance between mixed states (density matrices) is given by minimizing the distance between their respective *purifications* in a larger Hilbert space [19].

3.4. Quantum information theory

The concepts of Shannon's classical theory of communication have quantum analogues [9, 20]. But the quantum theory of communication is richer. Indeed, the key new notion in the quantum theory is entanglement (as already described). Schumacher studied the problem of quantum coding and, in particular, the problem of communication of an entangled state [20]. The technical name is *transposition*, since the copy of a quantum state is not possible, a fact that constitutes the no-cloning theorem [8]. His conclusion was that the von Neumann entropy of the state is the quantity that determines the *fidelity* of the transposition: it is possible to transpose the state with near-perfect fidelity if the signal can carry at least that information. The method is analogous to classical coding, that is, one has to discard small probabilities, but involves the use of the Schmidt decomposition.

Of course, fidelity and distinguishability are related concepts, and indeed the fidelity $F(\rho_1, \rho_2) = (1 - D^{(0)}(\rho_1, \rho_2)/4)^2$ [19]. Maximal fidelity ($F = 1$) is equivalent to perfect indistinguishability. Minimal fidelity ($F = 0$) takes place between maximally separated mixed states, since the Bures distance is bounded (this is obvious for pure states).

3.5. Quantum information interpretation of the density matrix renormalization group

Schumacher's approximate transposition of an entangled state is essentially identical to White's procedure. Moreover, White's distance criteria can be interpreted in terms of the distances defined in information geometry.

Let us recall White's prescription: select the block states to be kept by minimizing a 'distance' $S = ||\psi\rangle - |\tilde{\psi}\rangle|^2$ between the actual superblock state $|\psi\rangle$ and its approximation $|\tilde{\psi}\rangle$ (equation (1)). In terms of quantum information theory, we want to maximize the fidelity of the block mixed state or, in other words, to minimize the distance between the actual and approximated block mixed states. According to Jozsa's result [19], this minimization can be achieved by minimizing the distance between their respective *purifications* in a larger Hilbert space, which in this case is just the superblock. So it is correct to minimize the Hilbert space distance S.

The density matrix renormalization group for mixed states also corresponds to a distance minimization, in the space of superblock mixed states. The reference state is $\sum_k w_k|\psi^k\rangle\langle\psi^k|$ and its approximation is $\sum_k w_k|\tilde{\psi}^k\rangle\langle\tilde{\psi}^k|$. Then, according to equation (11),

$$D^{(0)}(\rho_1, \rho_2) = 2\sum_k w_k||\psi^k\rangle - |\tilde{\psi}^k\rangle|^2.$$

4. White's algorithm and angular quantization

White's density matrix renormalization group algorithm, exposed in section 2.2, can be purely justified on a quantum information basis as follows. If ρ_1 and ρ_2 are two mixed states of a Hilbert space $\mathcal{H}$, then $\mathcal{H}\otimes\mathcal{H}$ is the smallest Hilbert space that contains purifications of both states [19]. Therefore, for a block of given size, the most economical 'rest of the universe' is a reflection of the block (with the same size).

However, the particular geometry in White's algorithm lends itself to a more fruitful connection, namely, the connection with angular quantization [4]. Before explaining angular quantization, we must introduce the corner transfer matrix, a method of solving two-dimensional classical systems that turns out to be related to White's algorithm.

4.1. Corner transfer matrix and density matrix

Let us first recall the connection between quantum mechanics and classical statistical mechanics in one more dimension, realized by the Euclidean path integral. For spin chains, the equivalent classical system is defined on a two-dimensional lattice and the partition function can be conveniently expressed in terms of the *transfer matrix*. This matrix evolves the system from one row to the next one. In addition to the row-to-row transfer matrix, there was defined the *corner transfer matrix* (in the context of soluble models). This matrix evolves the system from one side of the corner to the other side. The formulation by Baxter of the corner transfer matrix for soluble models is old, but its importance in our context was realized later, in a paper by Thacker [21], in which he showed that the relevant symmetry is best understood in the continuum limit, as we shall see.

To introduce the corner transfer matrix, a site in the middle of the two-dimensional lattice is chosen as the origin, and then the spins (or other site variables) are fixed along the vertical and horizontal axes. Four different corner transfer matrices, say A, B, C, D, are defined by summing over the remaining site variables in each quadrant. Then the partition function is $Z = \mathrm{Tr}(ABCD)$. The matrix $ABCD$ represents the transfer from one side of the right

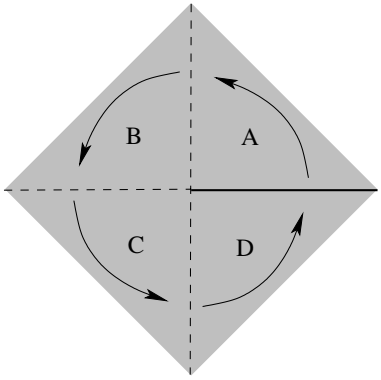


Figure 2. Composition of four corner transfer matrices in the direction of the arrows.

horizontal semiaxis to the other side, as shown in figure 2. If we define the state on the horizontal axis by the vector $\Psi(\boldsymbol{\sigma}_L, \boldsymbol{\sigma}_R)$ (splitting it on both semiaxes), we have

$$ABCD = \sum_{\sigma_L} \Psi^*(\boldsymbol{\sigma}_L, \boldsymbol{\sigma}_R)\Psi(\boldsymbol{\sigma}_L, \boldsymbol{\sigma}'_R).$$

Of course, this matrix (with entries $\boldsymbol{\sigma}_R$ and $\boldsymbol{\sigma}'_R$) corresponds to the density matrix of the right horizontal semiaxis in the environment provided by the left horizontal semiaxis. This connection was realized by Nishino and Okunishi [22] and developed by Peschel and collaborators [23].

On an isotropic lattice, the four corner transfer matrices can be arranged to coincide, so

$$\rho_R = A^4 = \exp(-H_{\mathrm{CTM}}),$$

defining a sort of corner-transfer-matrix Hamiltonian, such that $Z = \mathrm{Tr}\exp(-H_{\mathrm{CTM}})$ [23]. Roughly speaking, this Hamiltonian adopts the form $H_{\mathrm{CTM}} = \sum_n n H_n$, where H_n is a quasi-local Hamiltonian and the index n runs over sites. In comparison with the standard Hamiltonian for the row-to-row transfer matrix, we remark that the low-energy contribution of local states is depressed as we move away from the origin, due to the factor n (and vice versa). To substantiate this intuitive picture, we need to explain angular quantization in the continuum limit (field theory).

4.2. Field theory half-space density matrix

Let us consider, for definiteness, a chain of coupled oscillators. In the continuum limit, that is, for large correlation length, the model becomes simpler, in spite of having a non-denumerable set of degrees of freedom. The action for this model, namely, a one-dimensional scalar field, is (after a few redefinitions)

$$A[\varphi(x,t)] = \int \mathrm{d}t\,\mathrm{d}x \left(\frac{1}{2}[(\partial_t\varphi)^2 - (\partial_x\varphi)^2] - V(\varphi)\right), \tag{12}$$

where φ is the field. So the chain is described by a *relativistic* 1+1 field theory (relativistic with respect to the sound speed, normalized to one).

Let us obtain a path integral representation for the density matrix on the half-line with respect to the ground state (the vacuum) of action (12) [24–26]. In the continuum limit, the half-line density matrix is a functional integral,

$$\rho[\varphi_R(x), \varphi'_R(x)] = \int D\varphi_L(x)\psi_0[\varphi_L(x), \varphi_R(x)]\psi_0^*[\varphi_L(x), \varphi'_R(x)], \tag{13}$$

where the subscripts refer to the left or right position of the coordinates with respect to the boundary (the origin). Now, we must express the ground-state wavefunction as a path integral,

$$\psi_0[\varphi_L(x), \varphi_R(x)] = \int D\varphi(x,t)\exp(-A[\varphi(x,t)]), \tag{14}$$

where $t \in (-\infty, 0]$ and with boundary conditions $\varphi(x,0) = \varphi_L(x)$ if $x < 0$, and $\varphi(x,0) = \varphi_R(x)$ if $x > 0$. The conjugate wavefunction is given by the same path integral and boundary conditions but with $t \in [0, \infty)$. Substituting into equation (13) and performing the integral over $\varphi_L(x)$, one can express $\rho(\varphi_R, \varphi'_R)$ as a path integral over $\varphi(x,t)$, with $t \in (-\infty, \infty)$, and boundary conditions $\varphi_R(x, 0+) = \varphi'_R(x)$, $\varphi_R(x, 0-) = \varphi_R(x)$. In other words, $\rho(\varphi_R, \varphi'_R)$ is represented by a single path integral covering the entire plane with a cut along the positive semiaxis, where the boundary conditions are imposed.

Next, we need to calculate the density matrix, which we can do by diagonalizing it in the appropriate basis.

4.3. Angular quantization and Rindler space

Two-dimensional relativistic field theory has Lorentz symmetry, which becomes just rotational symmetry in its Euclidean version. The generator of rotations in the (x,t) plane is given by

$$\mathcal{L} = \int \mathrm{d}x(xT_{00} - tT_{11}), \tag{15}$$

in terms of the components of the stress tensor computed from the action (12). Of course, T_{00} is the Hamiltonian density and T_{11} is the momentum density. To simplify, one can evaluate $\mathcal{L}$ at $t = 0$, obtaining a Hamiltonian that we recognize as the continuum limit of H_{CTM}, defined in section 4.1.

For quantization, let us consider a free action [$V(\varphi) = 0$]. In the Schrödinger representation, we should replace the momentum $\Pi = \partial_t\varphi$ with $\Pi(x) = \mathrm{i}\delta/\delta\varphi(x)$. However, as in canonical quantization, one rather uses the second-quantization method, which diagonalizes the Hamiltonian by solving the classical equations of motion and quantizing the corresponding normal modes. Let us recall that, in canonical quantization, if we disregard anharmonic terms, the classical equations of motion in the continuum limit become the Klein–Gordon field equation, giving rise to the usual Fock space. In an angular analogy, the eigenvalue equation for $\mathcal{L}$ leads to the Klein–Gordon equation in polar coordinates in the (x,t) plane. The free field equation in polar coordinates,

$$(\Delta + m^2)\varphi = \left(\frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\phi^2} + m^2\right)\varphi = 0, \tag{16}$$

can be solved by separating the angular variable: it becomes a Bessel differential equation in the r coordinate with complex solutions $I_{\pm\mathrm{i}\ell}(mr)$, ℓ being the angular frequency. We have a continuous spectrum, which becomes discrete on introducing boundary conditions. One of them must be set at a short distance from the origin, to act as an ultraviolet regulator [24–26], necessary in the continuum limit.

Therefore, the second-quantized field is (on the positive semiaxis $t = 0 \Leftrightarrow \phi = 0$, $x \equiv r$)

$$\varphi(x) = \int \frac{\mathrm{d}\ell}{2\pi}\frac{b_\ell I_{\mathrm{i}\ell}(mx) + b_\ell^\dagger I_{-\mathrm{i}\ell}(mx)}{\sqrt{2\sinh(\pi\ell)}}, \tag{17}$$

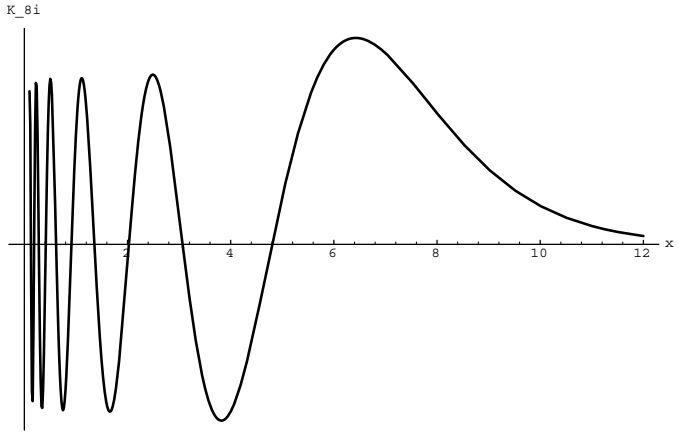


Figure 3. Angular-quantization wave $K_{8i}(x)$. Note the behaviour near $x = 0$.

where we have introduced annihilation and creation operators and where the term that appears in the denominator is just for normalization, to ensure that those operators satisfy canonical conmutation relations. There is an associated Fock space built by acting with $b_\ell^\dagger$ on the 'vacuum state'. These states constitute the spectrum of eigenstates of $\mathcal{L}$, which adopts the form $\mathcal{L} = \int \mathrm{d}\ell\ell b_\ell^\dagger b_\ell$ (where the integral is replaced with a sum for discrete ℓ). They are the density matrix eigenstates as well.

Let us remark that the functions $I_{\pm \mathrm{i}\ell}(mx)$ have wavelengths that increase with x. It is illustrative to represent a real 'angular-quantization wave',

$$K_{\mathrm{i}\ell}(mx) = \frac{\mathrm{i}\pi}{2\sinh(\pi\ell)}[I_{\mathrm{i}\ell}(mx) - I_{-\mathrm{i}\ell}(mx)].$$

This solution is oscillatory for $x < \ell/m$, with a wavelength proportional to x, and decays exponentially for $x > \ell/m$ (figure 3). Actually, for $x \ll \ell/m$, wavefunctions behave like $x^{\pm \mathrm{i}}$, that is, like trigonometric functions of $\ln x$.

This type of quantization was first introduced in the context of quantization in curved space, in particular, in Rindler space [5]. Rindler space is just Minkowski space and, therefore, *not curved*, but in coordinates such that the time is the proper time of a set of *accelerated* observers. Its remarkable feature is the appearance of an *event horizon*, which implies that the ground state (the Minkowski vacuum) is a mixed (thermal) state (the Unruh effect). The fact that wavelengths vanish at $x = 0$ is to be expected from the Rindler space viewpoint, because the origin corresponds to the horizon: the quantum states concentrate on it. The connection with black-hole entropy and Hawking radiation is very briefly explained in the next section.

4.4. Black-hole entropy

The motivation to study accelerated observers was, of course, the problem of black-hole entropy and Hawking radiation. This radiation is perceived by static observers but not by inertial (free-falling) observers. In fact, what a static observer close to the horizon can see is well described by the Rindler geometry. In other words, the large black-hole mass M limit of the Schwarzschild geometry is the Rindler geometry. To realize this limit, it is convenient to use the Kruskal–Szekeres coordinates u, v, instead of the Schwarzschild coordinates t, r (r is the radial distance which together with time are the only relevant variables of the Schwarzschild

geometry in any dimension) [5]. For small values of these coordinates (equivalently, $M \to \infty$), the curvature can be neglected and the geometry becomes locally the Rindler geometry.

Once established that the geometry near the black-hole horizon is locally the Rindler geometry of the preceding section, we can readily transfer the form of the density matrix of a scalar field therein, where we now ignore (trace over) the degrees of freedom inside the horizon. Hence, we can define a von Neumann entropy associated with this density matrix. Furthermore, in so doing, we can appreciate that the concept of black-hole entropy takes a new meaning: in addition to being of *quantum origin*, this entropy is related to shared properties between the interior and exterior, namely, to the horizon. Moreover, the radial vacuum is a thermal state with respect to the original Schwarzschild coordinates, giving rise to Hawking radiation [5].

4.5. Geometric entropy

We have seen that the half-line density matrix of a field theory has a geometric interpretation in Rindler space. Furthermore, the entropy of black holes can be understood as a generalization to a more complicated (curved) geometry. Since the important feature is just the existence of a horizon, we may wonder if further generalization is possible.

Indeed, the notion of 'geometric entropy' has been introduced by C Callan and F Wilczek [26], as the entropy 'associated with a pure state and a geometrical region by forming the pure state density matrix, tracing over the field variables inside the region to create an 'impure' density matrix'. They computed the Rindler space case (like Bombelli *et al* [24]) and further proposed a generalization to different *topologies*.

A different notion of geometric entropy can be deduced by purely geometrical means from the presence of horizons, namely, as associated with a spacetime topology that does not admit a trivial *Hamiltonian foliation* [27]. This type of topology prevents unitary evolution and produces mixed states. In fact, it is only this second type of entropy that leads to the famous 'one-quarter area law' for black holes, due to its origin in purely gravitational concepts. On the contrary, the first notion of geometric entropy needs an auxiliary field theory, involves UV divergences and needs renormalization before a comparison with the gravitational notion can be made.

5. Conclusions

We have seen that a density matrix renormalization group transformation amounts to a Hilbert space reduction that essentially preserves the information, that is, preserves the entropy. In this sense, it can be understood as a quantum coding operation, namely, a quantum data compression, which is not lossless but nearly so. So a density matrix renormalization group transformation is analogous to standard compression of classical data, such as it is routinely used in everyday data processing. Indeed, the singular value decomposition is used for classical data compression when data can be arranged in matrix form. Therefore, it is natural that it can also be applied to quantum data compression.

The density matrix renormalization group's ability to keep a constant and small Hilbert space size while the system size grows is crucial for its approaching a fixed point that represents the infinite size system. This limit is necessary to study quantum phase transitions, for example. The promising interface between quantum phase transitions and information theory is just beginning to be studied [2, 11]. Furthermore, in the limit of large correlation length, the relevant dynamics, as given by the ground state and the lowest excited states, can be described by a relativistic quantum field theory (in which the mass of the excitations decreases with the

correlation length). Since relativistic quantum field theories in 1+1 dimensions can be treated with powerful mathematical methods, we can expect that they are a suitable ground to explore the connection of quantum dynamics with information theory. We remark, of course, that one can stop the renormalization group iteration at any desired point, when some predetermined size is reached (the 'finite system method' [1]).

The density matrix renormalization group ability to keep a constant and small Hilbert space size relies on having a distribution of density matrix eigenvalues in which most of them are actually negligible. In fact, their typical distribution decays exponentially. White proposed the analogy with an ordinary statistical system with the canonical distribution [1] (which was the original motivation of Feynman's density matrix philosophy). We have seen that it is more than a mere analogy: White's algorithm is equivalent to the calculation of the density matrix in angular quantization. The connection with the Unruh effect reveals that one can indeed associate a particular thermodynamical picture and a temperature with angular quantization. This picture can be generalized in terms of the concept of geometric entropy and, in fact, connects with the notion of *holography* in quantum gravity [28].

Finally, regarding angular quantization and its associated distribution of quantum states, let us remark how it helps us to understand the efficiency of White's algorithm: it is very efficient because it employs the smallest number of boundaries allowed, namely, just one boundary, unlike other renormalization group formulations. For example, the block partitioning technique produces a very large number of boundaries.

Acknowledgments

My work is supported by MEC 'Ramón y Cajal' Program and grant BFM2002-01014.

References

[1] White S R 1992 *Phys. Rev. Lett.* **69** 2863
White S R 1993 *Phys. Rev.* B **48** 10345
[2] Schollwöck U 2005 *Rev. Mod. Phys.* **77** 259–315
[3] Gaite J and O'Connor D 1996 *Phys. Rev.* D **54** 5163–73
Gaite J 1998 *Phys. Rev. Lett.* **81** 3587
[4] Gaite J 2001 *Mod. Phys. Lett.* A **16** 1109
[5] Birrell N D and Davies P C W 1982 *Quantum Fields in Curved Space* (Cambridge: Cambridge University Press)
[6] Plenio M B, Eisert J, Dreissig J and Cramer M 2005 *Phys. Rev. Lett.* **94** 060503
Verstraete F, Cirac J I, Latorre J I, Rico E and Wolf M M 2005 *Phys. Rev. Lett.* **94** 140601
Wolf M M 2005 *Phys. Rev. Lett.* **96** 010404
Verstraete F and Cirac J I 2005 *Preprint* cond-mat/0505140
[7] White S R and Noack R M 1992 *Phys. Rev. Lett.* **68** 3487
[8] Peres A 1995 *Quantum Theory: Concepts and Methods* (Dordrecht: Kluwer)
[9] Preskill J 1998 *Physics (Lecture Notes* vol 229*)* (on line at http://www.theory.caltech.edu/people/preskill/ph229/)
[10] Osborne T J and Nielsen M A 2002 *Quant. Inf. Proc.* **1** 45
[11] Vidal G, Latorre J I and Rico E 2003 *Phys. Rev. Lett.* **90** 227902
Latorre J I, Rico E and Vidal G 2004 *Quant. Inf. Comp.* **4** 48–92
[12] Jaynes E T 1983 *Papers on Probability Theory, Statistics and Statistical Physics* (Dordrecht: Reidel)
[13] Amari Sh-i and Nagaoka H 2000 *Methods of Information Geometry* (Oxford: Oxford University Press)
[14] Brody D C and Hughston L P 1996 *Phys. Rev. Lett.* **77** 2851
[15] Hübner M 1992 *Phys. Lett.* A **163** 239
[16] Wootters W K 1981 *Phys. Rev.* D **23** 357
[17] Gibbons G W 1992 *J. Geom. Phys.* **8** 147
[18] Braustein S L and Caves C M 1994 *Phys. Rev. Lett.* **72** 3439
[19] Jozsa R 1994 *J. Mod. Opt.* **41** 2315

[20] Schumacher B 1995 *Phys. Rev.* A **51** 2783
[21] Thacker H B 1986 *Physica* D **18** 348
[22] Nishino T and Okunishi K 1997 *Corner Transfer Matrix Algorithm for Classical Renormalization Group Preprint* cond-mat/9705072, and references therein
[23] Peschel I, Kaulke M and Legeza O 1999 *Ann. Phys.,Lpz.* **8** 153
[24] Bombelli L, Koul R K, Lee J and Sorkin R D 1986 *Phys. Rev.* D **34** 373–83
Srednicki M 1993 *Phys. Rev. Lett.* **71** 666
[25] Kabat D and Strassler M J 1994 *Phys. Lett.* B **329** 46–52
[26] Callan C and Wilczek F 1994 *Phys. Lett.* B **333** 55–61
[27] Hawking S W and Hunter C J 1999 *Phys. Rev.* D **59** 044025
[28] Bousso R 2002 *Rev. Mod. Phys.* **74** 825–74