Minireview

Cayley Trees and Bethe Lattices: A concise analysis for mathematicians and physicists

M. Ostilli *

Department of Computational & Theoretical Sciences, International Islamic University Malaysia, P.O. Box, 141, 25710, Kuantan, Pahang, Malaysia
Statistical Mechanics and Complexity Center (SMC), INFM-CNR SMC, Rome, Italy

ARTICLE INFO

ABSTRACT

We review critically the concepts and the applications of Cayley Trees and Bethe Lattices in statistical mechanics in a tentative effort to remove widespread misuse of these simple, but yet important – and different – ideal graphs. We illustrate, in particular, two rigorous techniques to deal with Bethe Lattices, based respectively on self-similarity and on the Kolmogorov consistency theorem, linking the latter with the Cavity and Belief Propagation methods, more known to the physics community.

© 2012 Elsevier B.V. All rights reserved.

Contents

1. Introduction

After many years since their introduction, Cayley Trees (CT) [1] and Bethe Lattices (BL) [2] still play an important role as prototypes of graphs. In fact, even if one can say, from the point of graph theory, that these ideal graphs[1] are by now obsolete objects which have been replaced by more realistic random graphs, like the classical random graph since a few decades ago [3],[2] and by complex networks more recently [4], their key feature, that is the fact they are exact tree, *i.e.*, cycles-free, makes CT and BL very instructive examples where exact calculations can be done [5]. However, even though there are

---

[1] We will see soon that, both CT's and BL's have little numerical availability: on one hand a CT is very sensitive to the boundary conditions while, on the other hand, a BL, due to the fact that is an infinite graph, cannot be simulated (represented) on a PC.

[2] Note that nowadays the name Bethe Lattice is often used to indicate the regular random graph, which is a tree-like graph, but not an exact tree (a tree-like graph is a graph containing only long cycles so that locally it looks like a tree [4]). However, in this paper we will reserve the name Bethe Lattice to indicate an exact (and then infinite) tree.

several excellent works on their applications, there is still a quite widespread confusion about their exact definition and use. In particular, by a rapid survey (September 2011), ranging from Wikipedia to famous textbooks in statistical mechanics, as well as many papers in referred journals of mathematics or physics, one finds statements claiming, for example, that the BL is the thermodynamic limit of the CT, or one finds that a BL is the interior of a large CT, that can be then analyzed by introducing large but finite subtrees, etc... Such false statements and misuses are not just formal mistakes, but serious conceptual misunderstandings that may lead in turn to fatal errors. The difference between a CT and a BL was emerged long ago in Refs. [6–10], but, nevertheless, confusion on the subject remained around over the years, both among mathematicians and physicists. We think that the main reason for that is due to an incorrect mathematical approach to the BL, and to scarce communication between the physics and mathematics communities. Moreover, at the time of the Refs. [6–10], a proper nomenclature for the two kinds of graphs was not yet consolidated, causing further confusion.

The aim of this paper is to give a concise definition of CT and BL and to illustrate ambiguities-free mathematical tools to be used for statistical mechanical models built over CT and BL. We will show two rigorous techniques to deal with BL: self-similarity and the Kolmogorov consistency theorem, pointing out that the latter is equivalent to the Cavity and Belief Propagation methods, more known among physicists.

2. Cayley trees and Bethe lattices: definition and basic properties

Both CT and BL are simple connected undirected graphs $G = (V, E)$ (V set of vertices, E set of edges) with no cycles (a cycle is a closed path of different edges), *i.e.*, they are trees.

A CT of order q with n shells is defined in the following way. Given a root vertex (0), we link (0) with q new vertices by means of q edges. This first set of q vertices constitutes the shell $n = 1$ of the CT. Then, to build the shell $l \geq 2$, each vertex of the shell $l - 1$ is linked to $q - 1$ new vertices. Note that the vertices in the last shell n have degree $q_n = 1$, while all the other vertices have degree $q_l = q, l = 0, \ldots, n - 1$ (the shell $l = 0$ is represented by the single root vertex (0)).

The BL of degree q is instead defined as a tree in which any vertex has degree q, so that there is no boundary and no central vertex and, as a consequence, the main difference between CT and BL is simply that CT is finite while BL is infinite:

$$\text{CT}: \quad |V|, |E| < \infty, \tag{1}$$

$$\text{BL}: \quad |V|, |E| = \infty. \tag{2}$$

Note that Eqs. (1) and (2) imply an important difference for the average connectivity c (the connectivity of a given vertex is defined as the number of edges emanating from it) between a CT and a BL. In fact, for any finite tree, and in particular a CT, it holds $|V| = |E| + 1$. Therefore

$$\text{CT}: \quad c = 2 - \frac{2}{|V|}, \tag{3}$$

$$\text{BL}: \quad c = q, \tag{4}$$

where in deriving c for the CT case we have used $c = 2|E|/|V|$ (valid for any graph). In Sections 3 and 4 we will see that the difference between Eqs. (3) and (4) has a dramatic consequence in statistical mechanics.

In probability theory, an important distinction is in order between a sequence of probability spaces of increasing size that eventually diverges, and a probability space that is infinite by definition. Similarly, in statistical mechanics, one can be more interested in studying the thermodynamic limit of the density free energy of the system, or else in studying the physical properties of a system which is defined from the very beginning as an infinite space (physical or abstract). Despite only the former kind of infinity leads to a constructive theory of statistical mechanics and seems physically relevant (in the real world nothing is really infinite, not even the universe), the second kind of infinity may still be, not only mathematically convenient, but even physically important (we will see this soon). From Eqs. (1) and (2), we see that, when we study a model of statistical mechanics built on a CT, we can have access only to the thermodynamic limit of the system, while when we study a model on a BL, by definition, we have access only to the physical properties of the model as defined over an infinite space. Even though most of the models in physics show equivalence between the two different "kinds of infinity", in the case of CT and BL such equivalence is lost. The reason for such a difference is easily seen even without entering into the details of a specific model. In fact, when we study a model on the CT we need to specify the boundary conditions of the model, while on the BL, by definition, there are no boundary conditions. Given a CT of degree $q > 2$ with n shells, the number of vertices on the l-th shell is given by $N_l = q(q-1)^{l-1}$, while the total number of vertices of the CT is $N = q((q-1)^n - 1)/(q-2)$. We see therefore that the ratio N_n/N of the CT, in the thermodynamic limit, $N \to \infty$, does not approach zero (for $q > 2$). This situation is very different from what happens in a d-dimensional regular lattice box, where the ratio of the number of boundary vertices (vertices on the $d - 1$ dimensional surface), with respect to the total number of vertices, for $N \to \infty$ reaches zero as fast as $1/N^{1/d}$. Therefore, while the thermodynamic limit of a model built on increasing boxes subsets of Z^d, is equivalent to the physical properties of the model defined on the infinite space Z^d, the boundary conditions becoming negligible, the thermodynamic limit of a model built on increasing CT's is not equivalent to the physical properties of the model defined on the infinite space BL where, by definition, there are no boundaries. No matter how large N is, the model built on the CT will heavily depend on the boundary conditions, a feature that makes the model on the CT rarely representative of a real world physical system, so that the model on the BL is often preferred. In the next sections the non equivalence between CT and BL will be made concrete with the example of the Ising model.

3. Ising model on Cayley trees and on general trees

Given a CT (V, E) of degree $q > 2$ and n shells, we want to analyze the Ising model built on it, having Hamiltonian

$$H[\{\sigma_i\}] \overset{def}{=} -J \sum_{(i,j)\in E} \sigma_i \sigma_j - H_0 \sum_i \sigma_i, \tag{5}$$

and partition function

$$Z_n \overset{def}{=} \sum_{\{\sigma_i\}} e^{-\beta H[\{\sigma_i\}]}, \tag{6}$$

where $\beta = 1/T$ is the inverse temperature, J the coupling constant, H_0 the external field, the σ's $= \pm 1$s are the spin variables, and Z_l stands for the partition function of a CT with l shells. Let us consider free boundary conditions, and, for simplicity, the case $H_0 = 0$. Since the CT is finite, we can in particular start to perform the summation of (6) by summing over the boundary spins, *i.e.*, by summing over the spins on the n-th shell. By using $\cosh(\beta J \sigma) = \cosh(\beta J)$ we get the recursive equation

$$Z_n = Z_{n-1}[\cosh(\beta J)]^{N_n}, \tag{7}$$

from which, by iterating, we arrive at

$$Z_n = [\cosh(\beta J)]^{N_n + N_{n-1} + \cdots + N_1}, \tag{8}$$

that is

$$Z_n = [\cosh(\beta J)]^{|V|}. \tag{9}$$

From Eq. (9) we see that the density free energy f

$$-\beta f \overset{def}{=} \lim_{|V|\to\infty} \frac{\log(Z_n)}{|V|} = \log[\cosh(\beta J)] \tag{10}$$

is an analytic function of β for any $\beta < \infty$, therefore the Ising model on the CT does not give rise to a spontaneous magnetization. Eq. (10) has actually a more general counterpart, as it holds for any tree, *i.e.*, with no cycles. This can be easily seen, for example, by using the high temperature expansion of the free energy [11]. By using the high temperature expansion, for any graph $G = (V; E)$, f can be written as

$$-\beta f = \lim_{|V|\to\infty} \frac{|E|}{|V|} \log[\cosh(\beta J)] + \varphi_G, \tag{11}$$

where φ_G is the non trivial part of the expansion and is defined as the sum over all the closed non overlapping paths C of G of $\tanh(\beta J)^{l(C)}$, where $l(C)$ is the length of the closed path C. Now, if G is a tree, there are no closed paths, therefore $\varphi_G = 0$, and from Eq. (1) it follows

$$-\beta f = \log[\cosh(\beta J)] \lim_{|V|\to\infty} \frac{|E|}{|V|}. \tag{12}$$

The last factor $|E|/|V|$ is nothing other than $c/2$, c being the average connectivity of G. By using Eq. (3) we see that the free energy of any tree does not depend on the details of the graph. In conclusion, the thermodynamic limit of the Ising model built on trees does not have a spontaneous magnetization. However, it can be shown [7,8,10] that the free energy density f of a CT with $q > 2$ is a non analytic function of the external field h for $T \leq T_{BL}$, where T_{BL} will be introduced in the next section.

4. Ising model on the Bethe lattice

We now illustrate how to solve an Ising model on the BL by using two methods which are free of ambiguities.

4.1. Self-similarity method

Let us consider for simplicity a BL lattice G with degree $q = 3$ with the exception of one single vertex (0) that instead has degree 2 (see Fig. 1). We introduce this little irregularity on our BL just to make things simpler (later on we will show how to restore the fully regular BL). From the root (0) emanate 2 edges pointing at the vertices 1 and 2. Since G is infinite, it turns out that the two equivalent infinite subgraphs G_1 and G_2 that we obtain by eliminating from G the vertex (0) together with its two edges, are also both equivalent to the original G, *i.e.*, up to a change of name of the vertices, we have the self-similarity
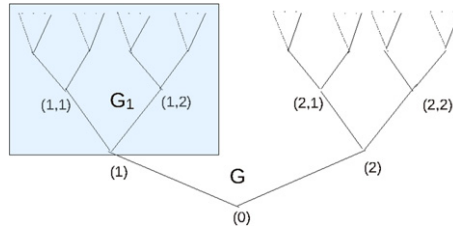
Fig. 1. Example of (a portion of) an irregular Bethe Lattice G of degree $q = 3$ with the root vertex (0) having only 2 first neighbors. The numbers (1) and (2) in the figure represent the vertices in the shell $n = 1$, the numbers (1,1), (1,2), (2,1), and (2,2), the vertices on the second shell, and so on. Up to a change of names of the vertices, one has $G = G_1$, where G_1 is the (infinite) subgraph in the shaded area.

$G = G_1 = G_2$. Note, furthermore, that G_1 and G_2 are each other disconnected. If we define the conditional partition function of the system $Z(\sigma_0)$ with respect to the root (0) as

$$Z(\sigma_0) \stackrel{def}{=} \sum_{\{\sigma_i, i \neq 0\}} e^{-\beta H[\{\sigma_i\}]}, \tag{13}$$

where H is defined similarly to Eq. (5), by using the self-similarity and the fact that G_1 and G_2 are disconnected, we get (again for simplicity we consider here only the case with no external field $H_0 = 0$)

$$Z(\sigma_0) = \sum_{\sigma_1, \sigma_2} e^{\beta J \sigma_0 (\sigma_1 + \sigma_2)} Z(\sigma_1) Z(\sigma_2). \tag{14}$$

One can feel unease with Eqs. (13) and (14) due to the fact that G is infinite, as well as H, and the Z's. However we can get rid of the Z's and any ill defined quantity when we consider the finite ratio $P(-)/P(+) = Z(-)/Z(+)$, where $P(\sigma_0)$ stands for the probability that the spin at (0) has value σ_0. If we define

$$e^{-2h} \stackrel{def}{=} \frac{P(-)}{P(+)}, \tag{15}$$

from Eq. (14) we arrive at the equation for h

$$e^{-2h} = \frac{e^{2\beta J - 4h} + e^{-2\beta J} + 2e^{-2h}}{e^{2\beta J} + e^{-2\beta J - 4h} + 2e^{-2h}}. \tag{16}$$

Eq. (16) has always the trivial solution $h = 0$ corresponding to a zero magnetization. However, by expanding Eq. (16) for small h, it is easy to see that a non trivial solution $h \neq 0$ is present when $T < T_{BL}$, where T_{BL} is the critical temperature given as solution of the equation

$$(q - 1) \tanh(\beta J) = 1, \tag{17}$$

which has solution as soon as $q > 2$. In other words, the value $q = 2$ represents the critical number of neighbors above which in the system there exists a phase transition. We can now physically understand why, unlike the BL, in the CT we cannot never have a spontaneous magnetization. From Eqs. (3) and (4) we see that, unlike the BL, on average, no matter how large the degree q of the CT is, the connectivity c of the CT is always strictly below the value $c = 2$, therefore, in the CT, no matter how large its degree q, on average any spin is surrounded by an insufficient number of neighbors so that the system cannot have a phase transition (at zero external field H).[3] Similar conclusions apply to the Potts model [10]. The average magnetization of the spin σ_0 can be calculated from Eq. (15) and gives $\langle \sigma_0 \rangle = \tanh(h)$, while, the average magnetization for the fully regular BL with $q = 3$ is simply $\langle \sigma_0 \rangle = \langle \sigma \rangle = \tanh(h * 3/2)$. Notice however the necessity to consider the irregular BL: self-similarity, and then the possibility to have an equation for the effective field h, applies only to the irregular BL, not to the fully regular BL. The physicist familiar with the Cavity Method (see next subsection) recognizes in this step another way to see that h is the strength felt by a spin due to all the other spins in the absence ("the cavity") of the spin itself.

We end this subsection with the following critical note. In many textbooks of statistical mechanics, as well as in many papers that deal with a BL (regular or not), in the place of Eqs. (14) or (16), we find respectively

$$Z_{n+1}(\sigma_0) = \sum_{\sigma_1, \sigma_2} e^{\beta J \sigma_0 (\sigma_1 + \sigma_2)} Z_n(\sigma_1) Z_n(\sigma_2), \tag{18}$$

and

$$e^{-2h_{n+1}} = \frac{e^{2\beta J - 4h_n} + e^{-2\beta J} + 2e^{-2h_n}}{e^{2\beta J} + e^{-2\beta J - 4h_n} + 2e^{-2h_n}}, \tag{19}$$

---

[3] More precisely, since in the thermodynamic limit for the CT it holds $c = 2$, from Eq. (17) applied with $q = c = 2$, we see that the Ising model on the CT is critical only at $T = 0$.

and it is often wrongly said that Z_n stands for the partition function of that finite portion of the BL having only n shells. From what we have seen all above, it should be clear that such an approach for the BL is conceptually quite wrong: Z_n cannot be interpreted as the partition function of a finite subgraph of the BL, otherwise for Z_n we would get again Eq. (9). At this level, the only correct meaning we can attribute here to $Z_n(\sigma)$ is that of the conditional partition function of an infinite subgraph self-similar to that associated to $Z_{n+1}(\sigma)$, as shown above, and, most importantly, $Z_{n+1}(\sigma)$ and $Z_n(\sigma)$ are the same thing. In this context, the index n appearing in Eqs. (18) or (19) is a quite misleading symbol. More precisely, one should use instead $Z(\sigma_{n+1})$ and $Z(\sigma_n)$ but, as a consequence, for the ratios we would get the same value regardless of the index: $h_n = h_{n+1}$. Rather, looking at Eqs. (14) or (16), as a mere numerical tool, it can be convenient to introduce the recurrent Eqs. (18) or (19) since its fixed points are equivalent to Eqs. (14) or (16), respectively. A more rigorous and general meaning to h_n can be attributed by following the next measure-theoretical approach.

4.2. Kolmogorov's condition

A more rigorous approach to the BL comes from the Kolmogorov consistency theorem (see *e.g.*, Ref. [12,13]) by which we can avoid to introduce any ill-defined quantity but, most importantly, as we will see, the method attributes a rigorous and physical meaning to Eq. (19) and to its generalizations.

From the point of view of probability theory, solving the Ising model on the BL amounts to find all the marginal probabilities of an infinite probability space, *i.e.*, a probability space characterized by an infinite number of random variables (in our case the spins with indices in V), a suitable sigma-algebra [12], and a measure μ. The marginal probabilities can be calculated in turn from an infinite set of given finite dimensional marginal distributions $\mu_n = \mu_{|V_n}$, where V_n is the set of vertices that are at distance not greater than n from 0. Note however that, when we consider μ_n, we are not dealing with a finite disconnected subgraph of the given BL; the spins of V_n that are located on the shell n are in fact in turn connected to the spins of V_{n+1}. The key point to be used is that, given the values of the spins on $V \setminus V_n$, the spins on V_n are conditioned only by the neighboring spins, *i.e.*, the spins of the $n + 1$-th shell. In turn, given a spin σ_i, with $i \in V_n$, the effects of all its neighbor spins that are on $V \setminus V_n$, can be encoded via a single effective external field h_i acting only on σ_i. Therefore, for the μ_n, we look for solutions of the form

$$\mu_n(\{\sigma_i, i \in V_n\}) = e^{-\beta H_n + \sum_{i \in W_n} h_i \sigma_i} / Z_n \tag{20}$$

where H_n is given by Eq. (5), $W_n = V_n \setminus V_{n-1}$ stands for the set of vertices of the shell n, and Z_n is the normalization constant. Physically, Eq. (20) represents the first step of the Bethe–Peierls approach [2] (nowadays also known as the Cavity-Method in physics [16], or Belief-Propagation in computer science [18]) in which the effects of all the spins other than those of the set V_n are encoded in the fields $\{h_i\}$ to be determined self-consistently. From a strict probabilistic point of view, Eqs. (5) and (20) represent the definition of the Ising model itself on the BL. Now, as we have mentioned in Section 2, the distinction between a sequence of probability spaces of increasing size and an infinite probability space is important. The sequence of probability spaces may or may not converge to anything, and even if it converges, it may happen that it does not converge to a probability space. The tool to investigate whether a sequence of increasing probability spaces converges to an infinite probability space is provided by the Kolmogorov consistency theorem [12]. It should be clear that on an infinite space like the BL, the probability that a given configuration of spins $\{\sigma_i, i \in \mathbb{N}\}$ (a point-like event) is realized is zero: $\mu(\{\sigma_i, i \in \mathbb{N}\}) = 0$. Therefore, when one considers an infinite probability space, the introduction of a proper sigma-algebra is not just a formal necessity. However, since here we do not need to work directly with the measure μ of the infinite space, for what follows, we can skip the introduction of the sigma-algebras. The Kolmogorov consistency theorem says that a sequence of probability spaces (V_n, μ_n) converges to the probability space on BL, (V, μ), if for any $n \geq 1$ one has

$$\sum_{\{\sigma_i, i \in W_{n+1}\}} \mu_{n+1}(\{\sigma_i, i \in V_{n+1}\}) = \mu_n(\{\sigma_i, i \in V_n\}). \tag{21}$$

Eqs. (20) and (21) give the functional equations for the fields h_i, $i \in W_n$.

Due to the tree-like structure of the BL, Eq. (21) factorizes over independent branches allowing for a fundamental simplification. Let us apply Eqs. (20) and (21) to the same model analyzed in the previous subsection (a BL with $q = 3$ in which the root vertex (0) has only two first neighbors) [14,15]. For any given $i_0 \in W_n$ we indicate with i_1 and i_2 the indices of the first two neighbors of i_0 that belong to W_{n+1}. We have ($H_0 = 0$)

$$\frac{e^{-\beta H_n(\{\sigma_l, l \in V_n\})}}{Z_{n+1}} \prod_{i_0 \in W_n} \sum_{\sigma_{i_1}, \sigma_{i_2}} [e^{\beta J \sigma_{i_0}(\sigma_{i_1} + \sigma_{i_2}) + h_{i_1}\sigma_{i_1} + h_{i_2}\sigma_{i_2}}] = \frac{e^{-\beta H_n(\{\sigma_l, l \in V_n\})}}{Z_n} \prod_{i_0 \in W_n} e^{h_{i_0}\sigma_{i_0}}. \tag{22}$$

From Eq. (22), for any $i_0 \in W_n$ and for any values of σ_{i_0}, we have an equation for the field h_{i_0} as a function of the fields h_{i_1} and h_{i_2}

$$\frac{Z_n}{Z_{n+1}} \sum_{\sigma_{i_1}, \sigma_{i_2}} e^{\beta J \sigma_{i_0}(\sigma_{i_1} + \sigma_{i_2}) + h_{i_1}\sigma_{i_1} + h_{i_2}\sigma_{i_2}} = e^{h_{i_0}\sigma_{i_0}},$$

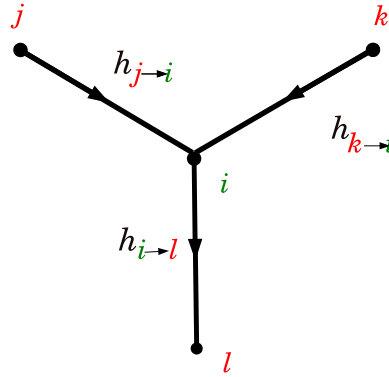**Fig. 2.** Graphical message-passing interpretation of the effective fields. Each field $h_{i \to l}$ is in correspondence with an edge $(i, l) = (l, i) \in E$ (recall that our BL G is an undirected graph). The field/message $h_{i \to l}$, emanating from i, as seen from Eqs. (24), is a function of all the other fields/messages arriving to i: $h_{j \to i}$ and $h_{k \to i}$. The correspondence between the fields appearing in Eqs. (23) and those of Eqs. (24) is: $h_{i \to l} = h_{i_0}$, $h_{j \to i} = h_{i_1}$, and $h_{k \to i} = h_{i_2}$. Note that, in general, we can attribute a direction to the fields/messages $h_{i \to l}$ even if the graph G is an undirected graph (*i.e.*, for any $(i, j) \in E$ we have $(i, j) = (j, i) \in E$).

from which, by evaluating the cases $\sigma_{i_0} = +1$ and $\sigma_{i_0} = -1$ we get

$$e^{-2h_{i_0}} = \frac{e^{2\beta J - h_{i_1} - h_{i_2}} + e^{-2\beta J + h_{i_1} + h_{i_2}} + e^{h_{i_1} - h_{i_2}} + e^{h_{i_2} - h_{i_1}}}{e^{2\beta J + h_{i_1} + h_{i_2}} + e^{-2\beta J - h_{i_1} - h_{i_2}} + e^{h_{i_1} - h_{i_2}} + e^{h_{i_2} - h_{i_1}}}. \tag{23}$$

It is immediate to verify that Eqs. (23) reduce to Eq. (19) for the choice $h_{i_0} = h_{i_1} = h_{i_2} = h_n$. Besides to be rigorous, this approach, based on the Kolmogorov consistency theorem, shows us that a possible non homogeneous solution for the effective fields $\{h_i\}$, periodic or not, is not an artifact of the theory: an effective field which depends on the index vertex is in correspondence with a non homogeneous marginal probability which depends on the index vertex too.

In the context of the Cavity and Belief Propagation methods, Eqs. (23) are better known in another form more suited for an intuitive interpretation of the effective fields as messages $h_{i \to j}$ passing from a vertex i to a neighbor vertex j; see Fig. 2. By using $\tanh^{-1}(x) = 1/2 \log[(1 + x)/(1 - x)]$, valid for $|x| < 1$, it is easy to see that Eqs. (23) can be rewritten as

$$h_{i \to l} = \tanh^{-1}[\tanh(\beta J) \tanh(h_{j \to i})] + \tanh^{-1}[\tanh(\beta J) \tanh(h_{k \to i})]. \tag{24}$$

More in general, the equation for a an Ising model built on a generic BL, regular or not, having generic couplings J_{ij} and an external field H_0, reads as

$$h_{i \to j} = \sum_{k \in \partial i \setminus j} \tanh^{-1}[\tanh(\beta J_{ij}) \tanh(\beta H_0 + h_{k \to i})], \tag{25}$$

where ∂i stands for the set of the first neighbors of the vertex i.

The developments about the convergence and algorithmic issues around Eqs. (25), as well as their applications in physics, computer science, and statistical inference are huge (see *e.g.* Refs. [16–18]). Here we wanted just to stress the equivalence between the Cavity/Belief-Propagation approaches and a rigorous method based on the Kolmogorov consistency theorem.

5. Last observation

From the previous section, we see that we can re-interpret the (infinite) BL, as the thermodynamic limit of (finite) CT's having specific "boundary conditions" determined by the fields h_i solution of Eqs. (20) and (21), as shown in [19]. However, this point of view is a quite poor physical one and, pedagogically, very inconvenient.

6. Conclusions

We have reviewed critically the concepts and the applications of CT and BL in statistical mechanics emphasizing their very different features. We have pointed out that careful attention must be paid especially in the case of a BL, an infinite space, where serious misuses and misunderstandings are currently seen in both textbooks and journal papers due to an incorrect mathematical approach to the BL. We have then illustrated for the BL case two alternative approaches which are rigorous and free of dangerous ambiguities based respectively on self-similarity and the Kolmogorov consistency theorem, pointing out the link of the latter with the Cavity and Belief Propagation methods, more known to the physics community. We hope that this critical review paper, where concepts and tools from physics and mathematics (too often kept apart) are used, might reduce the quite widespread misuse and conceptual errors around CT and BL.

Acknowledgments

References

[1] A. Cayley, Desiderata and suggestions: no. 2 the theory of groups: graphical representation, Amer. J. Math. (2) (1878) 174.
[2] H.A. Bethe, Statistical theory of superlattices, Proc. Roy. Soc. London Ser A 150 (1935) 552.
[3] P. Erdös, A. Rényi, Publ. Math. Debrecen 6 (1959) 290.
[4] R. Albert, A.L. Barbási, Rev. Mod. Phys. 74 (2002) 47;
S.N. Dorogovtsev, J.F.F. Mendes, Evolution of Networks, University Press, Oxford, 2003;
M.E.J. Newman, SIAM Review 45 (2003) 167;
S.N. Dorogovtsev, Lectures on complex networks, in: Oxford Master Series in Statistical, Computational, and Theoretical Physics, 2010;
S.N. Dorogovtsev, A.V. Goltsev, J.F.F. Mendes, Rev. Mod. Phys. 80 (2008) 1275.
[5] R.J. Baxter, Exact Solved Models in Statistical Mechanics, Academic Press, London, 1982.
[6] T.P. Eggarter, Phys. Rev. B 9 (1974) 2989.
[7] H. Matsuda, Progr. Theor. Phys. 51 (1973) 1053.
[8] E. Muller-Hartmann, J. Zittartz, Phys. Rev. Lett. 33 (1974) 893.
[9] F.Y. Wu, J. Phys. A: Math. Gen. 9 (1976) 593.
[10] L. Turban, Phys. Lett. 78A (1980) 404.
[11] C. Domb, in: C. Domb., M.S. Green (Eds.), Ising Model Phase Transitions and Critical Phenomena, vol. 3, Academic Press, London, 1974, p. 357.
[12] A.N. Shiryaev, Probability, Springer, New York, 1984.
[13] Y.G. Sinai, Theory of Phase Transitions: Rigorous Results, Pergamon Press, Oxford, New York, 1982.
[14] H.O. Georgii, Gibbs Measures and Phase Transitions, Walter de Gruyter, Berlin, 1988.
[15] P.M. Bleher, J. Ruiz, V.A. Zagrebnov, J. Stat. Phys. 79 (1995) 473.
[16] M. Mezard, G. Parisi, Eur. Phys. J. B 20 (2001) 217–233.
[17] R. Zecchina, O.C. Martin, R. Monasson, Statistical mechanics methods and phase transitions in optimization problems, Theoret. Comput. Sci. 3 (2001) 265.
[18] J.S. Yedidia, W.T. Freeman, Y. Weiss, Understanding belief propagation and its generalizations, in: Gerhard Lakemeyer, Bernhard Nebel (Eds.), Exploring Artificial Intelligence in the New Millennium, Morgan Kaufmann, 2003, p. 239.
[19] C.J. Thomposon, J. Stat. Mech. 27 (1982) 441.