

A Correlation Decay Theorem at High Temperature

M. Lehto^{1,*}, H. B. Nielsen², and Masao Ninomiya³

1 Nordita, Blegdamsvej 17, DK-2100, Copenhagen Ø, Denmark

2 The Niels Bohr Institute and Nordita, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark

3 Department of Physics, Brown University, Providence, RI 02912, USA

Abstract. We derive a theorem of exponential decay of correlation functions at high temperature for a general statistical mechanical system following the approach introduced by L. Gross. The theorem is formulated so as to be useful for locality problems in lattice quantum gravity.

1. Introduction

The purpose of the present article is to formulate a theorem of exponential decay of correlation functions at high temperature in a statistical mechanical system so as to be applicable to a lattice gravity model [1]. In fact we want to show that a lattice quantum gravity model based on Regge calculus [2] with an ultraviolet cutoff of the order of the Planck length leads to Einstein's gravity theory at large distances. The effective action derived from the lattice theory should be general coordinate invariant and local with respect to the gravitational field. The latter will be proved by the theorem of the present article in a subsequent paper [1].

There are several articles in literature discussing classical statistical mechanical systems which under certain conditions (e.g. high temperature or low activity) do not exhibit long range correlations. We refer for example to the papers of Gross [3], Israel [4], and Klein [5]. In our case we are mainly concerned with a rather generally formulated theorem introduced by Gross [3]. Gross used the Dobrushin uniqueness method [6] to derive exponential decay rates for two point correlation functions at high temperature in classical statistical mechanical lattice models. The result of Gross can be expressed in the form (Theorem 1 in [3]): If $\alpha < 1$, then for any functions $f, g \in C(\Omega)$,

$$|\sigma(fg) - \sigma(f)\sigma(g)| \leq e^{-d(a,b)} \|f\|_a \|g\|_b (1 - \alpha)^{-2} (1 - \alpha^2)^{-1}. \quad (1.1)$$

Ω is the space of all functions s from a countable set L ($a, b \in L$) into a compact metric space X ($s(a) \in X$ for $a \in L$). σ is a probability measure on Ω , $d(\cdot, \cdot)$ is a

* Address after 1.1.1984: University of Cambridge, Department of Applied Mathematics and Theoretical Physics, Silver Street, Cambridge CB3 9EW, England

pseudometric on L and the norm of f is given by

$$\|f\|_a = \sum_{b \in L} e^{d(a,b)} \sup \{|f(s) - f(t)| \mid s, t \in \Omega \text{ and } s = t \text{ off } \{b\}\}.$$

The parameter α is defined in terms of conditional probabilities $\mu_a(\cdot \mid \cdot)$, so that

$$\alpha = (1/2) \sup_{b \in L} \sum_{a \in L} [\sup \{\|\mu_b(\cdot \mid s) - \mu_b(\cdot \mid t)\|_{\text{var}} \mid s = t \text{ off } \{a\}\}] e^{d(a,b)}. \quad (1.2)$$

The main point of the present article is to slightly improve upon the Gross theorem (1.1) so that we can use it later on in applications (especially in lattice gravity).

It is the parameter α that provides a measure for how nonlocal the action of a lattice field theory is. The Gross theorem (1.1), which we shall refer as G. Th. hereafter, presupposes that α is less than unity. However, the problem with G. Th. is that the definition of α contains taking a supremum over pairs of microstates for the statistical mechanical system. This means that microstates that are extremely unlikely because of the weight factor e^{-S} (S is the action) being small may contribute significantly to α . Intuitively that seems unreasonable: one would think that rarely occurring states should not be of great importance for the range of correlations. Of course that is not quite true if the rare states by themselves have infinite range correlations. But if they have only finite, although somewhat long range, correlations we expect their contribution to the range of correlations for the total system to be controllable. It is basically this expectation we want to justify by formulating and proving an improved version of G. Th..

Before going into detailed definitions (Sect. 2) we first briefly explain some features of our way to improve G. Th..

We consider a general lattice field theory having field configurations $s : L \rightarrow X$ just like in G. Th., but we choose a pseudometric $d(\cdot, \cdot)$ which depends on the state of the statistical mechanical system. That is to say, for each state s there is defined a pseudometric $d_s(\cdot, \cdot)$ obeying as a function of the lattice points the usual rules for a pseudometric. Actually we are not so interested in the case in which there really is a new pseudometric for each state s . Rather we take a pseudometric which as a function of the states is a piecewise constant function. In other words there are “large” pieces γ of the space of states $\Omega (s \in \Omega, \gamma \subset \Omega)$ over which the pseudometric $d(\cdot, \cdot)$ does not vary. The point of this kind of an equivalence class division of Ω in the improved theorem is that if the classes γ are small but numerous our assumptions (see Sect. 2) will be very strong. Therefore one could expect that it will be easier to argue for the applicability of our theorem in the case where $d(\cdot, \cdot)$ is piecewise constant with quite large pieces. We mention that G. Th. itself may be considered as a special case corresponding to just one piece. Furthermore we notice that the set of all pseudometrics can be interpreted as a lattice in the mathematical sense of the word [7] and we use this property by introducing “the meet pseudometric” $d_s \wedge d_t(\cdot, \cdot)$ and “the join pseudometric” $d_s \vee d_t(\cdot, \cdot)$.

We also make another kind of an equivalence class division by dividing the lattice L , which is a countable set, into sets R by using the pseudometric nature of $d_s(\cdot, \cdot)$. Here the equivalence relation is determined by the pseudometric in such a way that R is given by those lattice points for which the pseudometric vanishes. So

effectively we are putting some points together and therefore, in a sense, we can treat the set R as if it were just one point. It turns out that we can do the derivation of the improved theorem with the sets R easily because the presence of these sets doesn't affect the derivation at all. However, in (future) applications we shall make use of the possibility that we are able to divide the lattice L into equivalence classes R , and therefore all our definitions in Sect. 2 already take this division into account. It is worth emphasizing that the classes γ mentioned previously are of great importance *both* in the derivation *and* in the applications.

In Sect. 2 we present the basic definitions following the notations of Gross [3] as closely as possible. The improved theorem is derived in Sect. 3 by making use of G. Th., and the main result is given by Theorem I. Then finally in Sect. 4 we give some concluding remarks about the result and possible future applications.

2. Notations and Definitions

Let our "lattice" be a countable set L of "points" $a \in L$, and let X denote a compact metric space. Then we have a general statistical mechanical system described by the function $s : L \rightarrow X$, and we define the configuration space for L with the product topology to be $\Omega = X^L$, so that $s \in \Omega$. Furthermore, we have a probability measure σ on the Borel sets of Ω so that

$$\langle f \rangle = \sigma(f) = \int f(s) e^{-S[s]} \mathcal{D}s, \tag{2.1}$$

where $f \in C(\Omega)$. Here $C(\Omega)$ denotes the space of continuous real valued functions on Ω .

Let $d_s(\cdot, \cdot)$ be a state-dependent pseudometric for the set L , i.e., there exists $d_s(a, b) \in \mathbb{R}^+ \cup \{0\}$ for any state $s \in \Omega$ and for any pair $a, b \in L$ such that $d_s(a, b) = d_s(b, a)$, $d_s(a, b) + d_s(b, c) \geq d_s(a, c)$, $d_s(a, b) = 0$ if $a = b$, but it is possible to have $d_s(a, b) = 0$ even if $a \neq b$. Corresponding to each pseudometric d_s we define a division of L into a set of equivalence classes $R \subset L$ by

$$L_{d_s} = \{R \mid R \text{ is an equivalence class under the relation } \sim_{d_s}\}. \tag{2.2}$$

Here $a \sim_{d_s} b$ iff $d_s(a, b) = 0$ for $\forall a, b \in L$. This division can be considered as some kind of a patching.

We can now define a new configuration space by

$$\Omega_{d_s} = \{t \mid t : L_{d_s} \rightarrow X_{d_s}\}, \tag{2.3}$$

where

$$X_{d_s} = \bigcup_{R \in L_{d_s}} X^R, \tag{2.4}$$

and where for any equivalence class $R \in L_{d_s}$ the field value $s_R \in X^R$. The correspondence between Ω and Ω_{d_s} can be expressed by a function $F_{d_s} : \Omega \rightarrow \Omega_{d_s}$, so that $F_{d_s}(t)_R = t|_R$ for each $R \in L_{d_s}$. Similarly we may introduce a meet metric $d_{s_1} \wedge d_{s_2}$ and the corresponding division $L_{d_{s_1} \wedge d_{s_2}}$ by

$$d_{s_1} \wedge d_{s_2} = \inf\{d_{s_1}, d_{s_2}\}, \tag{2.5}$$

$$L_{d_{s_1} \wedge d_{s_2}} = \{R | R \text{ is an equivalence class under the relation } \overbrace{\hspace{1.5cm}}^{d_{s_1} \wedge d_{s_2}}\}. \tag{2.6}$$

It is also possible to write $d_{s_1} \wedge d_{s_2}(a, b)$ in terms of the length of a chain from a to b as follows (see Appendix):

$$d_{s_1} \wedge d_{s_2}(a, b) = \inf\{l(\mathcal{C}) | \mathcal{C} \text{ is a chain from } a \text{ to } b\}. \tag{2.7}$$

We can also divide Ω into equivalence classes γ_i by putting

$$\gamma_i = \{s \in \Omega | d_s(a, b) = d_{s'}(a, b) \text{ for } \forall \text{ pairs } a, b \in L\}, \tag{2.8}$$

where the different equivalence classes are numbered by the value of i , s' is a representative of the i^{th} class and furthermore $\bigcup_i \gamma_i = \Omega$ and $\gamma_i \cap \gamma_j = \emptyset$ for $i \neq j$.

However, when making the division of Ω (which is a topological space) into equivalence classes γ_i we use closures $\bar{\gamma}_i$ because γ_i itself is neither closed nor open. Then the closures $\bar{\gamma}_i$, which are closed sets, can be taken to be Borel sets of Ω , and we define a probability measure σ in each of these Borel sets. The closures $\bar{\gamma}_i$ satisfy $\bigcup_i \bar{\gamma}_i = \Omega$ and we assume that $\sigma(\bar{\gamma}_i \cap \bar{\gamma}_j) = 0$ for $i \neq j$, i.e., the boundaries of $\bar{\gamma}_i$ and the overlapping parts of various $\bar{\gamma}_i$ give just a zero measure. So in spite of the fact that $\bar{\gamma}_i \cap \bar{\gamma}_j$ ($i \neq j$) can be non-void we can essentially work as if we had a division of Ω generated by γ_i .

To each ‘‘point’’ $R \in L_{d_{\gamma_1} \wedge d_{\gamma_2}}$ and to each state $s \in \gamma_1 \cup \gamma_2 \subset \Omega$, we associate a probability measure $\mu_R^{\gamma_1 \gamma_2 \varepsilon}(\cdot | s)$ on the Borel sets of $X_{d_{\gamma_1} \wedge d_{\gamma_2}}$. Here $\varepsilon \in [0, 1]$ is a weight coefficient in the case of two equivalence classes γ_1 and γ_2 . Let $C(\bar{\gamma}_1 \cup \bar{\gamma}_2)$ denote the space of continuous real-valued functions on $\bar{\gamma}_1 \cup \bar{\gamma}_2$. Then we define a mapping $\tau_R^{\bar{\gamma}_1 \bar{\gamma}_2 \varepsilon} : C(\bar{\gamma}_1 \cup \bar{\gamma}_2) \rightarrow C(\bar{\gamma}_1 \cup \bar{\gamma}_2)$ by

$$(\tau_R^{\bar{\gamma}_1 \bar{\gamma}_2 \varepsilon} f)(s) = \int_Y f(x \vee \hat{s}) \mu_R^{\gamma_1 \gamma_2 \varepsilon}(dx | \hat{s}), \tag{2.9}$$

where

$$\hat{s} = s|_{L \setminus R}, \tag{2.10a}$$

$$(x \vee \hat{s})_a = \begin{cases} x_a & \text{for } a \in R \\ \hat{s}_a & \text{for } a \in L \setminus R \end{cases}, \tag{2.10b}$$

$$Y = \{x_R \in X^R | x \vee \hat{s} \in \bar{\gamma}_1 \cup \bar{\gamma}_2\}, \tag{2.10c}$$

and f is a continuous function $f : \bar{\gamma}_1 \cup \bar{\gamma}_2 \rightarrow \mathbb{R}$. We say that the probability measures $\mu_R^{\gamma_1 \gamma_2 \varepsilon}$ form a set of ‘‘conditional probabilities’’ for the statistical mechanical system with a probability measure σ on the Borel sets of Ω ($\bar{\gamma}_1$ and $\bar{\gamma}_2$) provided

$$\sigma[(\varepsilon \chi_{\bar{\gamma}_1} + (1 - \varepsilon) \chi_{\bar{\gamma}_2}) \tau_R^{\bar{\gamma}_1 \bar{\gamma}_2 \varepsilon} f] = \sigma[(\varepsilon \chi_{\bar{\gamma}_1} + (1 - \varepsilon) \chi_{\bar{\gamma}_2}) f] \tag{2.11}$$

for all $R \in L_{d_{\gamma_1} \wedge d_{\gamma_2}}$ and for all $f \in C(\bar{\gamma}_1 \cup \bar{\gamma}_2)$. Here $\chi_{\bar{\gamma}_i}$ is the characteristic function given by

$$\chi_{\bar{\gamma}_i}(s) = \begin{cases} 1 & \text{for } s \in \bar{\gamma}_i \\ 0 & \text{for } s \notin \bar{\gamma}_i \end{cases} \quad (i = 1, 2). \tag{2.12}$$

We may consider $\tau_R^{\gamma_1 \gamma_2 \varepsilon} f$ the conditional expectation $E(f | \mathcal{F})$, where \mathcal{F} is the set of measurable sets F on Ω such that a change in s_b for $b \in R$ does neither bring any s out of nor into the set F . According to the Radon-Nikodym theorem we see that

such a “conditional expectation” exists. We remark that we can do such an argument for any value of ε separately.

To every pair γ_1, γ_2 and to every weight coefficient $\varepsilon \in [0, 1]$ we define

$$\alpha^{\gamma_1 \gamma_2 \varepsilon} = \sup_{R'} \sum_R \beta_{R, R'}^{\gamma_1 \gamma_2 \varepsilon}, \tag{2.13}$$

where $R, R' \in L_{d_{\gamma_1} \wedge d_{\gamma_2}}$ and

$$\beta_{R, R'}^{\gamma_1 \gamma_2 \varepsilon} = (1/2) \sup \{ \|\mu_{R'}^{\gamma_1 \gamma_2 \varepsilon}(\cdot | s) - \mu_{R'}^{\gamma_1 \gamma_2 \varepsilon}(\cdot | t)\|_{\text{var}} \cdot e^{d_{\gamma_1} \wedge d_{\gamma_2}(R, R')} | s, t \in \gamma_1 \cup \gamma_2 \text{ and } s = t \text{ off } \{R\} \}. \tag{2.14}$$

$\alpha^{\gamma_1 \gamma_2 \varepsilon}$ defined in Eq. (2.13) is analogous to the α defined by Gross [3]. However, for our proof in Sect. 3, we have to introduce new α 's as follows

$$\alpha^{\gamma_1 \gamma_2} = \sup_{\varepsilon \in [0, 1]} \alpha^{\gamma_1 \gamma_2 \varepsilon}, \tag{2.15}$$

and

$$\tilde{\alpha} = \sup_{\gamma_1, \gamma_2} \alpha^{\gamma_1 \gamma_2}, \tag{2.16}$$

and furthermore we assume that

$$\tilde{\alpha} < 1. \tag{2.17}$$

Inequality (2.17) is the basic assumption of our theorem.

For the norm of f we take the following expression

$$\|f\|_a = \sum_{b \in L} \sup \{ |f(s) - f(t)| e^{d_s \vee d_t(a, b)} | s = t \text{ off } \{b\} \}, \tag{2.18}$$

where $f \in C(\Omega)$ and the join metric is

$$d_s \vee d_t = \sup \{d_s, d_t\}. \tag{2.19}$$

3. Derivation

In this section we derive an upper bound for the correlation $|\langle fg \rangle - \langle f \rangle \langle g \rangle|$, where $f, g \in C(\Omega)$, $f \equiv f(s)$, $g \equiv g(s)$ and where the average is defined by

$$\langle \mathcal{O}(s) \rangle = \frac{\int \mathcal{O}(s) e^{-S[s]} \mathcal{D}s}{\int e^{-S[s]} \mathcal{D}s}. \tag{3.1}$$

We use the notation

$$\text{cor} = \langle fg \rangle - \langle f \rangle \langle g \rangle = \langle f(s) g(s) \rangle - \langle f(s) \rangle \langle g(s) \rangle. \tag{3.2}$$

Equation (3.2) can be rewritten in a form

$$\text{cor} = \langle\langle f(s) g(s) - f(t) g(s) \rangle\rangle, \tag{3.3}$$

where the double average is defined by

$$\langle\langle \mathcal{O}(s, t) \rangle\rangle = \frac{\int \mathcal{O}(s, t) e^{-S[s] - S[t]} \mathcal{D}s \mathcal{D}t}{\int e^{-S[s] - S[t]} \mathcal{D}s \mathcal{D}t}. \tag{3.4}$$

This double average can also be expressed in terms of the probability measures $\sigma(\gamma_i)$ for the equivalence classes defined by Eq. (2.8). We put

$$\Gamma = \{\gamma_i | i \in I\}, \tag{3.5}$$

where I is some index set parametrizing the family. Equation (3.1) reads

$$\langle \mathcal{O}(s) \rangle = \frac{\sum_{\gamma_i \in \Gamma} \langle \mathcal{O}(s) \rangle_{\gamma_i} \sigma(\gamma_i)}{\sum_{\gamma_i \in \Gamma} \sigma(\gamma_i)}, \tag{3.6}$$

and similarly

$$\langle\langle \mathcal{O}(s, t) \rangle\rangle = \frac{\sum_{\gamma_i, \gamma_j \in \Gamma} \langle\langle \mathcal{O}(s, t) \rangle\rangle_{\gamma_i \times \gamma_j} \sigma(\gamma_i) \sigma(\gamma_j)}{\sum_{\gamma_i, \gamma_j \in \Gamma} \sigma(\gamma_i) \sigma(\gamma_j)}. \tag{3.7}$$

Here $\langle \mathcal{O}(s) \rangle_{\gamma_i}$ means the average of $\mathcal{O}(s)$ which is defined only for those states s belonging to γ_i (analogously $\langle\langle \mathcal{O}(s, t) \rangle\rangle_{\gamma_i \times \gamma_j}$). Actually we should use here $\sigma(\bar{\gamma}_i)$, where $\bar{\gamma}_i$ is the closure but according to the assumption we made below Eq. (2.8) we put $\sigma(\text{boundary of } \gamma_i) = 0$. Equations (3.3) and (3.7) give

$$\text{cor} = \frac{\sum_{\gamma_i, \gamma_j \in \Gamma} \langle\langle f(s)g(s) - f(t)g(s) \rangle\rangle_{\gamma_i \times \gamma_j} \sigma(\gamma_i) \sigma(\gamma_j)}{\sum_{\gamma_i, \gamma_j \in \Gamma} \sigma(\gamma_i) \sigma(\gamma_j)}. \tag{3.8}$$

Next we split cor in Eq. (3.8) into two parts such that

$$\begin{aligned} \text{cor} = & \frac{\sum_{\gamma_i, \gamma_j \in \Gamma} \langle\langle f(s)g(s) \rangle\rangle_{\gamma_i \times \gamma_j} \sigma(\gamma_i) \sigma(\gamma_j)}{\sum_{\gamma_i, \gamma_j \in \Gamma} \sigma(\gamma_i) \sigma(\gamma_j)} \\ & - \frac{\sum_{\gamma_i, \gamma_j \in \Gamma} \langle\langle f(t)g(s) \rangle\rangle_{\gamma_i \times \gamma_j} \sigma(\gamma_i) \sigma(\gamma_j)}{\sum_{\gamma_i, \gamma_j \in \Gamma} \sigma(\gamma_i) \sigma(\gamma_j)}, \end{aligned} \tag{3.9}$$

and denote

$$\text{cor} = \text{cor}_1 - \text{cor}_2, \tag{3.10}$$

where cor_1 and cor_2 refer to the first and second terms on the right-hand side of Eq. (3.9), respectively. We immediately see from the definition of the double average that cor_1 can be written as

$$\text{cor}_1 = \frac{\sum_{\gamma_i \in \Gamma} \langle fg \rangle_{\gamma_i} \sigma(\gamma_i)}{\sum_{\gamma_i \in \Gamma} \sigma(\gamma_i)}. \tag{3.11}$$

Hereafter, we drop the explicit state-dependence. Similarly a straightforward calculation gives

$$\text{cor}_2 = \frac{\sum_{\gamma_i, \gamma_j \in \Gamma} \langle f \rangle_{\gamma_i} \langle g \rangle_{\gamma_j} \sigma(\gamma_i) \sigma(\gamma_j)}{\sum_{\gamma_i, \gamma_j \in \Gamma} \sigma(\gamma_i) \sigma(\gamma_j)}. \tag{3.12}$$

Now using Eqs. (3.11) and (3.12) in Eq. (3.10) we finally obtain

$$\text{cor} = \frac{\sum_{\gamma_i \in \Gamma} [\langle fg \rangle_{\gamma_i} - \langle f \rangle_{\gamma_i} \langle g \rangle_{\gamma_i}] \sigma(\gamma_i)}{\sum_{\gamma_i \in \Gamma} \sigma(\gamma_i)} + \frac{\sum_{\gamma_i, \gamma_j \in \Gamma} [\langle f \rangle_{\gamma_i} \langle g \rangle_{\gamma_i} - \langle f \rangle_{\gamma_i} \langle g \rangle_{\gamma_j}] \sigma(\gamma_i) \sigma(\gamma_j)}{\sum_{\gamma_i, \gamma_j \in \Gamma} \sigma(\gamma_i) \sigma(\gamma_j)}. \tag{3.13}$$

This result could have been obtained of course directly from Eqs. (3.2) and (3.6), but our purpose here was just to introduce some use of the double average which occurs in the final result [see Eq. (3.31)]. The first term on the right-hand side of Eq. (3.13) can be bounded by using the original theorem (G. Th.) because the quantity inside the square brackets includes only one equivalence class. Hence, we get

$$\begin{aligned} & \left| \frac{\sum_{\gamma_i \in \Gamma} [\langle fg \rangle_{\gamma_i} - \langle f \rangle_{\gamma_i} \langle g \rangle_{\gamma_i}] \sigma(\gamma_i)}{\sum_{\gamma_i \in \Gamma} \sigma(\gamma_i)} \right| \\ & \leq \frac{\sum_{\gamma_i \in \Gamma} |\langle fg \rangle_{\gamma_i} - \langle f \rangle_{\gamma_i} \langle g \rangle_{\gamma_i}| \sigma(\gamma_i)}{\sum_{\gamma_i \in \Gamma} \sigma(\gamma_i)} \\ & \leq \frac{\sum_{\gamma_i \in \Gamma} e^{-d_{\gamma_i}(a,b)} \Phi(\alpha^{\gamma_i}) \sigma(\gamma_i)}{\sum_{\gamma_i \in \Gamma} \sigma(\gamma_i)} \Phi_{\parallel \parallel} = \langle e^{-d_{\gamma_i}(a,b)} \Phi(\alpha^{\gamma_i}) \rangle \Phi_{\parallel \parallel}, \end{aligned} \tag{3.14}$$

where $\alpha^{\gamma_i} < 1$ for $\forall i \in I(\alpha^{\gamma_i} < \alpha^{\gamma_i \gamma_j})$ and

$$\Phi(\alpha^{\gamma_i}) = (1 - \alpha^{\gamma_i})^{-2} (1 - (\alpha^{\gamma_i})^2)^{-1}, \tag{3.15a}$$

$$\Phi_{\parallel \parallel} = \|f\|_a \|g\|_b. \tag{3.15b}$$

This is valid for arbitrary pairs a, b of lattice points.

In order to find a bound for the second term on the right-hand side of Eq. (3.13) we can restrict our analysis to two equivalence classes γ_1 and γ_2 . This is why we gave the basic definitions in Sect. 2 in terms of only γ_1 and γ_2 (and ε) [see Eqs. (2.9)–(2.16)]. Making use of Eq. (3.6) we can write down a formula for the average of f in the case of two equivalence classes γ_1 and γ_2 (likewise we could do the following analysis for any pair γ_i, γ_j)

$$\langle f \rangle^{(2)} = \varepsilon \langle f \rangle_{\gamma_1} + (1 - \varepsilon) \langle f \rangle_{\gamma_2}. \tag{3.16}$$

Here the superscript (2) means that we deal only with γ_1 and γ_2 and $\varepsilon \in [0, 1]$ is the weight factor used in Eq. (2.11), so that the correspondence is $\sigma[(\varepsilon \chi_{\gamma_1} + (1 - \varepsilon) \chi_{\gamma_2}) f] = \langle f \rangle^{(2)}$.

It is clear that we also have

$$\langle fg \rangle^{(2)} = \varepsilon \langle fg \rangle_{\gamma_1} + (1 - \varepsilon) \langle fg \rangle_{\gamma_2}. \tag{3.17}$$

By means of Eqs. (3.16) and (3.17) we get after some algebra

$$|\langle fg \rangle^{(2)} - \langle f \rangle^{(2)} \langle g \rangle^{(2)}| = |\varepsilon A + (1 - \varepsilon) \varepsilon B + (1 - \varepsilon) C|, \tag{3.18}$$

where

$$A = \langle fg \rangle_{\gamma_1} - \langle f \rangle_{\gamma_1} \langle g \rangle_{\gamma_1}, \tag{3.19a}$$

$$B = \langle f \rangle_{\gamma_1} \langle g \rangle_{\gamma_1} - \langle f \rangle_{\gamma_1} \langle g \rangle_{\gamma_2} + \langle f \rangle_{\gamma_2} \langle g \rangle_{\gamma_2} - \langle f \rangle_{\gamma_2} \langle g \rangle_{\gamma_1}, \tag{3.19b}$$

$$C = \langle fg \rangle_{\gamma_2} - \langle f \rangle_{\gamma_2} \langle g \rangle_{\gamma_2}. \tag{3.19c}$$

We can now apply G. Th. to the left-hand side of Eq. (3.18) due to the result

$$\alpha^{\gamma_1 \gamma_2} < 1, \tag{3.20}$$

which follows from Eqs. (2.16) and (2.17) (and of course we use all the previous definitions for the case of two equivalence classes in Sect. 2). Therefore we obtain

$$|\varepsilon A + (1 - \varepsilon)\varepsilon B + (1 - \varepsilon)C| \leq K, \tag{3.21}$$

where

$$K = e^{-d_{\gamma_1} \wedge d_{\gamma_2}(a,b)} \Psi. \tag{3.22}$$

Here the meet metric is

$$d_{\gamma_1} \wedge d_{\gamma_2} = \inf\{d_{\gamma_1}, d_{\gamma_2}\}, \tag{3.23}$$

and

$$\Psi = \|f\|_a \|g\|_b (1 - \alpha^{\gamma_1 \gamma_2})^{-2} (1 - (\alpha^{\gamma_1 \gamma_2})^2)^{-1}. \tag{3.24}$$

Actually we are interested in finding a bound for B in Eq. (3.19b), and therefore we use the following lemma.

Lemma. *If $|\varepsilon A + (1 - \varepsilon)\varepsilon B + (1 - \varepsilon)C| \leq K$ for $\forall \varepsilon \in [0, 1]$, then $|A| \leq K$, $|B| \leq 8K$, and $|C| \leq K$.*

Hence

$$|\langle f \rangle_{\gamma_1} \langle g \rangle_{\gamma_1} - \langle f \rangle_{\gamma_1} \langle g \rangle_{\gamma_2} + \langle f \rangle_{\gamma_2} \langle g \rangle_{\gamma_2} - \langle f \rangle_{\gamma_2} \langle g \rangle_{\gamma_1}| \leq 8K. \tag{3.25}$$

First we multiply both sides of Eq. (3.25) by $\sigma(\gamma_1)\sigma(\gamma_2)$, then we sum over all $\gamma_i, \gamma_j \in \Gamma$ and finally divide by

$$\sum_{\gamma_i, \gamma_j \in \Gamma} \sigma(\gamma_i) \sigma(\gamma_j)$$

obtaining

$$\begin{aligned} & \frac{\sum_{\gamma_i, \gamma_j \in \Gamma} |\langle f \rangle_{\gamma_i} \langle g \rangle_{\gamma_i} - \langle f \rangle_{\gamma_i} \langle g \rangle_{\gamma_j}| \sigma(\gamma_i) \sigma(\gamma_j)}{\sum_{\gamma_i, \gamma_j \in \Gamma} \sigma(\gamma_i) \sigma(\gamma_j)} \\ & \leq 4 \cdot \frac{\sum_{\gamma_i, \gamma_j \in \Gamma} e^{-d_{\gamma_i} \wedge d_{\gamma_j}(a,b)} \Psi(\alpha^{\gamma_i \gamma_j}) \sigma(\gamma_i) \sigma(\gamma_j)}{\sum_{\gamma_i, \gamma_j \in \Gamma} \sigma(\gamma_i) \sigma(\gamma_j)} \Psi_{\parallel} \\ & = 4 \cdot \ll e^{-d_{\gamma_i} \wedge d_{\gamma_j}(a,b)} \Psi(\alpha^{\gamma_i \gamma_j}) \gg \Psi_{\parallel}, \end{aligned} \tag{3.26}$$

where

$$\Psi(\alpha^{\gamma_i \gamma_j}) = (1 - \alpha^{\gamma_i \gamma_j})^{-2} (1 - (\alpha^{\gamma_i \gamma_j})^2)^{-1}, \tag{3.27a}$$

$$\Psi_{\parallel \parallel} = \|f\|_a \|g\|_b = \Phi_{\parallel \parallel}. \tag{3.27b}$$

Thus, this result gives the required bound for the second term on the left-hand side of Eq. (3.13).

Equations (3.13), (3.14), and (3.26) together give

$$|\text{cor}| \leq \{ \langle e^{-d_{\gamma_i(a,b)}} \Phi(\alpha^{\gamma_i}) \rangle + 4 \cdot \langle\langle e^{-d_{\gamma_i \wedge d_{\gamma_j(a,b)}}} \Psi(\alpha^{\gamma_i \gamma_j}) \rangle\rangle \} \cdot \|f\|_a \|g\|_b. \tag{3.28}$$

Using the definition (2.16) we can show that $\Psi(\alpha^{\gamma_i \gamma_j}) \leq \Psi(\tilde{\alpha})$ for $\forall i, j$ and $\Phi(\alpha^{\gamma_i}) \leq \Phi(\tilde{\alpha})$ for $\forall i$ (here we assume that $\alpha^{\gamma_i \gamma_i} = \alpha^{\gamma_i}$), where of course $\Phi(\tilde{\alpha}) = \Psi(\tilde{\alpha})$. Then from Eq. (3.28) it immediately follows that

$$|\text{cor}| \leq \{ \langle e^{-d_{\gamma_i(a,b)}} \rangle + 4 \cdot \langle\langle e^{-d_{\gamma_i \wedge d_{\gamma_j(a,b)}}} \rangle\rangle \} \cdot \Psi(\tilde{\alpha}) \|f\|_a \|g\|_b. \tag{3.29}$$

Furthermore because the meet metric $d_{\gamma_i \wedge d_{\gamma_j}}$ is given as an infimum [see Eq. (3.23)] we see that

$$\langle e^{-d_{\gamma_i(a,b)}} \rangle \leq \langle\langle e^{-d_{\gamma_i \wedge d_{\gamma_j(a,b)}}} \rangle\rangle. \tag{3.30}$$

Thus we have

$$|\langle fg \rangle - \langle f \rangle \langle g \rangle| \leq 5 \cdot \langle\langle e^{-d_{\gamma_i \wedge d_{\gamma_j(a,b)}}} \rangle\rangle \|f\|_a \|g\|_b \cdot (1 - \tilde{\alpha})^{-2} (1 - \tilde{\alpha}^2)^{-1}. \tag{3.31}$$

Now we can state our main result:

Theorem I. *If $\tilde{\alpha} < 1$ then for any functions $f, g \in C(\Omega)$*

$$|\langle fg \rangle - \langle f \rangle \langle g \rangle| \leq 5 \cdot \langle\langle e^{-d_{\gamma_i \wedge d_{\gamma_j(a,b)}}} \rangle\rangle \|f\|_a \|g\|_b \cdot (1 - \tilde{\alpha})^{-2} (1 - \tilde{\alpha}^2)^{-1},$$

where the double average of $e^{-d_{\gamma_i \wedge d_{\gamma_j(a,b)}}$ is given by

$$\langle\langle e^{-d_{\gamma_i \wedge d_{\gamma_j(a,b)}}} \rangle\rangle = \frac{\int e^{-\min\{d_s(a,b), d_t(a,b)\}} e^{-S[s] - S[t]} \mathcal{D}s \mathcal{D}t}{(\int e^{-S[s]} \mathcal{D}s)^2}.$$

This theorem is the improved version of G. Th.

4. Conclusions

Looking at the right-hand side of Eq. (3.31) the numerical factor of 5 seems slightly odd because strictly speaking we don't get back G. Th. as a limiting case, but we believe, however, that we can use this particular form of inequality in our applications. Therefore it is not necessary to get rid of this factor by e.g. inventing some new lemma instead of that given in Sect. 3.

What is most important for us is the possible exponential decay of the correlation. However, it seems that this property somehow is in disguise because instead of the ordinary metric we have the meet (infimum) metric, and in addition, we have to perform the double averaging in the exponential factor. How to calculate the double average and how to interpret the meet metric (for one

suggestion see Appendix) is evidently the problem of applications and therefore we refer to [1]. We also refer to [1] concerning the estimation of the parameter $\tilde{\alpha}$. Nevertheless we emphasize that it was crucial for our derivation in Sect. 3 that $\tilde{\alpha}$ is smaller than unity. Hence it is clear that the property $\tilde{\alpha} < 1$ induces strong restrictions for the action $S[s]$ because essentially $\tilde{\alpha}$ depends only on $S[s]$. In [1] we shall show that the Regge calculus action has indeed the desired property of allowing $\tilde{\alpha}$ to be smaller than unity (actually the situation is more involved because it turns out that we need some extra terms in the action in addition to the Regge calculus one).

As it is remarked in [3] it might be interesting to generalize G. Th., and also our version of the theorem, in such a way that the space X would be any measurable space and that the function f would be a bounded measurable function of Ω although, frankly speaking, we really don't have any need for such a generalization, at least in applications. One such approach was put forward in [5].

Finally we would like to mention that it might be useful to have a genuine continuum version of the theorem. Then, probably, we could apply the theorem more easily because we wouldn't have the extra worry how to relate the lattice points of L and the points in a Riemannian manifold, say.

Appendix

We prove that the meet metric $d_1 \wedge d_2(a, b)$ can be expressed in terms of the length of a chain from a to b where $a, b \in L$.

Define a chain \mathcal{C} from a to b to be a sequence of points all belonging to the lattice L , i.e.,

$$\mathcal{C} = (a \equiv c_0, c_1, \dots, c_{n-1}, c_n \equiv b), \tag{A.1}$$

and the length $l(\mathcal{C})$ of a chain \mathcal{C} to be given by

$$l(\mathcal{C}) = \sum_{i=1}^n \min \{d_1(c_{i-1}, c_i), d_2(c_{i-1}, c_i)\}. \tag{A.2}$$

Consider just a special chain $\mathcal{C} = (a, b)$, that is to say a "direct chain" from a to b , so that

$$l(\mathcal{C}) = \min \{d_1(a, b), d_2(a, b)\}. \tag{A.3}$$

We immediately see that

$$l(\mathcal{C}) \leq d_i(a, b) \quad (i = 1, 2), \tag{A.4}$$

and furthermore that

$$\inf \{l(\mathcal{C})\} \leq d_i(a, b) \quad (i = 1, 2). \tag{A.5}$$

This means that we have found a lower bound for $\{d_1, d_2\}$. Now we have to show that $\inf \{l(\mathcal{C})\}$ is the greatest lower bound, i.e., the infimum.

Let us introduce a metric d_k in such a way that $d_k \leq d_1$ and $d_k \leq d_2$. By using the triangle inequality and Eq. (A.2) we find

$$d_k(a, b) \leq \sum_{i=1}^n d_k(c_{i-1}, c_i) \leq \sum_{i=1}^n \min \{d_1(c_{i-1}, c_i), d_2(c_{i-1}, c_i)\} = l(\mathcal{C}). \tag{A.6}$$

Inequality (A.6) is true for all \mathcal{C} and we can conclude that

$$d_k(a, b) \leq \inf\{l(\mathcal{C})\}. \quad (\text{A.7})$$

Then from (A.5) and (A.7) we obtain

$$d_k(a, b) \leq \inf\{l(\mathcal{C})\} \leq d_i(a, b) \quad (i = 1, 2), \quad (\text{A.8})$$

which gives the desired result [Eq. (2.7) in the text]

$$d_1 \wedge d_2(a, b) = \inf\{d_1(a, b), d_2(a, b)\} = \inf\{l(\mathcal{C}) \mid \mathcal{C} \text{ is a chain from } a \text{ to } b\}. \quad (\text{A.9})$$

We mention that Eq. (A.9) is very useful in connection with applications of our improved theorem.

Acknowledgements. We would like to thank H. Schlereth for useful criticism and discussions. One of us (M. N.) thanks the Niels Bohr Institute and Nordita for financial support for his visit. He is supported by a grant of the U.S. Department of Energy under Contract No. DE-AC02-76ERO3130.A011-Task A.

References

1. Lehto, M., Nielsen, H.B., Ninomiya, M.: In preparation
2. Regge, T.: General relativity without coordinates. *Nuovo Cimento* **19**, 558 (1961)
3. Gross, L.: Decay of correlations in classical lattice models at high temperature. *Commun. Math. Phys.* **68**, 9 (1979)
4. Israel, R.B.: High-temperature analyticity in classical lattice systems. *Commun. Math. Phys.* **50**, 245 (1976)
5. Klein, D.: Dobrushin uniqueness techniques and the decay of correlations in continuum statistical mechanics. *Commun. Math. Phys.* **86**, 227 (1982)
6. Dobrushin, R.L.: The description of a random field by means of conditional probabilities and conditions on its regularity. *Theor. Prob. Appl.* **13**, 197 (1968)
7. Kelley, J.L.: *General topology*. New York: Van Nostrand 1955

Communicated by G. Mack

Received July 4, 1983; in revised form January 16, 1984

