Commun. Math. Phys. 92, 81–119 (1983)

Communications in Mathematical Physics

Charged States in $\mathbb{Z}_2$ Gauge Theories

Klaus Fredenhagen* and Mihail Marcu
Fakultät für Physik, Universität Freiburg, D-7800 Freiburg, Federal Republic of Germany

Abstract. Charged translation covariant states with finite energy are constructed in the Higgs phase of the $\mathbb{Z}_2$ gauge theory coupled to a $\mathbb{Z}_2$ matter field.

1. Introduction

The spectrum of charges in gauge theories depends on subtle properties of the dynamics, and it is extremely difficult to get information on them within the known approximation schemes. An easier problem is the investigation of lattice gauge theories where the powerful methods of classical statistical mechanics can be applied, and one may hope that certain properties of the lattice theory will survive in the continuum limit. However, on the lattice the investigation of the charge structure is not easy either. The emphasis has rather been on confinement criteria, the most prominent one being the Wilson criterion [1]. This criterion checks whether the energy of a system with two external charges increases linearly with the distance between the charges. If this happens, it is interpreted as a sign that it is impossible to create single charges with finite energy. Unfortunately, the Wilson criterion cannot be used to test the existence of dynamical charges in gauge theories with matter fields [2]. This is however the more interesting question. Criteria which are applicable in this case have been proposed by Mack and Meyer [3] and Bricmont and Fröhlich [4]. However, the implications of their criteria for the confinement problem are not clear.

The aim of this work is to understand the nature of charges in a gauge theory with matter fields. Our analysis leads to confinement criteria which directly test the existence of charges [5] and some of their most relevant properties.

In gauge theories with a discrete gauge group there exists a weak coupling expansion. In the pure gauge theory the Wilson loop obeys a perimeter law [6–8, 2]. One may ask whether charged states with finite energy will exist if the gauge field is coupled to a matter field. This has been conjectured by several

* Heisenberg fellow

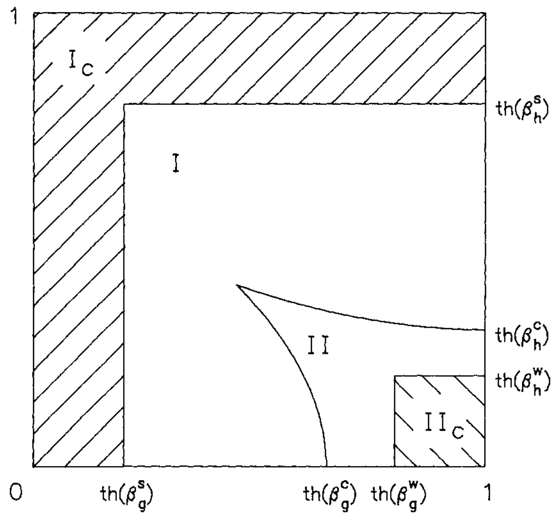

Fig. 1.1. Conjectured phase diagram of the gauge invariant Ising model

authors [9, 10], however a proof is still lacking. The main problem consists in the fact that charges in gauge theories are sources of electric flux lines; an isolated charge is therefore tied to an infinitely extended string of electric flux. It is not clear from the beginning whether such a string can be chosen such that the charged state has finite energy, and it is not easy to guess the dependence of the state on the form of the string.

In continuum quantum field theory the localization properties of charged particles have been analyzed in a general context, and it has been found that in the absence of physical massless particles a charged particle can always be localized in a stringlike region [11]. In this context, the expected behaviour of charges in gauge theories is the worst possible case. The asymptotic direction of the string, however, can never be observed. On the other hand, the derived localization properties lead already to the usual structure in the set of particle states, hence there seem to be no reasons against the occurrence of particles which are only stringlike localizable; the most natural candidates are charged particles in gauge theories.

In this paper we construct charged states for the $\mathbb{Z}_2$ gauge theory coupled to a $\mathbb{Z}_2$ Higgs field in $(d+1)$ dimensions $(d \geqq 2)$. The Euclidean version of this model is the so-called gauge invariant Ising model which has first been discussed by Wegner [6]. Its phase structure is relatively well known (Fig. 1.1). In Region I one expects that there exist no charged particles with finite energy (confinement/screening). Region II (the Higgs phase in the terminology of 't Hooft [12]) may contain charged particles. We shall exploit the fact that both regions contain subregions $\mathrm{I_c}$ and $\mathrm{II_c}$, respectively (the shaded areas in Fig. 1.1), where convergent expansions are known. We shall study a sequence of states which describe a pair of charges separated by an increasing distance and which are regularized such that their energy is uniformly bounded. In Region $\mathrm{II_c}$ the sequence will converge to a state ω orthogonal to the vacuum sector, whereas in Region $\mathrm{I_c}$ the limit state remains in the vacuum sector. One observes several remarkable differences between the Regions I_c and II_c. In I_c the vacuum component of the vectors of the approximating sequence does not vanish in the limit, in obvious contrast to the behaviour in Region II_c. The ratio of the expectation value of the charge operator in the state ω to that in vacuum is -1 in Region II_c and $+1$ in Region I_c. A third distinctive feature is the occurrence of strong correlations of electric fluxes in Region II_c but not in I_c. These strong correlations prevent the determination of the asymptotic direction of the string accompanying a charged particle and are a necessary condition for the existence of those particles according to the general analysis in [11]. One may use these properties as criteria distinguishing the confinement/screening phase from the phase where charges exist. The use of these criteria for numerical studies will be discussed in a forthcoming paper [13].

The present paper is organized as follows. In Sect. 2 the algebra of the model is introduced, and it is shown that the model cannot possess charged states which are created by local fields. Hence any charge must be of the gauge type, i.e. it can be determined via a version of Gauß' law in the complement of any finite region. In Sect. 3 we introduce the dynamics of the model by the Euclidean method [14, 15] and express the ground state in terms of a Gibbs state of the gauge invariant Ising model. In Sect. 4 the construction of the Gibbs state of the gauge invariant Ising model in the Higgs phase using the expansion of Marra and Miracle-Solé [10] is reviewed. Section 5 is devoted to the construction of charged states. Their behaviour under lattice translations and imaginary time translations is discussed in Sect. 6. The real time translations are studied in Sect. 7; this finally leads to the proof that the charged states constructed in Sect. 5 are orthogonal to the vacuum sector. The proofs rely heavily on the method of polymer expansions (see e.g. [2]). For the convenience of the reader an outline of the method is given in Appendix 1. In Appendix 4 we give an estimate of the convergence Region II_c. We also present a bound on the parameter appearing in the perimeter law for the Wilson loop in the pure gauge theory (Appendix 2) and some combinatorial and geometrical estimates which are needed several times (Appendices 3–5).

2. Fields and Observables

The model is defined on the hypercubic lattice $\mathbb{Z}^d$, $d \geqq 2$. We shall denote the distance between the lattice points $\mathbf{x}=(x^1, \ldots, x^d)$ and $\mathbf{y}=(y^1, \ldots, y^d)$ by $|\mathbf{x}-\mathbf{y}|$,

$$|\mathbf{x}-\mathbf{y}| = \sum_{i=1}^{d} |x^i - y^i| . \tag{2.1}$$

For each lattice point $\mathbf{x}$ we have hermitean operators, $\sigma_3(\mathbf{x})$ and $\sigma_1(\mathbf{x})$. They are the discrete analogues of the Higgs field and its canonically conjugate momentum and have the algebraic properties of Pauli matrices,

$$\begin{aligned} \sigma_1(\mathbf{x})^2 = \sigma_3(\mathbf{x})^2 = 1 , \\ \sigma_1(\mathbf{x})\sigma_3(\mathbf{x}) = -\sigma_3(\mathbf{x})\sigma_1(\mathbf{x}) . \end{aligned} \tag{2.2}$$

For each lattice bond $\mathbf{b}$ (i.e. a pair of lattice points with distance 1) we have hermitean operators, $\tau_3(\mathbf{b})$ and $\tau_1(\mathbf{b})$, representing the $\mathbb{Z}_2$ gauge field and the corresponding electric field, respectively, again with the algebraic relations

$$\begin{aligned} \tau_1(\mathbf{b})^2 &= \tau_3(\mathbf{b})^2 = 1\,, \\ \tau_1(\mathbf{b})\tau_3(\mathbf{b}) &= -\tau_3(\mathbf{b})\tau_1(\mathbf{b})\,. \end{aligned} \tag{2.3}$$

Operators associated to different lattice points or bonds commute.

There is a unique C^*-algebra $\mathfrak{F}$ which contains as a norm dense subalgebra the $*$-algebra $\mathfrak{F}_0$ generated by all these operators. The subalgebras $\mathfrak{F}(\Lambda)$ associated to subsets $\Lambda \subset \mathbb{Z}^d$ are C^*-algebras generated by $\sigma_i(\mathbf{x})$, $\tau_j(\mathbf{b})$, $i,j=1,3$, $\mathbf{x}\in\Lambda$, $\mathbf{b}\subset\Lambda$. Clearly, $\mathfrak{F}_0 = \bigcup_{|\Lambda|<\infty} \mathfrak{F}(\Lambda)$. The "local net" $\Lambda \to \mathfrak{F}(\Lambda)$ is covariant under lattice translations,

$$\alpha_{\mathbf{x}}(\mathfrak{F}(\Lambda)) = \mathfrak{F}(\Lambda + \mathbf{x})\,. \tag{2.4}$$

Here $\alpha_{\mathbf{x}}$, $\mathbf{x}\in\mathbb{Z}^d$, denotes the $*$-automorphism of $\mathfrak{F}$ with the properties

$$\alpha_{\mathbf{x}}(\sigma_i(\mathbf{y})) = \sigma_i(\mathbf{y}+\mathbf{x})\,, \qquad \alpha_{\mathbf{x}}(\tau_i(\mathbf{b})) = \tau_i(\mathbf{b}+\mathbf{x})\,, \qquad i=1,3\,. \tag{2.5}$$

The dynamics which will be introduced in Sect. 3 is gauge invariant, i.e. invariant under the automorphisms implemented by the unitary operators

$$\hat{q}(\mathbf{x}) = \sigma_1(\mathbf{x})\delta^*\tau_1(\mathbf{x})\,, \tag{2.6}$$

where $\delta^*\tau_1(\mathbf{x}) = \prod_{\mathbf{b}\ni\mathbf{x}} \tau_1(\mathbf{b})$ denotes the divergence of τ_1. Therefore only the gauge invariant part $\mathfrak{A}$ of $\mathfrak{F}$,

$$\mathfrak{A} = \{A\in\mathfrak{F}, \hat{q}(\mathbf{x})A = A\hat{q}(\mathbf{x}) \text{ for all } \mathbf{x}\in\mathbb{Z}^d\}\,, \tag{2.7}$$

is considered as the algebra of observables, whereas $\mathfrak{F}$ is called the field algebra.

The algebra of observables $\mathfrak{A}$ has a nontrivial center which is generated by the operators $\hat{q}(\mathbf{x})$, $\mathbf{x}\in\mathbb{Z}^d$. $\hat{q}(\mathbf{x})$ is interpreted as the observable which measures an external charge at the point $\mathbf{x}$. In a factorial[1] representation π of $\mathfrak{A}$ the external charges are multiples of the identity,

$$\pi(\hat{q}(\mathbf{x})) = q_\pi(\mathbf{x}) \quad \mathbb{1}\,, \tag{2.8}$$

and $q_\pi : \mathbb{Z}^d \to \{\pm 1\}$ is called the external charge configuration in π. The main interest is on representations π without external charges, $q_\pi \equiv 1$. In these representations the two-sided norm closed ideal J of $\mathfrak{A}$ generated by $\hat{q}(\mathbf{x}) - 1$ for all $\mathbf{x}\in\mathbb{Z}^d$ is annihilated. Hence the relevant algebra of observables is the quotient

$$\mathfrak{B} = \mathfrak{A}/J\,. \tag{2.9}$$

Both $\mathfrak{A}$ and $\mathfrak{B}$ inherit a local structure from $\mathfrak{F}$, which is again translation covariant,

$$\mathfrak{A}(\Lambda) = \mathfrak{F}(\Lambda)\cap\mathfrak{A},\ \mathfrak{B}(\Lambda) = \{A+J, A\in\mathfrak{A}(\Lambda)\}\,. \tag{2.10}$$

The algebra $\mathfrak{B}$ is generated by $u_1(\mathbf{b}) = \tau_1(\mathbf{b}) + J$, $u_3(\mathbf{b}) = \delta\sigma_3(\mathbf{b})\tau_3(\mathbf{b}) + J$[2], where $\mathbf{b}$ runs over the set of lattice bonds. $u_1(\mathbf{b})$ and $u_3(\mathbf{b})$ have again the algebraic

---

1 A representation π is called factorial if the center of the weak closure of $\pi(\mathfrak{A})$ consists of multiples of the identity

2 $\delta\sigma_3(\mathbf{b}) = \prod_{\mathbf{x}\in\mathbf{b}} \sigma_3(\mathbf{x})$ is the exterior derivative of σ_3

properties of Pauli matrices. The local algebra $\mathfrak{B}(\Lambda)$ is generated by $u_1(\mathbf{b})$, $u_3(\mathbf{b})$, $\mathbf{b}\subset\Lambda$. Let $\mathfrak{B}(\Lambda^c)$ denote the C^*-algebra generated by $u_1(\mathbf{b})$, $u_3(\mathbf{b})$, $\mathbf{b}\not\subset\Lambda$, and let $\mathfrak{B}(\Lambda)^c$ denote the relative commutant of $\mathfrak{B}(\Lambda)$,

$$\mathfrak{B}(\Lambda)^c=\{A\in\mathfrak{B}, AB=BA \text{ for all } B\in\mathfrak{B}(\Lambda)\}. \tag{2.11}$$

These algebras are related by the following proposition (compare [16, Theorem 2.5.10]).

Proposition 2.1.

(i) $\mathfrak{B}(\Lambda)^c=\mathfrak{B}(\Lambda^c)$.

(ii) $\mathfrak{B}(\Lambda)=\mathfrak{B}(\Lambda^c)^c$.

(iii) *If π_1 and π_2 are disjoint representations of $\mathfrak{B}$ (i.e. no subrepresentation of π_1 is equivalent to a subrepresentation of π_2), then also their restrictions to $\mathfrak{B}(\Lambda^c)$ for any finite $\Lambda\subset\mathbb{Z}^d$ are disjoint.*

Proof. (i), (ii) For a set $\mathbf{L}$ of lattice bonds, let $\mathfrak{B}(\mathbf{L})$ denote the C^*-algebra generated by $u_1(\mathbf{b})$, $u_3(\mathbf{b})$, $\mathbf{b}\in\mathbf{L}$, and let $\mathbf{L}^c$ denote the complement of $\mathbf{L}$ in the set of all lattice bonds in $\mathbb{Z}^d$. Then (i) and (ii) are special cases of the relation

$$\mathfrak{B}(\mathbf{L})^c=\mathfrak{B}(\mathbf{L}^c). \tag{$*$}$$

Let $A\in\mathfrak{B}(\mathbf{L})^c$ and $\varepsilon>0$. Then there exists some $B\in\mathfrak{B}(\mathbf{M})$ for some finite set of bonds $\mathbf{M}$ such that $\|A-B\|<\varepsilon$. Let G denote the finite group generated by $u_1(\mathbf{b})$, $u_3(\mathbf{b})$, $\mathbf{b}\in\mathbf{M}\cap\mathbf{L}$, and consider the following mean on $\mathfrak{B}$,

$$m(C)=|G|^{-1}\sum_{g\in G} gCg^{-1}, \qquad C\in\mathfrak{B}.$$

Clearly $m(B)\in\mathfrak{B}(\mathbf{L}^c\cap\mathbf{M})$ and $m(A)=A$, hence

$$\|m(B)-A\|=\|m(B-A)\|\leqq\|B-A\|<\varepsilon,$$

and A is in the norm closure of $\bigcup_{|\mathbf{M}|<\infty}\mathfrak{B}(\mathbf{L}^c\cap\mathbf{M})$.

This proves (i) and (ii).

(iii) Two representations π_1 and π_2 of $\mathfrak{B}$ are disjoint if and only if there exists a net (A_λ) in $\mathfrak{B}$ such that

$$\pi_1(A_\lambda)\underset{w}{\rightarrow}1, \qquad \pi_2(A_\lambda)\underset{w}{\rightarrow}-1$$

($\underset{w}{\rightarrow}$ means convergence in the weak operator topology) [16]. Let $\mathbf{L}$ be a finite set of lattice bonds, and let m denote the mean over the group generated by $u_1(\mathbf{b})$, $u_3(\mathbf{b})$, $\mathbf{b}\in\mathbf{L}$. Then $m(A_\lambda)\in\mathfrak{B}(\mathbf{L})^c=\mathfrak{B}(\mathbf{L}^c)$, and since m is continuous with respect to the weak operator topology

$$\pi_1(m(A_\lambda))\underset{w}{\rightarrow}1, \qquad \pi_2(m(A_\lambda))\underset{w}{\rightarrow}-1,$$

hence $\pi_1\restriction\mathfrak{B}(\mathbf{L}^c)$ and $\pi_2\restriction\mathfrak{B}(\mathbf{L}^c)$ are disjoint. q.e.d.

Proposition 2.1 (iii) shows that $\mathfrak{B}$ cannot possess charged states with a charge generated by local fields. Every charge must be of the gauge type, i.e. it can be measured in the complement of an arbitrarily large but finite region $\Lambda\subset\mathbb{Z}^d$. The same remark applies to the algebra $\mathfrak{F}$.

It is possible to use only the algebra $\mathfrak{B}$ in an investigation of the charge structure of the model. However, the redundant description in terms of $\mathfrak{F}$ and $\mathfrak{A}$ is much more transparent; therefore we shall work in this framework, keeping in mind that the physical relevance of a statement has to be checked in terms of $\mathfrak{B}$.

3. Dynamics and Ground State

The dynamics of the model shall be introduced by the Euclidean method. The local Hamiltonian H_Λ, $|\Lambda|<\infty$, is defined implicitly in terms of a local transfer matrix $T_\Lambda = e^{-H_\Lambda}$,

$$T_\Lambda = e^{\frac{1}{2}A_\Lambda} e^{B_\Lambda} e^{\frac{1}{2}A_\Lambda}, \tag{3.1}$$

with $A_\Lambda = \sum_{\mathbf{p}\in P(\Lambda)} \beta_g \delta\tau_3(\mathbf{p}) + \sum_{\mathbf{b}\in B(\Lambda)} \beta_h \delta\sigma_3(\mathbf{b})\tau_3(\mathbf{b})$,

$$B_\Lambda = \sum_{\mathbf{b}\in B(\Lambda)} \beta_g^* \tau_1(\mathbf{b}) + \sum_{\mathbf{x}\in\Lambda} \beta_h^* \sigma_1(\mathbf{x}),$$

[$B(\Lambda)$ set of bonds $\mathbf{b}\subset\Lambda$, $P(\Lambda)$ set of plaquettes (quadruples of lattice points spanning a square with side length 1) $\mathbf{p}\subset\Lambda$, $\delta\tau_3(\mathbf{p}) = \prod_{\substack{\mathbf{b}\in B(\Lambda)\\ \mathbf{b}\subset\mathbf{p}}} \tau_3(\mathbf{b})$, $\beta^* = -\frac{1}{2}\ln\operatorname{th}\beta$,

$0<\beta_g, \beta_h<\infty$]. T_Λ is a gauge invariant, positive, invertible operator in $\mathfrak{A}(\Lambda)$. It implements a (non $*-$) automorphism of $\mathfrak{F}$

$$\alpha_i^\Lambda(A) = T_\Lambda A T_\Lambda^{-1}, \quad A\in\mathfrak{F}. \tag{3.2}$$

Since $\alpha_i^{\Lambda_1}(A) = \alpha_i^{\Lambda_2}(A)$, $A\in\mathfrak{F}(\Lambda)$, if $\Lambda_1, \Lambda_2 \supset \hat{\Lambda}$, $\hat{\Lambda} = \{\mathbf{x}\in\mathbb{Z}^d,\ \operatorname{dist}(\mathbf{x},\Lambda)\leqq 2\}$, the automorphisms α_i^Λ converge to an automorphism α_i of $\mathfrak{F}_0$ as Λ tends to $\mathbb{Z}^d$. α_i may be interpreted as the time translation by one unit in imaginary direction [3].

Unfortunately it is not clear whether the real local time translations $\alpha_t^\Lambda(\cdot) = T_\Lambda^{-it} \cdot T_\Lambda^{it}$ also converge, and it is difficult to see whether α_i determines directly the global real time translations α_t (if this is the case, α_i is called the analytic generator of α_t [17]). However, it is possible to introduce the notion of a ground state using only α_i.

Definition. A ground state ω_0 of $\mathfrak{F}$ with respect to α_i is an α_i-invariant state with

$$0 \leqq \omega_0(A^*\alpha_i(A)) \leqq \omega_0(A^*A), \quad A\in\mathfrak{F}_0. \tag{3.3}$$

[The α_i-invariance is actually a consequence of (3.3).] If a ground state ω_0 is given one may introduce the global dynamics by the following method. In the GNS representation π_0 of $\mathfrak{F}$ in a Hilbert space $\mathscr{H}_0$ with a cyclic vector $\Omega\in\mathscr{H}_0$, which is characterized by

$$(\Omega, \pi_0(A)\Omega) = \omega_0(A), \quad A\in\mathfrak{F}, \tag{3.4}$$

one can define an operator T_0 by

$$T_0\pi_0(A)\Omega = \pi_0\alpha_i(A)\Omega, \quad A\in\mathfrak{F}_0. \tag{3.5}$$

---

3 α_i is not a $*$-automorphism but fulfills $(\alpha_i(A))^* = \alpha_i^{-1}(A^*)$. We shall use the notation $(\alpha_i)^n = \alpha_{in}$, $n\in\mathbb{Z}$

T_0 is well defined and hermitian and has a densely defined inverse,

$$T_0^{-1}\pi_0(A)\Omega = \pi_0\alpha_{-i}(A)\Omega\,, \qquad A\in\mathfrak{F}_0\,. \tag{3.6}$$

From (3.3)

$$0 \leqq T_0 \leqq 1\,. \tag{3.7}$$

The real time translations are defined by

$$\hat{\alpha}_t\pi_0(A) = T_0^{-it}\pi_0(A)T_0^{it}\,, \tag{3.8}$$

which is possible since T_0 has a densely defined inverse and therefore no zero eigenvalue. For $A\in\mathfrak{F}_0$

$$T_0\pi_0(A)T_0^{-1} = \pi_0\alpha_i(A)\,. \tag{3.9}$$

By interpolation arguments [18, Sect. IX.4]

$$\|T_0^z\pi_0(A)T_0^{-z}\| \leqq \|A\|^{1-\mathrm{Re}\,z}\|\alpha_i(A)\|^{\mathrm{Re}\,z}\,, \qquad 0 \leqq \mathrm{Re}\,z \leqq 1\,,$$

hence $z\to\hat{\alpha}_z\pi_0(A) = T_0^{-iz}\pi_0(A)T_0^{iz}$ is entirely analytic and

$$\hat{\alpha}_{in}\pi_0(A) = \pi_0\alpha_{in}(A)\,, \qquad n\in\mathbb{Z}\,. \tag{3.10}$$

This method of defining real times translations has first been discussed by Lüscher [15] and Osterwalder and Seiler [14] in their work on the quantum mechanical interpretation of Euclidean lattice gauge theories. Note that the problem of possible zero eigenvalues of the transfer matrix T_0 [19] does not appear in our case due to the following two facts [20]:

(i) The local transfer matrices are invertible, hence α_i^Λ exists.

(ii) The automorphisms α_i^Λ converge strongly on $\mathfrak{F}_0$ to an automorphism of $\mathfrak{F}_0$.

It is not known under which conditions the automorphisms $\hat{\alpha}_t$ leave the original algebra $\mathfrak{F}$ or at least its weak closure in the representation π_0 invariant, so it is not clear whether the dynamics is almost local, in contrast to the dynamics defined with local Hamiltonians where locality properties are under control [16]. We shall eventually consider the (probably larger) algebras $\hat{\mathfrak{F}}$, $\hat{\mathfrak{A}}$, and $\hat{\mathfrak{B}}$, where $\hat{\mathfrak{F}}$ is the C^*-algebra generated by $\hat{\alpha}_t\pi_0(A)$, $A\in\mathfrak{F}$, $t\in\mathbb{R}$, $\hat{\mathfrak{A}}$ is the gauge invariant part of $\hat{\mathfrak{F}}$ and $\hat{\mathfrak{B}} = \hat{\mathfrak{A}}/\hat{J}$ with $\hat{J}$ denoting the two-sided ideal in $\hat{\mathfrak{A}}$ generated by $\pi_0(J)$.

$\mathfrak{F}$ has always a ground state. This may be seen as follows. To each finite $\Lambda\subset\mathbb{Z}^d$ there exists a state ω_Λ on $\mathfrak{F}$ with

$$\omega_\Lambda(T_\Lambda) = \|T_\Lambda\|\,. \tag{3.11}$$

ω_Λ is a ground state of $\mathfrak{F}$ with respect to α_i^Λ in the sense of (3.3), and any ground state with respect to α_i^Λ fulfills (3.11). Since the set of states is compact in the weak-$*$-topology and $\mathfrak{F}$ is separable, there exists an increasing sequence (Λ_n) of finite subsets of $\mathbb{Z}^d$ such that $(\omega_{\Lambda_n}(A))$ converges for each $A\in\mathfrak{F}$. The limit state ω_0,

$$\omega_0(A) = \lim_{n\to\infty}\omega_{\Lambda_n}(A)\,, \tag{3.12}$$

satisfies (3.3) and is therefore a ground state of $\mathfrak{F}$ with respect to α_i. Actually, $\omega_\Lambda \restriction \mathfrak{F}(\Lambda)$ is a uniquely determined pure state, and the net (ω_Λ) is convergent.

This well known fact [14, 15] follows from the following argument. Let.

$$E^{\sigma}_{ss'}(\mathbf{x})=\sigma_1(\mathbf{x})^{\frac{1}{2}(1-s)}\tfrac{1}{2}(1+\sigma_3(\mathbf{x}))\sigma_1(\mathbf{x})^{\frac{1}{2}(1-s')},$$
$$E^{\tau}_{ss'}(\mathbf{b})=\tau_1(\mathbf{b})^{\frac{1}{2}(1-s)}\tfrac{1}{2}(1+\tau_3(\mathbf{b}))\tau_1(\mathbf{b})^{\frac{1}{2}(1-s')}, \tag{3.13}$$

$s, s'=\pm 1$, and let $\mathscr{C}_{\mathbf{\Lambda}}$ denote the set configurations

$$(\sigma,\tau):\mathbf{\Lambda}\times B(\mathbf{\Lambda})\to\{\pm 1\}\times\{\pm 1\}.$$

Then the partial isometries

$$E_{(\sigma,\tau)(\sigma',\tau')}=\prod_{\mathbf{x}\in\mathbf{\Lambda}}E^{\sigma}_{\sigma(\mathbf{x})\sigma'(\mathbf{x})}\prod_{\mathbf{b}\in B(\mathbf{\Lambda})}E^{\tau}_{\tau(\mathbf{b})\tau(\mathbf{b}')}, \tag{3.14}$$

(σ,τ), $(\sigma',\tau')\in\mathscr{C}_{\mathbf{\Lambda}}$, are a basis of matrix units for $\mathfrak{F}(\mathbf{\Lambda})$. The transfer matrix $T_{\mathbf{\Lambda}}$ has the following expansion:

$$T_{\mathbf{\Lambda}}=N_{\mathbf{\Lambda}}^{-1}\sum_{(\sigma,\tau),(\sigma',\tau')}T_{\mathbf{\Lambda}}(\sigma,\tau,\sigma',\tau')E_{(\sigma,\tau)(\sigma',\tau')}, \tag{3.15}$$

with $N_{\mathbf{\Lambda}}=(2\sinh 2\beta_h)^{|\mathbf{\Lambda}|/2}\,(2\sinh 2\beta)^{|B(\mathbf{\Lambda})|/2}$ and

$$T_{\mathbf{\Lambda}}(\sigma,\tau,\sigma',\tau')=\exp\left\{\beta_h\left(\sum_{\mathbf{b}\in B(\mathbf{\Lambda})}\tfrac{1}{2}(\delta\sigma(\mathbf{b})\tau(\mathbf{b})+\delta\sigma'(\mathbf{b})\tau'(b))+\sum_{\mathbf{x}\in\mathbf{\Lambda}}\sigma(\mathbf{x})\sigma'(x)\right)\right.$$
$$\left.+\beta_g\left(\sum_{\mathbf{p}\in P(\mathbf{\Lambda})}\tfrac{1}{2}(\delta\tau(\mathbf{p})+\delta\tau'(\mathbf{p}))+\sum_{\mathbf{b}\in B(\mathbf{\Lambda})}\tau(\mathbf{b})\tau'(\mathbf{b})\right)\right\}.$$

The expansion coefficients are strictly positive, thus from the theorem of Perron [21] the spectral projection $E_{\mathbf{\Lambda}}$ of $T_{\mathbf{\Lambda}}$ associated to the eigenvalue $\|T_{\mathbf{\Lambda}}\|$ is one-dimensional in $\mathfrak{F}(\mathbf{\Lambda})$, and $E_{\mathbf{\Lambda}}=\lim T^n_{\mathbf{\Lambda}}\|T_{\mathbf{\Lambda}}\|^{-n}$ has again an expansion with strictly positive coefficients. Thus $\omega_{\mathbf{\Lambda}}\restriction\mathfrak{F}(\mathbf{\Lambda})$ is unique and pure, and one may find it by the formula

$$\omega_{\mathbf{\Lambda}}(A)=\lim_{n\to\infty}(Z_{\mathbf{\Lambda},n})^{-1}\operatorname{Tr}_{\mathfrak{F}(\mathbf{\Lambda})}T^n_{\mathbf{\Lambda}}AT^n_{\mathbf{\Lambda}}E^{(0)}_{\mathbf{\Lambda}}, \tag{3.16}$$

with $Z_{\mathbf{\Lambda},n}=\operatorname{Tr}_{\mathfrak{F}(\mathbf{\Lambda})}T^{2n}_{\mathbf{\Lambda}}E^{(0)}_{\mathbf{\Lambda}}$, where $E^{(0)}_{\mathbf{\Lambda}}$ is any non-zero element of $\mathfrak{F}(\mathbf{\Lambda})$ with non-negative expansion coefficients. A convenient choice of $E^{(0)}_{\mathbf{\Lambda}}$ is

$$E^{(0)}_{\mathbf{\Lambda}}=\sum e_{\mathbf{\Lambda}}(\sigma,\tau)e_{\mathbf{\Lambda}}(\sigma',\tau')E_{(\sigma,\tau)(\sigma',\tau')}, \tag{3.17}$$

with $e_{\mathbf{\Lambda}}(\sigma,\tau)=\exp\left\{\tfrac{1}{2}\beta_h\sum_{\mathbf{b}\in B(\mathbf{\Lambda})}\delta\sigma(\mathbf{b})\tau(\mathbf{b})+\tfrac{1}{2}\beta_g\sum_{\mathbf{p}\in P(\mathbf{\Lambda})}\delta\tau(\mathbf{p})\right\}$. Equation (3.16) is the starting point for the transition to a classical statistical mechanics in $\mathbb{Z}^{d+1}$. Let $\Lambda_n=\{x=(x^0,\mathbf{x}),\,|x^0|\leqq n,\,\mathbf{x}\in\mathbf{\Lambda}\}$, and let $\mathscr{C}_{\Lambda_n}$ denote the set of configurations

$$(\sigma,\tau):\Lambda_n\times B(\Lambda_n)\to\{\pm 1\}\times\{\pm 1\},$$

with $\tau(b)=1$ for all vertical bonds (i.e. bonds of the form $b=\{(k,\mathbf{x}),\,(k+1,\mathbf{x})\}$). $Z_{\Lambda_n}=Z_{\mathbf{\Lambda},n}N^{2n}_{\mathbf{\Lambda}}$ is the partition function of the gauge invariant Ising model in the temporal gauge, with free [4] boundary conditions,

$$Z_{\Lambda_n}=\sum_{(\sigma,\tau)}e^{-H_{\Lambda_n}(\sigma,\tau)}, \tag{3.18}$$

4 This comes from the special choice of $E^{(0)}_{\mathbf{\Lambda}}$ in (3.17). Other choices lead to different boundary conditions in the 0-direction which turn out to be inconvenient for the construction of the thermodynamic limit by Griffiths inequalities and for the transition to the unitary gauge

where the Hamilton function is

$$H_{\Lambda_n}(\sigma,\tau) = -\left\{\beta_h \sum_{b\in B(\Lambda_n)} \delta\sigma(b)\tau(b) + \beta_g \sum_{p\in P(\Lambda_n)} \delta\tau(p)\right\}. \tag{3.19}$$

$\omega_{\mathbf{\Lambda}}(A)$ for $A\in\mathfrak{F}(\mathbf{\Lambda})$ can be computed by choosing functions $A^{\Lambda_n}:\mathscr{C}_{\Lambda_n}\to\mathbb{C}$ such that

$$\langle A^{\Lambda_n}\rangle_{\Lambda_n} = Z^{-1}_{\mathbf{\Lambda},n}\operatorname{Tr}_{\mathfrak{F}(\mathbf{\Lambda})} T^n_{\mathbf{\Lambda}} A T^n_{\mathbf{\Lambda}} E^{(0)}_{\mathbf{\Lambda}}, \tag{3.20}$$

where $\langle A^{\Lambda_n}\rangle_{\Lambda_n} = Z^{-1}_{\Lambda_n}\sum_{(\sigma,\tau)} A^{\Lambda_n}(\sigma,\tau)e^{-H_{\Lambda_n}(\sigma,\tau)}$ is the expectation value of A^{Λ_n} in the gauge invariant Ising model. A^{Λ_n} is not unique; a possible choice is the function $A^{\Lambda_n}_{(0)}$,

$$A^{\Lambda_n}_{(0)}(\sigma,\tau) = (AT_{\mathbf{\Lambda}})(\sigma^{(0)},\tau^{(0)},\sigma^{(1)},\tau^{(1)})/T_{\mathbf{\Lambda}}(\sigma^{(0)},\tau^{(0)},\sigma^{(1)},\tau^{(1)}), \tag{3.21}$$

where $(\sigma^{(k)},\tau^{(k)})\in\mathscr{C}_{\mathbf{\Lambda}}$ is the restriction of (σ,τ) to the $(x^0=k)$-hyperplane. We note the following rules for the choice of the classical functions A^{Λ_n}:

(i) For $m>n$ one may choose

$$A^{\Lambda_m}(\sigma,\tau) = A^{\Lambda_n}((\sigma,\tau)\upharpoonright\Lambda_n\times B(\Lambda_n)). \tag{3.22}$$

(ii) If $-n\leqq k_1<\ldots<k_l\leqq n-1$, $A^{(1)},\ldots,A^{(l)}\in\mathfrak{F}(\mathbf{\Lambda})$, and $A=\alpha_{ik_1}(A^{(1)})\ldots\alpha_{ik_l}(A^{(l)})$, a convenient choice for A^{Λ_n} is

$$A^{\Lambda_n}(\sigma,\tau) = A^{(1)\Lambda_n}_{(0)}(\sigma^{(k_1)},\tau^{(k_1)},\sigma^{(k_1+1)},\tau^{(k_1+1)}) \ldots A^{(l)\Lambda_n}_{(0)}(\sigma^{(k_l)},\tau^{(k_l)},\sigma^{(k_l+1)},\tau^{(k_l+1)}). \tag{3.23}$$

(iii) If $A\in\mathfrak{F}(\mathbf{\Lambda}_1)$, $B\in\mathfrak{F}(\mathbf{\Lambda}_2)$ and $\operatorname{dist}(\mathbf{\Lambda}_1,\mathbf{\Lambda}_2)\geqq 2$, $\mathbf{\Lambda}_1$, $\mathbf{\Lambda}_2\subset\mathbf{\Lambda}$, then

$$(A\,B)^{\Lambda_n}_{(0)} = A^{\Lambda_n}_{(0)}B^{\Lambda_n}_{(0)}. \tag{3.24}$$

(iv) For $A\in\mathfrak{F}(\mathbf{\Lambda})$ and $\mathbf{\Lambda}'\supset\hat{\mathbf{\Lambda}}$ with $\hat{\mathbf{\Lambda}}$ defined after Eq. (3.2)

$$A^{\Lambda'_n}_{(0)}(\sigma,\tau) = A^{\hat{\Lambda}_n}_{(0)}((\sigma,\tau)\upharpoonright\hat{\Lambda}_n\times B(\hat{\Lambda}_n)). \tag{3.25}$$

Equation (3.16) now becomes

$$\omega_{\mathbf{\Lambda}}(A) = \lim_{n\to\infty}\langle A^{\Lambda_n}\rangle_{\Lambda_n}. \tag{3.26}$$

The convergence of $\omega_{\mathbf{\Lambda}}$ for $\mathbf{\Lambda}\nearrow\mathbb{Z}^d$ is a consequence of Griffiths inequalities for the gauge invariant Ising model [22],

$$\omega_0(A) = \lim_{\mathbf{\Lambda}\nearrow\mathbb{Z}^d}\omega_{\mathbf{\Lambda}}(A), \qquad A\in\mathfrak{F}. \tag{3.27}$$

In the next section we shall use polymer expansions to compute ω_0. It is convenient to work in the unitary gauge. This amounts to the transformation of variables

$$(\sigma,\tau)\to(\varrho,u),$$

$\varrho:\mathbf{\Lambda}\to\{\pm1\}$, $u:B(\Lambda_n)\to\{\pm1\}$, $\varrho(\mathbf{x})=\sigma(0,\mathbf{x})$, $u(b)=\delta\sigma(b)\tau(b)$. Since the Hamilton function H_{Λ_n} is independent of ϱ we can replace the function $A^{\Lambda_n}(\varrho,u)$ by its mean $A^{\Lambda_n}(u)$ over ϱ. Then

$$\langle A^{\Lambda_n}\rangle_{\Lambda_n} = (Z_{\Lambda_n})^{-1}2^{|\mathbf{\Lambda}|}\sum_{u:B(\Lambda_n)\to\{\pm1\}} A^{\Lambda_n}(u) \cdot\exp\left[\beta_h\sum_b u(b) + \beta_g\sum_p \delta u(p)\right]. \tag{3.28}$$

4. The Gauge Invariant Ising Model

In a neighbourhood of $\beta_g = \infty$, $\beta_h = 0$ the gauge invariant Ising model admits a convergent expansion which has first been discussed by Marra and Miracle-Solé [10]. We want to use this expansion for a construction of charged states for the $\mathbb{Z}_2$ gauge theory. As a first step we review its use for the construction of the Gibbs state of the gauge invariant Ising model which gives the vacuum state of the $\mathbb{Z}_2$ gauge theory.

Let Λ be a box in $\mathbb{Z}^{d+1}$, and let

$$\chi_L(u) = \prod_{b \in L} u(b) \tag{4.1}$$

for $L \subset B(\Lambda)$ and the configuration $u: B(\Lambda) \to \{\pm 1\}$. Each function $A(u)$ is a linear combination of the characters χ_L. Given a gauge field configuration u, the corresponding field strength configuration $f = \delta u$ is uniquely determined by its support,

$$\operatorname{supp} f = \{p \in P(\Lambda), f(p) = -1\}, \tag{4.2}$$

and a configuration $f: P(\Lambda) \to \{\pm 1\}$ is a field strength configuration if and only if $\delta f(c) = 1$ for each elementary cube c in Λ. Thus a set of plaquettes $P \subset P(\Lambda)$ is the support of some field strength configuration if and only if its coboundary, i.e. the set of cubes in Λ with an odd number of faces in P, is empty (P is coclosed). Let $\mathscr{P}_\Lambda$ denote the set of all coclosed $P \subset P(\Lambda)$. Then, by a character expansion of $e^{\beta_h \Sigma u(b)}$,

$$\langle \chi_L \rangle_\Lambda = Z_\Lambda^{-1} \sum_{\substack{P \in \mathscr{P}_\Lambda \\ M \subset B(\Lambda)}} e^{-2\beta_g |P|} (\operatorname{th} \beta_h)^{|M|} (P, L \triangle M), \tag{4.3}$$

with $(P, L\triangle M) = 2^{1-|\Lambda|} \sum_{\substack{u: B(\Lambda) \to \{\pm 1\} \\ \operatorname{supp} \delta u = P}} \chi_{L \triangle M}(u)$.

$L \triangle M$ denotes the symmetric difference $(L \cup M) \backslash (L \cap M)$ of L and M. Z_Λ is fixed by the condition $\langle \chi_\emptyset \rangle_\Lambda = 1$. We have

$$(P, L\triangle M) = \begin{cases} 0, & \text{if } L\triangle M \text{ is not closed,} \\ \chi_{L\triangle M}(u) & \text{for all } u \text{ with } \operatorname{supp} \delta u = P, \text{ if } L \triangle M \text{ is closed.} \end{cases} \tag{4.4}$$

We may therefore restrict the summation in (4.3) to those M for which $L \triangle M$ is closed. $(P, L\triangle M)$ is then -1 if the winding number of P around $L \triangle M$ is odd, and $+1$ if it is even. Let ∂N for $N \subset B(\Lambda)$ denote the boundary of N, i.e. the set of points $x \in \Lambda$ which are endpoints of an odd number of bonds in N. We may decompose M in a unique way into connected components, i.e. the equivalence classes of the equivalence relation $\underset{\subset}{\simeq}$ in M generated by the relation

$$b \cap b' \neq \emptyset. \tag{4.5}$$

Some of the components of M have a nonvoid boundary, the other components are closed. Let $\operatorname{Conn}_\Lambda(L)$ denote the set of all $M \subset B(\Lambda)$ with $\partial M = \partial L$ such that no component of M is closed, and let $\operatorname{Disc}_\Lambda(M)$ denote the set of all $N \subset B(\Lambda)$ with $\partial N = \emptyset$ which are disconnected to M, i.e. $b \cap b' = \emptyset$ for all $b \in M$, $b' \in N$. Using the fact that $(P, L \triangle (M \cup N)) = (P, L\triangle M)(P, N)$ for $M \in \operatorname{Conn}_\Lambda(L)$ and

$N \in \mathrm{Disc}_\Lambda(M)$, one arrives at

$$\langle \chi_L \rangle_\Lambda = \sum_{M \in \mathrm{Conn}_\Lambda(L)} (\mathrm{th}\,\beta_h)^{|M|} \sigma_\Lambda(L, M), \tag{4.6}$$

with $\sigma_\Lambda(L, M) = Z_\Lambda^{-1} \sum_{\substack{P \in \mathscr{P}_\Lambda \\ N \in \mathrm{Disc}_\Lambda(M)}} (P, L \triangle M) \mu(P, N)$,

$$\mu(P, N) = (P, N) e^{-2\beta_g |P|} (\mathrm{th}\,\beta_h)^{|N|}.$$

For the weights $\sigma_\Lambda(L, M)$ we have the following uniform estimates. [By $\langle \ \rangle_\Lambda^0$ we shall denote the expectation value in the pure gauge theory ($\beta_h = 0$).]

Proposition 4.1. *Let $L, M \subset B(\Lambda)$ with $\partial L = \partial M$. Then*

(i) $$\sigma_\Lambda(L, M) \leqq \langle \chi_{L \triangle M} \rangle,$$

(ii) $$\sigma_\Lambda(L, M) \geqq (1 + \mathrm{th}\,\beta_h)^{-n(M)} \langle \chi_{L \triangle M} \rangle_\Lambda^0,$$

$n(M) = |\{b \in B(\Lambda), b \gtrsim M\}| \leqq (4d+3)|M|$,

(iii) $$\frac{\sigma_\Lambda(L \triangle L', M \triangle M')}{\sigma_\Lambda(L', M')} \leqq \frac{1}{\langle \chi_{L \triangle M} \rangle_\Lambda^0}, \quad L', M' \subset B(\Lambda), \quad \partial L' = \partial M', M' \cap M = \emptyset.$$

Proof. (i) The definition of $\sigma_\Lambda(L, M)$ in (4.6) may be rewritten in the following way

$$\sigma_\Lambda(L, M) = (Z'_\Lambda)^{-1} \sum_{N \in \mathrm{Disc}_\Lambda(M)} (\mathrm{th}\,\beta_h)^{|N|} \langle \chi_{L \triangle M \triangle N} \rangle_\Lambda^0, \tag{$*$}$$

Z'_Λ being fixed by the condition $\sigma_\Lambda(\emptyset, \emptyset) = 1$. From the first Griffiths inequality

$$\langle \chi_{L \triangle M \triangle N} \rangle_\Lambda^0 \geqq 0,$$

hence, as $\mathrm{Disc}_\Lambda(M) \subset \mathrm{Disc}_\Lambda(\emptyset)$,

$$\sigma_\Lambda(L, M) \leqq \sigma_\Lambda(L \triangle M, \emptyset) = \langle \chi_{L \triangle M} \rangle_\Lambda.$$

(ii) From the second Griffiths inequality

$$\langle \chi_{L \triangle M \triangle N} \rangle_\Lambda^0 \geqq \langle \chi_{L \triangle M} \rangle_\Lambda^0 \langle \chi_N \rangle_\Lambda^0,$$

hence from relation ($*$)

$$\sigma_\Lambda(L, M) \geqq \langle \chi_{L \triangle M} \rangle_\Lambda^0 (Z'_\Lambda)^{-1} \sum_{N \in \mathrm{Disc}_\Lambda(M)} (\mathrm{th}\,\beta_h)^{|N|} \langle \chi_N \rangle_\Lambda^0.$$

On the other hand,

$$(Z'_\Lambda)^{-1} \sum_{N \in \mathrm{Disc}_\Lambda(M)} (\mathrm{th}\,\beta_h)^{|N|} \langle \chi_N \rangle_\Lambda^0 = \left\langle \prod_{b \gtrsim M} (1 + \mathrm{th}\,\beta_h u(b))^{-1} \right\rangle_\Lambda,$$

which is bounded from below by $\inf_u \prod_{b \gtrsim M} (1 + \mathrm{th}\,\beta_h u(b))^{-1} = (1 + \mathrm{th}\,\beta_h)^{-n(M)}$.

(iii) Since $M \cap M' = \emptyset$, $\mathrm{Disc}_\Lambda(M \triangle M') \subset \mathrm{Disc}_\Lambda(M)$; from the Griffiths inequalities

$$0 \leqq \langle \chi_{L \triangle L' \triangle M \triangle M' \triangle N} \rangle_\Lambda^0 \leqq \frac{\langle \chi_{L \triangle M \triangle N} \rangle_\Lambda^0}{\langle \chi_{L' \triangle M'} \rangle_\Lambda^0}.$$

Inserting these relations in ($*$) proves statement (iii). q.e.d.

The thermodynamic limit

$$\sigma_\Lambda(L,M)=\Big\langle \prod_{b\supsetneq M}(1+\mathrm{th}\beta_h\chi_{\{b\}})^{-1}\chi_{L\triangle M}\Big\rangle_\Lambda\to\sigma(L,M) \tag{4.7}$$

exists always. Since there are at most $\frac{|\partial L|}{4}\left(\frac{m-1}{2}\right)^{|\partial L|/2-1}\times(2d+1)^m$ sets $M\in\mathrm{Conn}_\Lambda(L)$ with $|M|=m$ (Appendix A.3), the expansion

$$\langle\chi_L\rangle=\sum_{M\in\mathrm{Conn}(L)}(\mathrm{th}\beta_h)^{|M|}\sigma(L,M) \tag{4.8}$$

is convergent for $\mathrm{th}\beta_h<(2d+1)^{-1}$. [$\mathrm{Conn}(L)=\{M\subset B(\mathbb{Z}^{d+1}),\ M\in\mathrm{Conn}_\Lambda(L)$ for Λ sufficiently large$\}$.]

For a more detailed estimate of $\sigma(L,M)$ we consider $\sigma_\Lambda(L,M)$ as the ratio of partition functions of certain polymer models. Following Marra and Miracle-Solé [10] we associate to each pair $(P,N)\in\mathscr{P}_\Lambda\times\mathrm{Disc}_\Lambda(\emptyset)$ a graph γ_{PN} such that the connected components of N and the coconnected[5] components of P are the vertices of the graph, and two vertices are connected by a line if one is a component P_1 of P, the other a component N_1 of N and $(P_1,N_1)=-1$. The class of all these graphs is denoted by $\mathscr{G}_\Lambda$, and if $\gamma=\gamma_{PN}$ we set $P_\gamma=P$, $N_\gamma=N$. Γ_γ is the set of connected components of γ[6], and γ, γ' are called compatible, $\gamma\sim\gamma'$, if there is a $\gamma''\in\mathscr{G}_\Lambda$ with Γ_γ, $\Gamma_{\gamma'}\subset\Gamma_{\gamma''}$ and $\Gamma_\gamma\cap\Gamma_{\gamma'}=\emptyset$.

Now we assign an activity $\mu(\gamma)$ to each graph $\gamma\in\mathscr{G}_\Lambda$

$$\mu(\gamma)=(P_\gamma,N_\gamma)e^{-2\beta_g|P_\gamma|}(\mathrm{th}\beta_h)^{|N_\gamma|}\,; \tag{4.9}$$

$\mu(\gamma)$ is multiplicative,

$$\mu(\gamma)=\prod_{\gamma'\in\Gamma(\gamma)}\mu(\gamma')\,. \tag{4.10}$$

These activities define the following polymer model: if A is a function on $\mathscr{G}_\Lambda$, its expectation value in this model is

$$\langle A\rangle_{\mu,\Lambda}=\sum_{\gamma\in\mathscr{G}_\Lambda}A(\gamma)\mu(\gamma)\Big/\sum_{\gamma\in\mathscr{G}_\Lambda}\mu(\gamma)\,. \tag{4.11}$$

For functions A which are multiplicative in the sense of (4.10) the expectation value (4.11) can be considered as the ratio of partition functions; one has the following formula [23]

$$\ln\langle A\rangle_{\mu,\Lambda}=\sum_{\Gamma:\mathscr{G}^c_\Lambda\to\mathbb{Z}_+}c_\Gamma(A^\Gamma-1)\mu^\Gamma\,, \tag{4.12}$$

where $\mathscr{G}^c_\Lambda$ is the set of connected graphs $\gamma\in\mathscr{G}_\Lambda$ and where we used the convention for multi-indices

$$v^\Gamma=\prod_{\gamma\in\mathrm{supp}\,\Gamma}v(\gamma)^{\Gamma(\gamma)}\,,\qquad v:\mathscr{G}^c_\Lambda\to\mathbb{C}\,. \tag{4.13}$$

The c_Γ are purely combinatorial coefficients which are independent of μ,

$$c_\Gamma=\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\mathscr{N}_n(\Gamma)\,, \tag{4.14}$$

5 Two plaquettes p and p' are coconnected if they are faces of the same cube

6 We shall often identify the set Γ_γ with its characteristic function

where $\mathcal{N}_n(\Gamma)$ is the number of possibilities to write Γ in the form $\Gamma=\Gamma_{\gamma_1}+\ldots+\Gamma_{\gamma_n}$ with $\gamma_i\in\mathcal{G}_\Lambda$, $\gamma_i\neq\emptyset$, $i=1,\ldots,n$. (For details see [2].) They vanish if Γ can be decomposed into two parts Γ_1 and Γ_2 such that $\gamma_1\sim\gamma_2$ for all pairs $(\gamma_1, \gamma_2)\in\operatorname{supp}\Gamma_1\times\operatorname{supp}\Gamma_2$.

The thermodynamic limit can be controlled using the fact that ($\beta=\min(\beta_h^*, \beta_g)$)

$$\sum_{\substack{\gamma\in\mathcal{G}_\Lambda^c \\ \gamma\nsim\gamma'}} |\mu(\gamma)|\leqq F_0(2\beta)\,|\gamma'|,$$

where F_0 is Λ-independent, monotonically decreasing and convex, and $|\gamma'|=|N_{\gamma'}|+|P_{\gamma'}|$ (Appendix A.3). In Appendix A.1 it is shown that this leads to the estimate ($\beta\geqq\beta^c$)

$$\sum_{\Gamma\nsim\gamma} |c_\Gamma|\,|\mu^\Gamma|\leqq F_1(2\beta)\,|\gamma| \tag{4.15}$$

($\Gamma\nsim\gamma$ means that $\gamma'\nsim\gamma$ for some $\gamma'\in\operatorname{supp}\Gamma$). β^c and F_1 are defined in terms of F_0 (A.9)–(A.12).

For $\sigma(L,M)$ the polymer expansion yields

$$\ln\sigma(L,M)=\sum_\Gamma c_\Gamma(a_{L,M}^\Gamma-1)\mu^\Gamma \tag{4.16}$$

with

$$a_{L,M}(\gamma)=\begin{cases}0, & \text{if } N_\gamma \text{ is connected with } M\\ (P_\gamma, L\triangle M) & \text{otherwise}.\end{cases}$$

If $M=\emptyset$ we write $a_{L,\emptyset}=a_L$. The expansion (4.16) converges if $\beta_g\geqq\beta^c=\beta_g^w$ and $\beta_h\leqq-\frac{1}{2}\ln\operatorname{th}\beta^c=\beta_h^w$.

Bounds on β_g^w and β_h^w as well as a more detailed determination of the convergence region are given in Appendix A.4.

We shall use several times the following estimate on the contribution of large Γ which is an immediate consequence of (4.15) ($\|\Gamma\|=\sum\Gamma(\gamma)\,|\gamma|$):

$$\sum_{\substack{\Gamma\nsim\gamma \\ \|\Gamma\|\geqq n}} |c_\Gamma|\,|\mu^\Gamma|\leqq e^{-2(\beta-\beta^c)n}F_1(2\beta^c)\,|\gamma|. \tag{4.17}$$

Inequality (4.17) implies clustering of expectation values which means that the vacuum vector Ω of the $\mathbb{Z}_2$ gauge theory is the unique (up to a phase) ground state of T_0. Unfortunately, one cannot conclude that the vacuum representation π_0 is irreducible [25], since the invariance of the weak closure of $\pi_0(\mathfrak{F})$ under the time evolution is not known.

5. Construction of Charged States

The idea for the construction of a charged state is simple. One creates a charge at some point **x** together with a compensating charge and transports the compensating charge to infinity. In a gauge theory the charges are connected by electric flux lines, so one has to arrange these flux lines in such a way that the limit state has finite energy.

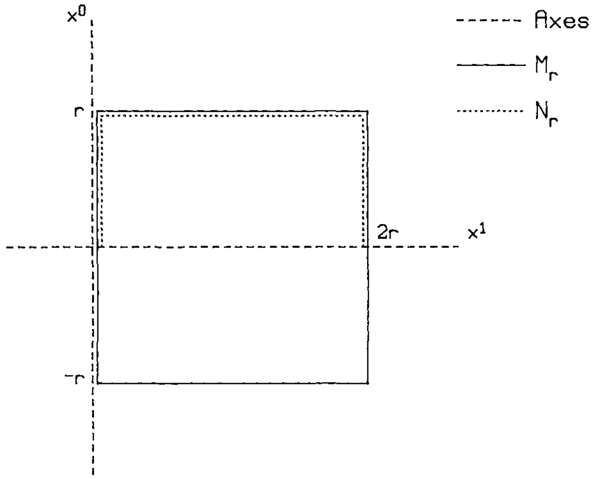

Fig. 5.1. The lines M_r and N_r in the $(1-0)$-plane

Let $\mathbf{x}_r = (2r, 0, \ldots, 0) \in \mathbb{Z}^d$, $r \in \mathbb{N}$, and let $\mathbf{L}_r$ denote the path along the 1-axis from the origin to $\mathbf{x}_r$. Let

$$\begin{aligned} \tau_3(\mathbf{L}_r) &= \prod_{\mathbf{b} \in \mathbf{L}_r} \tau_3(\mathbf{b}), \\ F_r &= \sigma_3(\mathbf{0})\sigma_3(\mathbf{x}_r)\alpha_{ir}(\tau_3(\mathbf{L}_r)). \end{aligned} \tag{5.1}$$

F_r is gauge invariant. Consider the states ω_r,[7]

$$\omega_r(A) = (F_r\Omega, AF_r\Omega)\,\|F_r\Omega\|^{-2}, \qquad A \in \mathfrak{F}. \tag{5.2}$$

Let $\varrho_{\mathbf{x}}$, $\mathbf{x} \in \mathbb{Z}^d$ denote the automorphism of $\mathfrak{F}$ which is implemented by $\sigma_3(\mathbf{x})$,

$$\varrho_{\mathbf{x}}(A) = \sigma_3(\mathbf{x})A\sigma_3(\mathbf{x}), \qquad A \in \mathfrak{F}. \tag{5.3}$$

For $A \in \mathfrak{F}_0$ and r large enough

$$\omega_r(A) = \frac{(\Omega, \tau_3(\mathbf{L}_r)T_0^r\varrho_{\mathbf{0}}(A)T_0^r\tau_3(\mathbf{L}_r)\Omega)}{(\Omega, \tau_3(\mathbf{L}_r)T_0^{2r}\tau_3(\mathbf{L}_r)\Omega)}. \tag{5.4}$$

Let $A^\varrho(u)$ denote the gauge invariant part of a classical function localized in the slice $-r < x^0 \leqq r$ which corresponds to the operator $\varrho_0(A)$. Then

$$\omega_r(A) = \langle A^\varrho \chi_{M_r} \rangle \langle \chi_{M_r} \rangle^{-1}, \tag{5.5}$$

where M_r is the square in the $(0-1)$-plane with edges $(r, \mathbf{0})$, $(r, \mathbf{x}_r)$, $(-r, \mathbf{x}_r)$ and $(-r, \mathbf{0})$ (Fig. 5.1). From the results of Sect. 4, for $L \subset B(\mathbb{Z}^{d+1})$, $|L| < \infty$

$$\langle \chi_L \chi_{M_r} \rangle \langle \chi_{M_r} \rangle^{-1} = \sum_{M \in \mathrm{Conn}(L)} (\mathrm{th}\,\beta_h)^{|M|} \frac{\sigma(L \triangle M_r, M)}{\sigma(M_r, \emptyset)}. \tag{5.6}$$

7 In the following we identify $\mathfrak{F}$ and $\pi_0(\mathfrak{F})$ by dropping the symbol π_0

From Proposition 4.1 (iii)

$$\frac{\sigma(L\triangle M_r, M)}{\sigma(M_r, \emptyset)} \leqq (\langle \chi_{M\triangle L}\rangle^0)^{-1},$$

in the pure gauge theory one has a perimeter law,

$$\langle \chi_{L\triangle M}\rangle^0 \geqq e^{-\alpha(\beta_g)|L\triangle M|} \tag{5.7}$$

with $\alpha(\beta_g)\to 0$ for $\beta_g\to\infty$ [8]. [For an estimate of $\alpha(\beta_g)$ see Appendix A.2.] Hence the expansion (5.6) converges uniformly in r provided $\mathrm{th}\beta_h<(2d+1)^{-1}e^{-\alpha(\beta_g)}$. On the other hand, for $\beta_g>\beta_g^w$, $\beta_h<\beta_h^w$ from (4.16)

$$\ln\frac{\sigma(L\triangle M_r, M)}{\sigma(M_r, \emptyset)} = \sum_\Gamma c_\Gamma(a^\Gamma_{L\triangle M_r, M} - a^\Gamma_{M_r})\mu^\Gamma$$

$$= \sum_\Gamma c_\Gamma(a^\Gamma_{L,M} - 1)a^\Gamma_{M_r}\mu^\Gamma. \tag{5.8}$$

The convergence of the right-hand side of (5.8) follows now simply from the fact that $|a_{M_r}(\gamma)|\leqq 1$ and $a_{M_r}(\gamma)\to a_{M_\infty}(\gamma)$ for each γ, where M_∞ is the 0-axis. Hence we arrive at the following theorem.

Theorem 5.1. *For $\beta_g>\beta_g^w$, $\beta_h<\beta_h^w$ there exists a state ω on $\mathfrak{F}$ such that for all $A\in\mathfrak{F}$*

$$\lim_{r\to\infty}\omega_r(A)=\omega(A).$$

The interpretation of ω as a charged state is supported by the following fact. Let $Q_\mathbf{\Lambda} = \prod_{\mathbf{x}\in\mathbf{\Lambda}} \delta^*\tau_1(\mathbf{x})$ be the charge operator associated to the region $\mathbf{\Lambda}\subset\mathbb{Z}^d$, and let $\partial^*\mathbf{\Lambda}$ denote the set of bonds with exactly one endpoint in $\mathbf{\Lambda}$ (the coboundary of $\mathbf{\Lambda}$). Then (Gauß' law)

$$Q_\mathbf{\Lambda} = \prod_{\mathbf{b}\in\partial^*\mathbf{\Lambda}} \tau_1(\mathbf{b}), \tag{5.9}$$

hence $Q_\mathbf{\Lambda}$ can be measured at the boundary of $\mathbf{\Lambda}$. Thus there are no local fields in the interior of $\mathbf{\Lambda}$ which create a charge. Actually, for every $A\in\mathfrak{F}_0$ with $A\Omega\neq 0$

$$\frac{(A\Omega, Q_\mathbf{\Lambda}A\Omega)}{\|A\Omega\|^2(\Omega, Q_\mathbf{\Lambda}\Omega)}\to 1, \tag{5.10}$$

if $\mathbf{\Lambda}$ tends to $\mathbb{Z}^d$ with $|\partial^*\mathbf{\Lambda}|<\mathrm{const}\times\mathrm{dist}(\partial^*\mathbf{\Lambda},\mathbf{0})^k$ for some $k\in\mathbb{N}$. But for the state ω which arose from a nonlocal operation on the vacuum we have

Theorem 5.2. *Let $\mathbf{\Lambda}\nearrow\mathbb{Z}^d$ such that $|\partial^*\mathbf{\Lambda}|<\mathrm{const}\,\mathrm{dist}(\partial^*\mathbf{\Lambda},\mathbf{0})^k$ for some $k\in\mathbb{N}$. Then for $\beta_g>\beta_g^w$, $\beta_h<\beta_h^w$*

$$\lim_\mathbf{\Lambda}\frac{\omega(Q_\mathbf{\Lambda})}{\omega_0(Q_\mathbf{\Lambda})}=-1.$$

Proof. A classical function corresponding to the operator $Q_\mathbf{\Lambda} = \varrho_0(Q_\mathbf{\Lambda})$ is

$$Q^\varrho_\mathbf{\Lambda}(u) = \prod_{p\in P_\mathbf{\Lambda}} e^{-2\beta_g\delta u(p)} \prod_{b\in B_\mathbf{\Lambda}} e^{-\beta_h u(b)},$$

where $P_{\mathbf{\Lambda}}$ is the (coclosed) set of plaquettes in $\mathbb{Z}^{d+1}$, which are spanned by $\{0\}\times\mathbf{b}$ and $\{1\}\times\mathbf{b}$, $\mathbf{b}\in\partial^*\mathbf{\Lambda}$, and $B_{\mathbf{\Lambda}}$ is the set of bonds $\{0\}\times\mathbf{b}$, $\mathbf{b}\in\partial^*\mathbf{\Lambda}$. The computation of the expectation value of $Q_{\mathbf{\Lambda}}$ in the vacuum can be replaced by the computation of the expectation value of the following function on polymers,

$$Q_{\mathbf{\Lambda}}(\gamma)=(\cosh\beta_h)^{-|B_{\mathbf{\Lambda}}|}b_{\mathbf{\Lambda}}(\gamma),$$

$$b_{\mathbf{\Lambda}}(\gamma)=\begin{cases}0, & N_\gamma\cap B_{\mathbf{\Lambda}}\neq\emptyset\\ (P_{\mathbf{\Lambda}},N_\gamma) & \text{otherwise},\end{cases}$$

hence $\omega_0(\mathrm{Q}_{\mathbf{\Lambda}})=(\cosh\beta_h)^{-|B_{\mathbf{\Lambda}}|}\exp\left\{\sum_\Gamma c_\Gamma(b_{\mathbf{\Lambda}}^\Gamma-1)\mu^\Gamma\right\}$. In the charged state ω one obtains instead

$$\tilde{Q}_{\mathbf{\Lambda}}(\gamma)=(P_{\mathbf{\Lambda}},M_\infty)Q_{\mathbf{\Lambda}}(\gamma),$$

and since $(P_{\mathbf{\Lambda}},M_\infty)=-1$ for $\mathbf{0}\in\mathbf{\Lambda}$,

$$\omega(Q_{\mathbf{\Lambda}})=-(\cosh\beta_h)^{-|B_{\mathbf{\Lambda}}|}\exp\left\{\sum_\Gamma c_\Gamma(b_{\mathbf{\Lambda}}^\Gamma-1)a_{M_\infty}^\Gamma\mu^\Gamma\right\}.$$

It remains to show that

$$\sum_\Gamma c_\Gamma(1-b_{\mathbf{\Lambda}}^\Gamma)(1-a_{M_\infty}^\Gamma)\mu^\Gamma$$

vanishes if $\mathbf{\Lambda}$ tends to $\mathbb{Z}^d$ in the mentioned way. But this follows from the fact that only those Γ can contribute to the sum with $\gamma_1,\gamma_2\in\operatorname{supp}\Gamma$, $(P_{\gamma_1},M_\infty)=-1$, $b_{\mathbf{\Lambda}}(\gamma_2)\neq1$; their length is at least $\|\Gamma\|=\operatorname{dist}(\mathbf{0},\partial^*\mathbf{\Lambda})$. Hence from (4.17), for some $\alpha>0$

$$\left|\ln\left\{\frac{-\omega(Q_{\mathbf{\Lambda}})}{\omega_0(Q_{\mathbf{\Lambda}})}\right\}\right|\leqq\text{const}\,|\partial^*\mathbf{\Lambda}|\,e^{-\alpha\operatorname{dist}(\mathbf{0},\partial^*\mathbf{\Lambda})}\quad\text{q.e.d.}$$

Unfortunately, this result does not exclude completely the possibility that ω is a state in the vacuum sector. Since $\omega_0(Q_{\mathbf{\Lambda}})\sim e^{-\alpha|\partial^*\mathbf{\Lambda}|}$ for some $\alpha>0$, the convergence $Q_{\mathbf{\Lambda}}\omega_0(Q_{\mathbf{\Lambda}})^{-1}\to1$ in the sense of matrix elements can be shown only on the dense set $\mathfrak{F}_0\Omega$, so the theorem would be compatible with the existence of a vector $\Phi\in\mathscr{H}_0$, $\Phi\notin\mathfrak{F}_0\Omega$, which induces the state ω.

A further indication that this hypothetical vector does not exist and ω is really a state in a new sector disjoint from the vacuum sector is the observation that the sequence $F_r\Omega\|F_r\Omega\|^{-1}$ becomes orthogonal to each vector in $\mathscr{H}_0$ in the limit $r\to\infty$. Actually, by the following proposition, it is sufficient to check whether it becomes orthogonal to the vacuum vector Ω.

Proposition 5.3. *Assume that* $(\Omega,F_r\Omega)\|F_r\Omega\|^{-1}\to0$. *Then* $(\Psi,F_r\Omega)\,\|F_r\Omega\|^{-1}\to0$ *for every* $\Psi\in\mathscr{H}_0$.

Proof. Since the sequence $F_r\Omega\|F_r\Omega\|^{-1}$ is bounded, it is sufficient to show the convergence $(\Psi,F_r\Omega)\|F_r\Omega\|^{-1}\to0$ for $\Psi\in\mathfrak{F}_0\Omega$. Then, denoting the part of M_r above the $x^0=0$ hyperplane by N_r (Fig. 5.1), we obtain

$$(\Psi,F_r\Omega)=\sum_{i=1}^n\lambda_i\langle\chi_{L_i}\chi_{N_r}\rangle$$

for some $n\in\mathbb{N}$, $\lambda_i\in\mathbb{C}$, $L_i\subset B(\mathbb{Z}^{d+1})$, $|L_i|<\infty$, $i=1,\dots,n$ and r sufficiently large. By Griffiths inequalities

$$\langle\chi_{L_i}\chi_{N_r}\rangle\leqq\langle\chi_{N_r}\rangle\langle\chi_{L_i}\rangle^{-1}\leqq\langle\chi_{N_r}\rangle(\operatorname{th}\beta_h)^{-|L_i|}.$$

Since $\langle\chi_{N_r}\rangle=(\Omega,F_r\Omega)$, the proposition follows. q.e.d.

Theorem 5.4. *If the perimeter law* (5.7) *holds for the pure gauge theory and if* $(2d+1)\operatorname{th}\beta_h e^{2\alpha(\beta_g)}<1$, *then* $F_r\Omega\|F_r\Omega\|^{-1}\underset{w}{\rightarrow}0$.

Proof. From Proposition 5.3 it is sufficient to show that $(\Omega,F_r\Omega)\|F_r\Omega\|^{-1}\rightarrow 0$. Now from (4.8), Proposition 4.1 and A.3.1 we get, for $(2d+1)\operatorname{th}\beta_h<1$,

$$\begin{aligned}(\Omega,F_r\Omega)=\langle\chi_{N_r}\rangle&\leqq\sum_{M\in\operatorname{Conn}(N_r)}(\operatorname{th}\beta_h)^{|M|}\\&\leqq\sum_{m=r}^{\infty}[(2d+1)\operatorname{th}\beta_h]^{2m}=[(2d+1)\operatorname{th}\beta_h]^{2r}(1-(2d+1)^2\operatorname{th}^2\beta_h)^{-1}.\end{aligned}$$

From (5.7)

$$\|F_r\Omega\|=\langle\chi_{M_r}\rangle^{1/2}\geqq e^{-\alpha(\beta_g)4r},$$

thus

$$\begin{aligned}(\Omega,F_r\Omega)\|F_r\Omega\|^{-1}\leqq&[(2d+1)\operatorname{th}\beta_h e^{2\alpha(\beta_g)}]^{2r}\\&\cdot(1-(2d+1)^2\operatorname{th}^2\beta_h)^{-1}.\quad\text{q.e.d.}\end{aligned}$$

The weak convergence of $F_r\Omega\|F_r\Omega\|^{-1}$ to 0 is a necessary condition for the limit points ω of the sequence ω_r to be disjoint from the vacuum sector. It is interesting that one can prove this weak convergence not only in Region II_c of the phase diagram but in a larger part of Region II. To get a feeling of whether the weak limit points of the sequence $F_r\Omega\|F_r\Omega\|^{-1}$ indicate the presence or absence of charges we analyse the convergence properties in Region I_c.

If β_g is sufficiently small one can map the model onto the following polymer model. Graphs are sets of plaquettes P, the plaquettes are the vertices of the graph and two plaquettes p,p' are joined by a line if there is a bond $b\subset\partial p\cap\partial p'$. The activity is

$$\mu(P)=(\operatorname{th}\beta_g)^{|P|}(\operatorname{th}\beta_h)^{|\partial P|}.\tag{5.11}$$

The computation of $\langle\chi_L\rangle$ amounts to the mean over the polymers of the function

$$(\operatorname{th}\beta_h)^{|L|}A_L(P)=(\operatorname{th}\beta_h)^{|L\triangle\partial P|-|\partial P|}=(\operatorname{th}\beta_h)^{|L|-2|L\cap\partial P|}.\tag{5.12}$$

If β_h is sufficiently large one defines graphs as sets M of lattice bonds with activities

$$\mu(M)=e^{-2\beta_h|M|}e^{-2\beta_g|\partial^*M|},\tag{5.13}$$

where two bonds $b,b'\in M$ are connected by a line if there exists a plaquette p with $b,b'\subset\partial p$. The expectation value of χ_L corresponds to the expectation value of the function on polymers

$$B_L(M)=(-1)^{L\cap M};\tag{5.14}$$

β_g^s and β_h^s, respectively, denote the border of the convergence regions of these expansions.

It will turn out that for $\beta_g < \beta_g^s$ or $\beta_h > \beta_h^s$ there exist eigenvectors of T_0 which have a single external charge. The following proposition provides a general criterion for the existence of such eigenvectors.

Proposition 5.5. *Let t_0 be the norm of the restriction of T_0 to the subspace of $\mathscr{H}_0$ with external charge configuration q, $\operatorname{supp} q = \{\mathbf{0}\}$. t_0 is an eigenvalue of T_0 if and only if*

$$\lim_{n\to\infty} \frac{(\sigma_3(\mathbf{0})\Omega, T_0^n \sigma_3(\mathbf{0})\Omega)^2}{(\sigma_3(\mathbf{0})\Omega, T_0^{2n}\sigma_3(\mathbf{0})\Omega)} \neq 0 .$$

Proof. Let

$$(\sigma_3(\mathbf{0})\Omega, T_0^n\sigma_3(\mathbf{0})\Omega)^2 (\sigma_3(\mathbf{0})\Omega, T_0^{2n}\sigma_3(\mathbf{0})\Omega)^{-1} \to c \neq 0 .$$

Then $\sigma_3(\mathbf{0})\Omega$ has a nonvanishing scalar product with an eigenvector Φ_0 of T_0; one may choose

$$\Phi_0 = \lim_{n\to\infty} T_0^n \sigma_3(\mathbf{0})\Omega \, \| T_0^n\sigma_3(\mathbf{0})\Omega\|^{-1} .$$

Actually, the corresponding eigenvalue is t_0. This can be seen as follows:

$$\begin{aligned} t_0 &= \sup_{A,B\in\mathfrak{A}} \lim_{n\to\infty} (A\sigma_3(\mathbf{0})\Omega, T_0^n B\sigma_3(\mathbf{0})\Omega)^{1/n} \\ &= \sup_{\substack{L,M\subset B(\mathbb{Z}^{d+1}) \\ |L|,|M|<\infty}} \lim_{n\to\infty} \langle \chi_{L\triangle(M+n(1,\mathbf{0}))\triangle L_n}\rangle^{1/n} , \end{aligned}$$

where L_n is the part of the 0-axis between 0 and n. From Griffiths inequalities

$$\begin{aligned} \langle \chi_{L\triangle(M+n(1,\mathbf{0}))\triangle L_n}\rangle &\leqq \langle \chi_{L_n}\rangle \langle \chi_{L\triangle(M+n(1,\mathbf{0}))}\rangle^{-1} \\ &\leqq \langle \chi_{L_n}\rangle (\operatorname{th}\beta_h)^{-(|L|+|M|)} . \end{aligned}$$

Thus

$$\begin{aligned} t_0 &= \lim_{n\to\infty} \langle \chi_{L_n}\rangle^{1/n} = \lim_{n\to\infty} (\sigma_3(\mathbf{0})\Omega, T_0^n\sigma_3(\mathbf{0})\Omega)^{1/n} \\ &= (\Phi_0, T_0\Phi_0) . \end{aligned}$$

Now let $(\sigma_3(\mathbf{0})\Omega, T_0^n\sigma_3(\mathbf{0})\Omega)^2(\sigma_3(\mathbf{0})\Omega, T_0^{2n}\sigma_3(\mathbf{0})\Omega)^{-1} \to 0$. Then $t_0^{-n/2}T_0^{n/2}\sigma_3(\mathbf{0})\Omega \underset{s}{\to} 0$. By the same calculation as in the first part of the proof, one finds

$$t_0^{-n/2} T_0^{n/2} A\sigma_3(\mathbf{0})\Omega \underset{s}{\to} 0$$

for all $A \in \mathfrak{A}$, hence t_0 is not an eigenvalue. q.e.d.

It is an easy consequence of the polymer expansions mentioned before Proposition 5.5 that for $\beta_g < \beta_g^s$ or $\beta_h > \beta_h^s$

$$\lim_{n\to\infty} (\sigma_3(\mathbf{0})\Omega, T_0^n\sigma_3(\mathbf{0})\Omega)^2 (\sigma_3(\mathbf{0})\Omega, T_0^{2n}\sigma_3(\mathbf{0})\Omega)^{-1} \neq 0 \tag{5.15}$$

(cf. the proof of Theorem 5.6).

Bricmont and Fröhlich [4] have recently proposed to use the behaviour of $(\sigma_3(\mathbf{0})\Omega, T_0^n\sigma_3(\mathbf{0})\Omega)$ for large n as a criterion distinguishing the confinement/screening phase from the phase where charges exist. In Phase I one expects

$$(\sigma_3(\mathbf{0})\Omega, T_0^n\sigma_3(\mathbf{0})\Omega) \sim e^{-n\,\text{const}} \tag{5.16}$$

[an immediate consequence of (5.15)]. For Phase II, they point out that one should expect a behaviour like

$$(\sigma_3(\mathbf{0})\Omega, T_0^n\sigma_3(\mathbf{0})\Omega) \sim n^{-d/2}e^{-n\,\text{const}}. \tag{5.17}$$

(5.17) implies that (5.15) is violated, so according to Proposition 5.5, t_0 is not an eigenvalue in this case.

We now return to the investigation of the sequence $F_r\Omega\|F_r\Omega\|^{-1}$ in Region I_c.

Theorem 5.6. *Let* $\beta_g < \beta_g^s$ *or* $\beta_h > \beta_h^s$.

(i) *For all* $A \in \mathfrak{F}$

$$\lim_{r\to\infty} \omega_r(A) = (\sigma_3(\mathbf{0})\Phi_0, A\sigma_3(\mathbf{0})\Phi_0),$$

(ii) $F_r\Omega\|F_r\Omega\|^{-1} \underset{w}{\to} (\sigma_3(\mathbf{0})\Phi_0, \Omega)\sigma_3(\mathbf{0})\Phi_0$.

Proof. First we consider the case $\beta_g < \beta_g^s$.

(i) We have to investigate the sequence

$$\langle \chi_{L\triangle M_r}\rangle \langle \chi_{M_r}\rangle^{-1}.$$

We can write the expansion for expectation values, similar to that in Region II_c, in the form

$$\langle \chi_{L\triangle M_r}\rangle = \sum_{P\in\mathrm{Conn}(L)} (\mathrm{th}\,\beta_h)^{|L\triangle M_r\triangle\partial P|}(\mathrm{th}\,\beta_g)^{|P|} \cdot \exp\left\{\sum_\Gamma c_\Gamma(b_{L,P}^\Gamma A_{M_r}^\Gamma - 1)\mu^\Gamma\right\},$$

where $\mathrm{Conn}(L) = \{P \subset P(\mathbb{Z}^{d+1}),\ |P|<\infty,\ \partial P_1 \cap L \neq \emptyset$ for each connected component P_1 of $P\}$, $b_{L,P}(P') = 0$ if $P'\cap P \neq \emptyset$ or $\partial P' \cap (\partial P \cup L) \neq \emptyset$ and $b_{L,P}(P') = 1$ otherwise, and $A_{M_r}(P)$ is defined in (5.12). A slight complication comes from the fact that A_{M_r} is not bounded by 1. But negative values of $n_r = |\partial P| - 2|\partial P \cap M_r|$ can occur only for large sets of plaquettes P with $|P| \geqq \frac{r}{2}|n_r|$ (Appendix A.5). Hence

$$\|A_{M_r}\mu\| = \sup_P |A_{M_r}(P)\mu(P)|^{1/|P|} \leqq \mathrm{th}\,\beta_g(\mathrm{th}\,\beta_h)^{-2/r},$$

which is bounded by $\mathrm{th}\,\beta_g'$ for some $\beta_g < \beta_g' < \beta_g^s$, provided r is large enough. Thus

$$\begin{aligned}\langle \chi_{L\triangle M_r}\rangle\langle\chi_{M_r}\rangle^{-1} &= \sum_{P\in\mathrm{Conn}(L)} (\mathrm{th}\,\beta_h)^{|L\triangle\partial P| - 2|(L\triangle\partial P)\cap M_r|}(\mathrm{th}\,\beta_g)^{|P|} \\ &\quad\cdot \exp\left\{\sum_\Gamma c_\Gamma(b_{L,P}^\Gamma - 1)A_{M_r}^\Gamma \mu^\Gamma\right\} \\ &\xrightarrow[r\to\infty]{} \sum_{P\in\mathrm{Conn}(L)} (\mathrm{th}\,\beta_h)^{|L\triangle\partial P| - 2|(L\triangle\partial P)\cap M_\infty|}(\mathrm{th}\,\beta_g)^{|P|} \\ &\quad\cdot \exp\left\{\sum_\Gamma c_\Gamma(b_{L,P}^\Gamma - 1)A_{M_\infty}^\Gamma \mu^\Gamma\right\}.\end{aligned}$$

(ii) We have to show that for $L\subset B(\mathbb{Z}^{d+1})$, $|L|<\infty$

$$\lim_{r\to\infty}\langle\chi_{L\triangle N_r}\rangle\langle\chi_{M_r}\rangle^{-1/2}=\lim_{r\to\infty}\langle\chi_{L\triangle L_{0,r}}\rangle\langle\chi_{L_{0,r}}\rangle\langle\chi_{L_{-r,r}}\rangle^{-1},$$

where $L_{k,l}$ is the part of the 0-axis between k and l. From the first part of the proof we have

$$\langle\chi_{L\triangle N_r}\rangle\langle\chi_{M_r}\rangle^{-1/2}=\sum_{P\in\mathrm{Conn}(L)}(\mathrm{th}\,\beta_h)^{|L\triangle N_r\triangle\partial P|-1/2|M_r|}(\mathrm{th}\,\beta_g)^{|P|}\cdot\exp\left\{\frac{1}{2}\sum_{\Gamma}c_\Gamma[b^\Gamma_{L,P}A^\Gamma_{N_r}+b^\Gamma_{\theta L,\theta P}A^\Gamma_{\theta N_r}-A^\Gamma_{M_r}]\mu^\Gamma\right\},$$

where θ is the reflection on the $(x^0=0)$-hyperplane, and

$$\begin{aligned}&\langle\chi_{L\triangle L_{0,r}}\rangle\langle\chi_{L_{0,r}}\rangle\langle\chi_{L_{-r,r}}\rangle^{-1}\\&=\sum_{P\in\mathrm{Conn}(L)}(\mathrm{th}\,\beta_h)^{|L\triangle L_{0,r}\triangle\partial P|-|L_{0,r}|}(\mathrm{th}\,\beta_g)^{|P|}\\&\cdot\exp\left\{\frac{1}{2}\sum_{\Gamma}c_\Gamma[b^\Gamma_{L,P}A^\Gamma_{L_{0,r}}+b^\Gamma_{\theta L,\theta P}A^\Gamma_{\theta L_{0,r}}\right.\\&\left.-A^\Gamma_{L_{-r,r}}-A^\Gamma_{L_{-r,r}+x_r}+A^\Gamma_{L_{0,r}+x_r}+A^\Gamma_{\theta L_{0,r}+x_r}]\mu^\Gamma\right\},\end{aligned}$$

($x_r=(0,\mathbf{x}_r)$). Both expressions become equal in the limit $r\to\infty$ since they differ only by the contributions of large P and $\Gamma(|P|,\|\Gamma\|\gtrsim r)$ which can be estimated by (4.17). This concludes the proof in the case $\beta_g<\beta^s_g$.

If $\beta_h>\beta^s_h$, one can proceed similarly. The argument is even simpler since the function $B_L(M)$ which was defined in (5.14) is multiplicative on components and bounded by 1. q.e.d.

From Theorems 5.4 and 5.6 and Proposition 5.3 we infer that the behaviour of $(\Omega,F_r\Omega)\|F_r\Omega\|^{-1}$ can be used as a criterion for the existence or absence of charges [5]. The criterion proposed by Bricmont and Fröhlich [4] is very similar as may be seen by formulating it in the following way: Let $G_r=\sigma_3(\mathbf{0})\alpha_{ri}(\sigma_3(\mathbf{0}))$. Then $G_r\Omega$ may be interpreted as a state with two charges where the second charge is not shifted to spacelike infinity but to infinity in positive Euclidean time which means physically that the accompanying gauge field configuration is minimized with respect to energy. Presumably, the sequence of states $(G_r\Omega,\cdot G_r\Omega)\|G_r\Omega\|^{-2}$ converges to the charged state ω in Region II$_c$. In Region I$_c$ one has from Proposition 5.5 and (5.15) $G_r\Omega\|G_r\Omega\|^{-1}\underset{s}{\rightarrow}\sigma_3(\mathbf{0})\Phi_\mathbf{0}$. Thus the Bricmont-Fröhlich criterion consists essentially in replacing the F_r in our criterion by G_r.

It is interesting to look at the expectation value of the charge operator in the state induced by $\sigma_3(\mathbf{0})\Phi_\mathbf{0}$. One finds for $\beta_g<\beta^s_g$ or $\beta_h>\beta^s_h$,

$$\lim_{\Lambda\nearrow\mathbb{Z}^d}\frac{(\sigma_3(\mathbf{0})\Phi_\mathbf{0},Q_\Lambda\sigma_3(0)\Phi_\mathbf{0})}{(\Omega,Q_\Lambda\Omega)}=1. \tag{5.18}$$

Thus also this behaviour is a good test for distinguishing the two phases of the model. A nice feature of this quantity is that it even shows a difference between the screening ($\beta_h > \beta_h^s$, β_g large) and the confinement ($\beta_g < \beta_g^s$, β_h small) regime. In the screening case for $\mathbf{\Lambda}$ sufficiently small, $\mathbf{0} \in \mathbf{\Lambda}$,

$$\frac{(\sigma_3(\mathbf{0})\Phi_\mathbf{0}, Q_\mathbf{\Lambda}\sigma_3(\mathbf{0})\Phi_\mathbf{0})}{(\Omega, Q_\mathbf{\Lambda}\Omega)} < 0. \tag{5.19}$$

Hence locally the state induced by $\sigma_3(\mathbf{0})\Phi_\mathbf{0}$ looks like a charged state, but charge measurements for large regions are completely screened by vacuum fluctuations. In the confinement case

$$\frac{(\sigma_3(\mathbf{0})\Phi_\mathbf{0}, Q_\mathbf{\Lambda}\sigma_3(\mathbf{0})\Phi_\mathbf{0})}{(\Omega, Q_\mathbf{\Lambda}\Omega)} \approx 1, \tag{5.20}$$

hence even locally $\sigma_3(\mathbf{0})\Phi_\mathbf{0}$ does not look like a charged state. To get an idea why the charge has disappeared one may look at the approximating sequence ω_r. One finds

$$\frac{\omega_r(Q_\mathbf{\Lambda})}{\omega_0(Q_\mathbf{\Lambda})} < 0 \tag{5.21}$$

for small r, provided $\mathbf{0} \in \mathbf{\Lambda}$ and $\mathbf{x}_r \notin \mathbf{\Lambda}$. One may interpret this result in the following way; charge separation with a fixed amount of energy leads to "charge fragmentation" when a critical distance has been reached.

A further feature that distinguishes the phase with charges from the confinement/screening phase is the different behaviour of the charge (= total electric flux) and the electric flux through nonclosed surfaces. Let $\mathbf{\Lambda}$ be a hypercube in $\mathbb{Z}^d$. Decompose $\partial^*\mathbf{\Lambda}$ into two parts $\mathbf{S}_1$ and $\mathbf{S}_2$, each consisting of the bonds intersecting with one half of the boundary of $\mathbf{\Lambda}$. Let $E(\mathbf{S}_i) = \prod_{\mathbf{b} \in \mathbf{S}_i} \tau_1(\mathbf{b})$, $i = 1, 2$. Then $Q_\mathbf{\Lambda} = E(\mathbf{S}_1)E(\mathbf{S}_2)$. In the charge phase one finds

$$-\ln \frac{\omega_0(E(\mathbf{S}_1))\omega_0(E(\mathbf{S}_2))}{\omega_0(Q_\mathbf{\Lambda})} \sim |\mathbf{S}_3|, \tag{5.22}$$

where $\mathbf{S}_3$ is the minimal set of bonds with $\partial^*\mathbf{S}_3 = \partial^*\mathbf{S}_1 (= \partial^*\mathbf{S}_2)$. This is a signal for the existence of large vacuum fluctuations of the electric flux through nonclosed surfaces which makes it impossible to measure the asymptotic direction of a string transporting the electric flux to infinity. As claimed in [11] this is a necessary condition for the existence of states with gauge charges. In the confinement/screening region we have instead

$$-\ln \frac{\omega_0(E(\mathbf{S}_1))\omega_0(E(\mathbf{S}_2))}{\omega_0(Q_\mathbf{\Lambda})} \sim |\partial^*\mathbf{S}_1|. \tag{5.23}$$

Equation (5.22) suggests that charged states, if they exist, could be further distinguished by the asymptotic direction of the string. This is impossible, in the general framework of quantum field theory, for particles in a massive theory [11] and probably excluded for all states with finite energy.

These "order parameters" which test the existence of charged states and their utility in numerical analysis will be discussed in more detail in [13].

6. Energy and Translations

Confinement in the sense of Wilson means that it is impossible to create a charged state with finite energy. Therefore, the sequence of vectors $(F_r\Omega)_{r\in\mathbb{N}}$ in Sect. 5 was chosen in such a way that their energy is uniformly bounded. More precisely, we have the following result.

Proposition 6.1. *For all* $n\in\mathbb{N}$ *there exists a constant* $c_n>0$ *such that for all* $r\in\mathbb{N}$

$$(F_r\Omega, T_0^{-n}F_r\Omega)\,\|F_r\Omega\|^{-2}<c_n\,.$$

Proof. From the definition of F_r in (5.1) we have

$$\begin{aligned}&(F_r\Omega, T_0^{-n}F_r\Omega)\,\|F_r\Omega\|^{-2}\\ &=(F_r\Omega, \alpha_{-in}(\sigma_3(\mathbf{0})\sigma_3(\mathbf{x}_r))T_0^{r-n}\tau_3(\mathbf{L}_r)\Omega)\,\|F_r\Omega\|^{-2}\\ &\leqq \|\alpha_{-in}(\sigma_3(\mathbf{0}))\|^2\,\|T_0^{r-n}\tau_3(\mathbf{L}_r)\Omega\|\,\|T_0^{r}\tau_3(\mathbf{L}_r)\Omega\|^{-1}\,.\end{aligned}$$

For $r<n$ there is nothing to show. For $r\geqq n$ we find by a repeated application of Schwartz' inequality

$$\frac{\|T_0^{r-n}\tau_3(\mathbf{L}_r)\Omega\|}{\|T_0^{r}\tau_3(\mathbf{L}_r)\Omega\|}\leqq\frac{\|T_0^{r-2n}\tau_3(\mathbf{L}_r)\Omega\|^{1/2}}{\|T_0^{r}\tau_3(\mathbf{L}_r)\Omega\|^{1/2}}\leqq\ldots\leqq\frac{\|T_0^{r-2^kn}\tau_3(\mathbf{L}_r)\Omega\|^{2^{-k}}}{\|T_0^{r}\tau_3(\mathbf{L}_r)\Omega\|^{2^{-k}}}\,,$$

hence choosing $k\in\mathbb{Z}_+$ such that $2^kn\leqq r<2^{k+1}n$, and keeping in mind that $\|T_0^r\tau_3(\mathbf{L}_r)\Omega\|\leqq 1$, we arrive at the estimate

$$\frac{\|T_0^{r-n}\tau_3(\mathbf{L}_r)\Omega\|}{\|T_0^{r}\tau_3(\mathbf{L}_r)\Omega\|}\leqq\|T_0^r\tau_3(\mathbf{L}_r)\Omega\|^{-\frac{2n}{r}}\,.$$

But $\|T_0^r\tau_3(\mathbf{L}_r)\Omega\|^2=\langle\chi_{M_r}\rangle$, and from Griffiths inequalities $\langle\chi_{M_r}\rangle\geqq(\mathrm{th}\,\beta_h)^{8r}$, hence

$$(F_r\Omega, T_0^{-n}F_r\Omega)\,\|F_r\Omega\|^{-2}\leqq\|\alpha_{-in}(\sigma_3(\mathbf{0}))\|^2\,(\mathrm{th}\,\beta_h)^{-8n}\,.\quad \text{q.e.d.}$$

Because of this behaviour we expect that whenever the limit state $\omega=\lim\omega_r$ will exist it will have finite energy in the sense that it is possible, in the cyclic representation induced by ω, to define a transfer matrix T which implements the imaginary time translation α_i. Since we are interested in the properties of the charged states we will from now on consider the case $\beta_g<\beta_g^w$ and $\beta_h>\beta_h^w$. We exploit the fact that the state $\omega\circ\varrho_0$ is invariant under imaginary time translations. It is even a ground state of $\mathfrak{A}$ as may be seen by its cluster properties under separation in Euclidean time. Let $(\pi,\mathscr{H},\Phi_0)$ denote the GNS triple associated to $\omega\circ\varrho_0$. $\Phi=\pi(\sigma_3(\mathbf{0}))\Phi_0$ is then the gauge invariant vector inducing the state ω. The transfer matrix T may be introduced by

$$T\pi(A)\Phi_0=\pi\alpha_i(A)\Phi_0\,,\qquad A\in\mathfrak{F}_0\,. \tag{6.1}$$

The properties of T and its relation to the transfer matrix T_0 in the vacuum sector are described in the following theorem.

Theorem 6.2. (i) *T is a bounded positive operator with a densely defined inverse such that*

$$T\pi(A)T^{-1}=\pi\alpha_i(A)$$

for all $A\in\mathfrak{F}_0$.

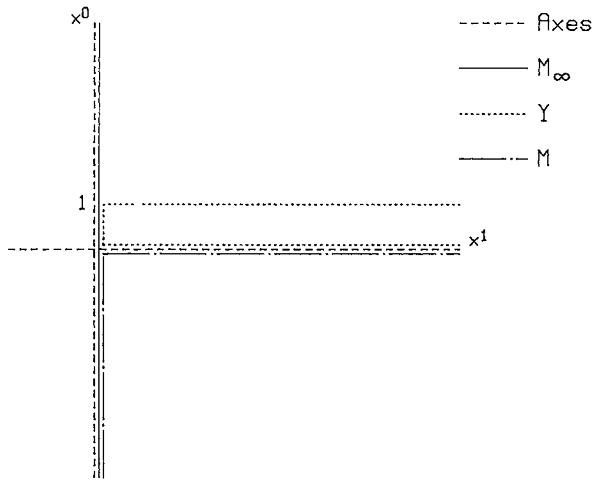

Fig. 6.1. The lines Y, M_∞, and M in the $(1-0)$-plane

(ii) *There exists a constant* $a>0$ *such that for* $A, B\in\mathfrak{F}$ *and* $n\in\mathbb{Z}_+$

$$\lim_{r\to\infty}(A\alpha_{\mathbf{x}_r}(B)F_r\Omega, T_0^n A\alpha_{\mathbf{x}_r}(B)F_r\Omega)\|F_r\Omega\|^{-2}$$
$$=a^{2n}(\pi(A)\Phi, T^n\pi(A)\Phi)(\pi(B)\Phi, T^n\pi(B)\Phi). \qquad (*)$$

(iii) *a is given by the formula*

$$a=\exp\left\{\sum_\Gamma c_\Gamma[(P_\Gamma, Y)-1]\tfrac{1}{2}[1-(P_\Gamma, M_\infty)](P_\Gamma, M)\mu^\Gamma\right\},$$

where Y is the border of the infinite rectangle in the $(0-1)$*-plane with* $0\leqq x^0\leqq 1$ *and* $x^1\geqq 0$, M_∞ *of the halfplane* $x^1\leqq 0$ *and M of the quadrant* $x^1\geqq 0$, $x^0\leqq 0$ *(Fig. 6.1)*

$$\left((P_\Gamma, Y)=\prod_\gamma (P_\gamma, Y)^{\Gamma(\gamma)}\right).$$

Proof. (i) Positivity and boundedness of T will follow from (ii) since $a>0$ by a slight generalization of Proposition 6.1 or by (iii). The inverse of T is given by

$$T^{-1}\pi(A)\Phi_0=\pi\alpha_{-i}(A)\Phi_0\,,\qquad A\in\mathfrak{F}_0\,,$$

and the implementation relation follows immediately. (ii) and (iii). Since the left hand side of $(*)$ is uniformly bounded by $\|A\|^2\|B\|^2$ it is sufficient to prove $(*)$ for $A, B\in\mathfrak{F}_0$. For this it is enough to check the relations $(L, N\subset B(\mathbb{Z}^{d+1}), |L|, |N|<\infty)$

$$\lim_{r\to\infty}\frac{\langle\chi_{L\triangle(N+x_r)\triangle M_r^{(n)}}\rangle}{\langle\chi_{M_r^{(n)}}\rangle}=\lim_{r\to\infty}\frac{\langle\chi_{L\triangle M_r}\rangle}{\langle\chi_{M_r}\rangle}\lim_{r\to\infty}\frac{\langle\chi_{N\triangle M_r}\rangle}{\langle\chi_{M_r}\rangle},$$

and

$$\lim_{r\to\infty}\frac{\langle\chi_{M_r^{(n)}}\rangle}{\langle\chi_{M_r}\rangle}=a^{2n},$$

where $M_r^{(n)}$ is the rectangle with edges $(-r,\mathbf{0})$, $(r+n,\mathbf{0})$, $(r+n,\mathbf{x}_r)$, $(-r,\mathbf{x}_r)$. Now the first of these relations is an easy consequence of the polymer expansion in the same way as in the proof of the existence of the charged state ω. To prove the second relation we need the following lemma.

Lemma 6.3. *Let*

$$a_{(k)}=\exp\left\{\sum_{\Gamma} c_\Gamma[(P_\Gamma, Y^{(k)})-1]\tfrac{1}{2}[1-(P_\Gamma, M_\infty)](P_\Gamma, M)\mu^\Gamma\right\},$$

where $Y^{(k)}$ is the boundary of the infinite rectangle $0\leqq x^0\leqq k$, $x^1\geqq 0$ in the $(0-1)$-plane, $k\in\mathbb{N}$. Then one has the product relation $a_{(k)}a_{(l)}=a_{(k+l)}$ and therefore $a_{(k)}=a^k$.

Proof of the Lemma. First we convince ourselves that the sum in the definition of $a_{(k)}$ converges. Only those Γ can contribute for which P_Γ contains a plaquette p in the interior of $Y^{(k)}$ and a plaquette in the halfplane $x^1\leqq 0$, hence $\frac{1}{2}\|\Gamma\|\geqq \operatorname{dist}(p, M_\infty)$. Thus from (4.17)

$$\begin{aligned}&\left|\sum_{\Gamma} c_\Gamma[(P_\Gamma, Y^{(k)})-1]\tfrac{1}{2}[1-(P_\Gamma, M_\infty)](P_\Gamma, M)\mu^\Gamma\right|\\ &\quad\leqq \sum_{r=1}^{\infty} 2k \sum_{\substack{\Gamma \nmid \partial p\\ \|\Gamma\|\geqq 2r}} |c_\Gamma|\,|\mu^\Gamma|\leqq 8kF_1(2\beta^c)(e^{4(\beta-\beta^c)}-1)^{-1}.\end{aligned}$$

From translation invariance of μ and M_∞ in the 0-direction we have

$$a_{(l)}=\exp\left\{\sum_{\Gamma} c_\Gamma[(P_\Gamma, Y^{(l)}+(k,\mathbf{0}))-1]\tfrac{1}{2}[1-(P_\Gamma, M_\infty)](P_\Gamma, M\triangle Y^{(k)})\mu^\Gamma\right\},$$

and the product relation follows from $(Y^{(l)}+(k,\mathbf{0}))\triangle Y^{(k)}=Y^{(k+l)}$. q.e.d.

To complete the proof of part (ii) and (iii) of the theorem it is sufficient to show that

$$\lim_{r\to\infty}\frac{\langle \chi_{M_r^{(n)}}\rangle}{\langle \chi_{M_r}\rangle}=a_{(n)}^2.$$

Let $\theta_r^{(n)}$ denote the reflection on the hyperplane $x^0=r+\dfrac{n}{2}$. Then

$$\begin{aligned}\ln\frac{\langle \chi_{M_r^{(n)}}\rangle}{\langle \chi_{M_r}\rangle}&=\sum_{\Gamma} c_\Gamma[(P_\Gamma, Y^{(n)}+(r,\mathbf{0}))-1](P_\Gamma, M_r)\mu^\Gamma\\ &=\sum_{\Gamma}[(P_\Gamma, Y^{(n)}+(r,\mathbf{0}))-1]\tfrac{1}{2}[(P_\Gamma, M_r)+(P_\Gamma, \theta_r^{(n)}M_r)]\mu^\Gamma,\end{aligned}$$

where the latter equation comes from the invariance of the activities under lattice reflections and from the invariance of $Y^{(n)}+(r,\mathbf{0})$ under $\theta_r^{(n)}$. Using the identity

$$(P_\Gamma, M_r)+(P_\Gamma, \theta_r^{(n)}M_r)=(P_\Gamma, M_r)(1-(P_\Gamma, M_r^{(n+2r)})),$$

which holds for $(P_\Gamma, Y^{(n)}+(r,\mathbf{0}))=-1$, and neglecting all Γ with $\|\Gamma\|\geqq r$ – their contribution vanishes in the limit $r\to\infty$ – the sum splits into two parts, one containing all Γ, where P_Γ intersects with the halfplane $x^1\leqq 0$, the other containing all Γ, where P_Γ intersects with the halfplane $x^1\geqq 2r$. Both parts converge to $\ln a_{(n)}$ in the limit $r\to\infty$, thus completing the proof of the theorem. q.e.d.

The constant $-\ln a$ can be interpreted as the minimal energy of a gauge field configuration accompanying an external charge. It coincides with the constant governing the perimeter law of the Wilson loop:

Corollary. *Let R_{kl} denote the boundary of a rectangle in a lattice plane with side lengths l and k. Then*

$$\lim_{k,l\to\infty} \langle \chi_{R_{kl}} \rangle a^{-|R_{kl}|} = \text{const} \neq 0 .$$

Proof. Let $f_{k,l} = \ln(\langle \chi_{R_{kl}} \rangle a^{-|R_{kl}|})$. In the proof of Theorem 6.2 we established the estimate

$$|f_{k+1,l} - f_{k,l}| \leqq \text{const}\, e^{-\text{const} \min(k,l)}$$

with positive constants. Thus

$$\begin{aligned} |f_{k,l} - f_{k',l'}| \leqq \lim_{n\to\infty} \{ &|f_{k,l} - f_{k+n,l+n}| + |f_{k+n,l+n} - f_{k'+n,l'+n}| \\ &+ |f_{k'+n,l'+n} - f_{k',l'}| \} \\ \leqq\, &4\,\text{const}\, e^{-\text{const} \min(k,l,k',l')} \sum_{m=0}^{\infty} e^{-\text{const}\, m} , \end{aligned}$$

which proves the corollary. q.e.d.

It is easy to find ground states of $\mathfrak{A}$ with general configurations of external charges. Let $q : \mathbb{Z}^d \to \{\pm 1\}$ be a function with finite support. A ground state ω_q with the external charge configuration

$$\omega_q(\hat{q}(\mathbf{x})) = q(\mathbf{x}) , \qquad \mathbf{x} \in \mathbb{Z}^d , \tag{6.2}$$

is given by $\omega_q(A) = \langle A_{cl} \rangle_q$, A_{cl} being a classical function corresponding to A according to the rules in Sect. 3, with the expectation values

$$\langle \chi_L \rangle_q = \sum_{M \in \text{Conn}(L)} (\text{th}\,\beta_h)^{|M|} \exp\left\{ \sum_{\Gamma} c_\Gamma (a^\Gamma_{L,M} - 1) a^\Gamma_{M_q} \mu^\Gamma \right\} , \tag{6.3}$$

where $M_q = \bigcup_{\mathbf{x} \in \text{supp}\, q} (M_\infty + (0, \mathbf{x}))$. [Compare (5.8).]

Theorem 6.4. *If $|\text{supp}\, q|$ is even, there exists a unique eigenvector Φ_q of T_0 in $\mathscr{H}_0$ inducing ω_q, with $\left(\prod_{\mathbf{x} \in \text{supp}\, q} \sigma_3(\mathbf{x}) \Omega, \Phi_q \right) > 0$. If $|\text{supp}\, q|$ is odd, there exists a unique eigenvector Φ_q of T in $\mathscr{H}$ inducing ω_q with $\left(\prod_{\mathbf{x} \in \text{supp}\, q \triangle \{\mathbf{0}\}} \pi(\sigma_3(\mathbf{x})) \Phi_{\mathbf{0}}, \Phi_q \right) > 0.$*

Proof. We already know the statement for $q \equiv 1$ ($\Phi_q = \Omega$) and for $q(\mathbf{x}) = (-1)^{\delta_{0,\mathbf{x}}}$ ($\Phi_q = \Phi_{\mathbf{0}}$). Let $|\text{supp}\, q|$ be even. Then there exists a finite set of bonds $\mathbf{L} \subset B(\mathbb{Z}^d)$ with $\partial \mathbf{L} = \text{supp}\, q$. Let $\tau_3(\mathbf{L}) = \prod_{\mathbf{b} \in \mathbf{L}} \tau_3(\mathbf{b})$.

Then $\alpha_{in}(\tau_3(\mathbf{L}))\Omega =: \Phi_q^{(n)}$ has an external charge configuration described by q. It is easy to see that

$$\lim_{n\to\infty} (\Phi_q^{(n)}, A \Phi_q^{(n)}) \, \|\Phi_q^{(n)}\|^{-2} = \omega_q(A) .$$

Furthermore, the sequence $\Phi_q^{(n)}\|\Phi_q^{(n)}\|^{-1}$ converges strongly. To see this we use the formula $(m>n)$

$$\begin{aligned}&\ln\left\{\frac{(\phi_q^{(n)},\phi_q^{(m)})}{\|\phi_q^{(n)}\|\,\|\phi_q^{(m)}\|}\right\}\\&=\ln\left\{\frac{(\Omega,\tau_3(\mathbf{L})T_0^{n+m}\tau_3(\mathbf{L})\Omega)}{(\Omega,\tau_3(\mathbf{L})T_0^{2n}\tau_3(\mathbf{L})\Omega)^{1/2}(\Omega,\tau_3(\mathbf{L})T_0^{2m}\tau_3(\mathbf{L})\Omega)^{1/2}}\right\}\\&=\tfrac{1}{2}\sum_{\Gamma}c_\Gamma\{(P_\Gamma,L_{-n,m})+(P_\Gamma,L_{-m,n})-(P_\Gamma,L_{-n,n})-(P_\Gamma,L_{-m,m})\}\mu^\Gamma\,,\end{aligned}$$

where for $k,\,l\in\mathbb{Z}$, $k<l$, L_{kl} is the set of bonds

$$L_{kl}=\{\{k\}\times\mathbf{b},\{l\}\times\mathbf{b},\mathbf{b}\in\mathbf{L}\}\cup\{\{(j,\mathbf{x}),(j+1,\mathbf{x})\},k\leqq j<l,\mathbf{x}\in\partial\mathbf{L}\}\,.$$

Clearly, $L_{kl}\triangle L_{lm}=L_{km}$. Hence

$$\begin{aligned}&(P_\Gamma,L_{-n,m})+(P_\Gamma,L_{-m,n})-(P_\Gamma,L_{-n,n})-(P_\Gamma,L_{-m,m})\\&\quad=[(P_\Gamma,L_{-m,-n})-1]\,(P_\Gamma,L_{-n,n})\,[1-(P_\Gamma,L_{n,m})]\,.\end{aligned}$$

Therefore only those Γ contribute to the sum for which $(P_\Gamma,L_{-m,-n})=(P_\Gamma,L_{n,m})=-1$. If we write

$$(P_\Gamma,L_{n,m})=\prod_{k=n}^{m-1}(P_\Gamma,L_{k,k+1})\,,$$

we see that the sum can be estimated by [cf. (4.17)]

$$\begin{aligned}\sum_{k=n}^{m-1}\sum_{\substack{\Gamma+L_{k,k+1}\\ \|\Gamma\|\geqq n+k}}|c_\Gamma|\,|\mu^\Gamma|&\leqq\sum_{k=n}^{\infty}|L_{k,k+1}|F_1(2\beta^c)e^{-2(\beta-\beta^c)(n+k)}\\&\leqq\mathrm{const}\,e^{-n\,\mathrm{const}}\,.\end{aligned}$$

This proves that $\Phi_q^{(n)}\|\Phi_q^{(n)}\|^{-1}$ is a Cauchy sequence, and we may identify the limit vector with Φ_q. Φ_q is an eigenvector of T_0 by construction. Φ_q is cyclic for $\mathfrak{F}$, since by the same argument as before $\alpha_{in}(\tau_3(\mathbf{L}))\Phi_q$ converges strongly to Ω. Hence from cluster properties of ω_q under separation in Euclidean time the uniqueness of Φ_q (up to a phase) follows. It remains to establish the positivity of the scalar product $\left(\prod_{\mathbf{x}\in\mathrm{supp}\,q}\sigma_3(\mathbf{x})\Omega,\,\Phi_q\right)$. We have [cf. (4.16)]

$$\begin{aligned}&\left(\prod_{\mathbf{x}\in\mathrm{supp}\,q}\sigma_3(\mathbf{x})\Omega,\Phi_q^{(n)}\right)\|\Phi_q^{(n)}\|^{-1}=\langle\chi_{K_n}\rangle\langle\chi_{L_{-n,n}}\rangle^{-1/2}\\&=\sum_{M\in\mathrm{Conn}(K_n)}(\mathrm{th}\,\beta_h)^{|M|}\exp\{\tfrac{1}{2}\textstyle\sum c_\Gamma[a^\Gamma_{K_n,M}+a^\Gamma_{\theta K_n,\theta M}-a^\Gamma_{L_{-n,n}}-1]\mu^\Gamma\}\,,\end{aligned}$$

where K_n is the subset of bonds of $L_{-n,n}$ which belong to the halfspace $x^0\geqq0$ and θ is the reflection on the hyperplane $x^0=0$. Since ∂K_n and hence also $\mathrm{Conn}(K_n)$ are independent of n, it is sufficient to prove the convergence of the sum in the exponent as n goes to infinity. Only those Γ contribute for which either there is some $\gamma\in\mathrm{supp}\,\Gamma$ with $N_\gamma\notin\mathrm{Disc}(M)\cap\mathrm{Disc}(\theta M)$ or $(P_\Gamma,\,M\triangle\{0\}\times\mathbf{L})=-1$ or $(P_\Gamma,\,\theta M\triangle\{0\}\times\mathbf{L})=-1$ or $(P_\Gamma,\,K_n\triangle\{0\}\times\mathbf{L})=-1$, $(P_\Gamma,\,L_{-n,n})=+1$. From (4.15) the contribution of those Γ which fulfill one of the first three conditions is bounded by $2(|M|+|\mathbf{L}|)F_1(2\beta^c)$; for the last condition one obtains the bound

$$\sum_{k=1}^{\infty} \sum_{\substack{\Gamma \neq L_{k-1,k} \\ \|\Gamma\| \geq k}} |c_\Gamma| |\mu^\Gamma| \leqq (2|L|+|\partial L|)F_1(2\beta^c)(e^{2(\beta-\beta^c)}-1)^{-1}.$$

If $|\mathrm{supp}\,q|$ is odd one makes the same construction in the charged Hilbert space $\mathscr{H}$ starting from Φ_0 and using the even set $\mathrm{supp}\,q \triangle \{\mathbf{0}\}$. q.e.d.

According to the preceding theorem, there are, for each $\mathbf{x} \in \mathbb{Z}^d$, vectors $\Phi_{q_\mathbf{x}} = \Phi_\mathbf{x}$, $q_\mathbf{x}(\mathbf{x}) = -1$, $q_\mathbf{x}(\mathbf{y}) = 1$ for $\mathbf{x} \neq \mathbf{y}$, which induce the state $\omega_{q_\mathbf{x}} = \omega \circ \varrho_0 \circ \alpha_{-\mathbf{x}}$. This leads to the following simple definition of translation operators $U(\mathbf{x})$,

$$U(\mathbf{x})\pi(A)\Phi_0 = \pi\alpha_\mathbf{x}(A)\Phi_\mathbf{x}, \qquad A \in \mathfrak{F}. \tag{6.4}$$

Theorem 6.5. (i) $\mathbf{x} \to U(\mathbf{x})$ *is a unitary representation of the group of lattice translations implementing the automorphisms* $\alpha_\mathbf{x}$,

$$U(\mathbf{x})\pi(A)U(-\mathbf{x}) = \alpha_\mathbf{x}(A).$$

(ii) $U(\mathbf{x})$ *commutes with* T.
(iii) $U(\mathbf{x}) \longrightarrow 0$ *as* $|\mathbf{x}| \to \infty$.

Proof. (ii) follows from $\alpha_\mathbf{x}\alpha_i = \alpha_i\alpha_\mathbf{x}$ and $T\Phi_\mathbf{x} = \Phi_\mathbf{x}$.
(i) It is sufficient to show

$$U(\mathbf{x})\Phi_\mathbf{y} = \Phi_{\mathbf{y}+\mathbf{x}}.$$

From (ii) and Theorem 6.4 it follows that $U(\mathbf{x})\Phi_\mathbf{y} = \lambda\Phi_{\mathbf{y}+\mathbf{x}}$ with $|\lambda| = 1$. Let $\mathbf{L}_\mathbf{z}$ for $\mathbf{z} \in \mathbb{Z}^d$ denote a path from $\mathbf{0}$ to $\mathbf{z}$ with $|\mathbf{L}_\mathbf{z}| = |\mathbf{z}|$. Then

$$\begin{aligned}
&(U(\mathbf{x})\Phi_\mathbf{y}, \Phi_{\mathbf{x}+\mathbf{y}}) \\
&= \lim_{n\to\infty} \lim_{m\to\infty} (\pi\alpha_{in}(\tau_3(\mathbf{L}_\mathbf{y}+\mathbf{x}))\pi\alpha_{im}(\tau_3(\mathbf{L}_\mathbf{x}))\Phi_0, \pi\alpha_{in}(\tau_3(\mathbf{L}_{\mathbf{x}+\mathbf{y}})\Phi_0) \\
&\quad \cdot \|\pi\alpha_{in}(\tau_3(\mathbf{L}_\mathbf{y}))\Phi_0\|^{-1} \|\pi\alpha_{im}(\tau_3(\mathbf{L}_\mathbf{x}))\Phi_0\|^{-1} \|\pi\alpha_{in}(\tau_3(L_{\mathbf{x}+\mathbf{y}}))\Phi_0\|^{-1},
\end{aligned}$$

which is nonnegative as may be seen by Griffiths inequalities or by the expansion of its logarithm. Thus $\lambda = 1$.

(iii) Let E_q denote the projection on vectors in $\mathscr{H}$ with external charge configuration q, $\mathrm{supp}\,q$ finite. Then $U(\mathbf{x})E_q \xrightarrow[w]{} 0$ if $q \neq 1$. It is therefore sufficient to look at matrix elements of $U(\mathbf{x})$ in the gauge invariant subspace of $\mathscr{H}$. Let A, $B \in \mathfrak{A} \cap \mathfrak{F}_0$. Then for $|\mathbf{x}|$ large enough

$$\begin{aligned}
&(\pi(A)\Phi, U(\mathbf{x})\pi(B)\Phi) \\
&= (\Phi_0, \pi(\sigma_3(\mathbf{0})A^*\alpha_\mathbf{x}(B\sigma_3(\mathbf{0}))\Phi_\mathbf{x}) \\
&= \lim_{n\to\infty} (\Phi_0, \pi(\varrho_0(A^*)\alpha_\mathbf{x}(B)\sigma_3(\mathbf{0})\sigma_3(\mathbf{x})\alpha_{in}(\tau_3(\mathbf{L}_\mathbf{x})))\Phi_0) \|\pi\alpha_{in}(\tau_3(\mathbf{L}_\mathbf{x}))\Phi_0\|^{-1} \\
&= \lim_{n\to\infty} \langle A^\varrho B_\mathbf{x} \chi_{K_n} \rangle_{q_0} \langle \chi_{L_{-n,n}} \rangle_{q_0}^{-1/2},
\end{aligned}$$

where $\mathbf{L}_\mathbf{x}$ was defined in the proof of (ii), K_n and $L_{-n,n}$ in the proof of Theorem 6.4, and where A^ϱ and $B_\mathbf{x}$ are classical functions with support in the time slice $-1 \leqq x^0$

$\leqq 0$ corresponding to $\varrho_0(A^*)$ and $\alpha_{\mathbf{x}}(B)$, respectively. But for finite sets R_1, $R_2 \subset B(\mathbb{Z}^{d+1})$ (cf. the proof of Proposition 5.3) one has

$$\begin{aligned} &\langle \chi_{R_1 \triangle (R_2 + (0,\mathbf{x})) \triangle K_n} \rangle_{q_0} \\ &= \lim_{r\to\infty} \langle \chi_{R_1 \triangle (R_2+(0,\mathbf{x}))\triangle K_n \triangle M_r} \rangle \langle \chi_{M_r} \rangle^{-1} \\ &\leqq \lim_{r\to\infty} \langle \chi_{K_n \triangle M_r} \rangle \langle \chi_{M_r} \rangle^{-1} \langle \chi_{R_1} \rangle^{-1} \langle \chi_{R_2} \rangle^{-1} \leqq (\operatorname{th}\beta_h)^{-|R_1|-|R_2|} \langle \chi_{K_n} \rangle_{q_0}. \end{aligned}$$

It is therefore sufficient to consider the case $A, B = 1$, i.e. to look at the "two-point function"

$$(\Phi, U(\mathbf{x})\Phi) = (\Phi_0, \pi(\sigma_3(0)\sigma_3(\mathbf{x}))\Phi_{\mathbf{x}}) = \lim_{n\to\infty} \langle \chi_{K_n} \rangle_{q_0} \langle \chi_{L_{-n,n}} \rangle_{q_0}^{-1/2}.$$

From (5.6)

$$\langle \chi_{K_n} \rangle_{q_0} = \sum_{M \in \operatorname{Conn}(K_n)} (\operatorname{th}\beta_h)^{|M|} \lim_{r\to\infty} \sigma(K_n \triangle M_r, M) \langle \chi_{M_r} \rangle^{-1},$$

from Proposition 4.1, (i) and Griffiths inequalities

$$\sigma(K_n \triangle M_r, M) \leqq \langle \chi_{K_n \triangle M_r \triangle M} \rangle \leqq \langle \chi_{L_{0,n} \triangle M_r} \rangle \langle \chi_{M \triangle \{0\} \times \mathbf{L}_{\mathbf{x}}} \rangle^{-1},$$

from Schwartz' inequality

$$\begin{aligned} &\lim_{r\to\infty} \langle \chi_{L_{0,n} \triangle M_r} \rangle \langle \chi_{M_r} \rangle^{-1} \\ &= \langle \chi_{L_{0,n}} \rangle_{q_0} = (\pi(\tau_3(\mathbf{L}_{\mathbf{x}}))\Phi_0, T^n \pi(\tau_3(\mathbf{L}_{\mathbf{x}})\Phi_0)) \\ &\leqq \| T^n \pi(\tau_3(\mathbf{L}_{\mathbf{x}}))\Phi_0 \| = \langle \chi_{L_{-n,n}} \rangle_{q_0}^{1/2}, \end{aligned}$$

and from the perimeter law for the Wilson loop in the pure gauge theory

$$\langle \chi_{M \triangle \{0\} \times \mathbf{L}} \rangle^{-1} \leqq (\langle \chi_{M \triangle \{0\} \times \mathbf{L}} \rangle^0)^{-1} \leqq e^{\alpha(\beta_g)(|M| + |\mathbf{x}|)},$$

hence

$$\begin{aligned} (\Phi, U(\mathbf{x})\Phi) &\leqq \sum_{M \in \operatorname{Conn}(K_n)} [(\operatorname{th}\beta_h) e^{\alpha(\beta_g)}]^{|M|} e^{\alpha(\beta_g)|\mathbf{x}|} \\ &\leqq [(2d+1)(\operatorname{th}\beta_h) e^{2\alpha(\beta_g)}]^{|\mathbf{x}|} [1 - (2d+1)^2 (\operatorname{th}\beta_h)^2 e^{2\alpha(\beta_g)}]^{-1}. \quad \text{q.e.d.} \end{aligned}$$

7. Dynamics in the Charged Sector

The transfer matrix T in the Hilbert space of charged states, $\mathscr{H}$, is positive and has a densely defined inverse, thus the time evolution in the charged sector may be described by

$$\check{\alpha}_t \pi(A) = T^{-it} \pi(A) T^{it}, \qquad A \in \mathfrak{F}. \tag{7.1}$$

Again it is an open question whether $\check{\alpha}_t \pi(\mathfrak{F})$ is contained in $\pi(\mathfrak{F})$ or at least in its weak closure.

For a physical interpretation it is essential that the dynamics in the charged sector can be compared with the dynamics in the vacuum sector. The most natural way to do this is, in our opinion, to look at the time invariant algebra $\hat{\mathfrak{F}}$ which was defined as the smallest C^*-algebra containing $\mathfrak{F}=\pi_0(\mathfrak{F})$ which is invariant under the time translations $\hat{\alpha}_t$ of the vacuum sector. That $\check{\alpha}_t$ and $\hat{\alpha}_t$ actually describe the same dynamics (in a sense to be specified) is the content of the following theorem.

Theorem 7.1. *There exists a unique representation $\hat{\pi}$ of $\hat{\mathfrak{F}}$ in $\mathscr{H}$ with the properties*

(i) $$\hat{\pi}(A)=\pi(A), \quad A\in\mathfrak{F}$$

(ii) $$\hat{\pi}\hat{\alpha}_t=\check{\alpha}_t\hat{\pi}, \quad t\in\mathbb{R}.$$

Proof. If a representation $\hat{\pi}$ with the properties (i) and (ii) exists it is of the form $(A_1, \ldots, A_n\in\mathfrak{F})$

$$\hat{\pi}(\hat{\alpha}_{t_1}(A_1)\ldots\hat{\alpha}_{t_n}(A_n))=\check{\alpha}_{t_1}\pi(A_1)\ldots\check{\alpha}_{t_n}\pi(A_n),$$

and therefore unique. The existence of $\hat{\pi}$ will follow from the convergence of "Green functions" [16, 6.3.4]

$$\lim_{r\to\infty}(F_r\Omega, \hat{\alpha}_{t_1}(A_1)\ldots\hat{\alpha}_{t_n}(A_n)F_r\Omega)\,\|F_r\Omega\|^{-2}=(\Phi, \check{\alpha}_{t_1}\pi(A_1)\ldots\check{\alpha}_{t_n}\pi(A_n)\Phi), \tag{$*$}$$

holding for all $A_1, \ldots, A_n\in\mathfrak{F}_0$, $t_1, \ldots, t_n\in\mathbb{R}$, $n\in\mathbb{N}$. To prove ($*$) we exploit the fact that T_0^{it+1} is a continuous function of T_0 and, according to Weierstrass' theorem, can be uniformly approximated by polynomials,

$$\left\|T_0^{it+1}-\sum_n a_n^{(\varepsilon)}(t)T_0^n\right\|<\varepsilon.$$

Using Theorem 6.1, the left hand side of ($*$) can be approximated uniformly in r by

$$\sum_{n_1,\ldots,n_{n+1}} a_{n_1}^{(\varepsilon)}(-t_1)a_{n_2}^{(\varepsilon)}(t_1-t_2)\ldots a_{n_{n+1}}^{(\varepsilon)}(t_n)$$
$$\cdot(F_r\Omega, T_0^{n_1}\alpha_{-i}(A_1)T_0^{n_2}\alpha_{-2i}(A_2)\ldots\alpha_{-ni}(A_n)T_0^{n_{n+1}}\alpha_{-(n+1)i}(F_r)\Omega)\,\|F_r\Omega\|^{-2}.$$

From Theorem 6.2 this converges for each $\varepsilon>0$ to

$$\sum_{n_1,\ldots,n_{n+1}} a_{n_1}^{(\varepsilon)}(-t_1)\ldots a_{n_{n+1}}^{(\varepsilon)}(t_n)$$
$$\cdot a^{2(\Sigma n_i-(n+1))}(\Phi, T^{n_1}\pi\alpha_{-i}(A_1)\ldots\pi\alpha_{-ni}(A_n)T^{n_{n+1}-(n+1)}\Phi)$$
$$\cdot(\Phi, T^{\Sigma n_i-(n+1)}\Phi).$$

Since $0\leqq a^2T\otimes T\leqq 1$ this converges for $\varepsilon\to 0$ to

$$(\Phi, T^{-it_1}\pi(A_1)\ldots\pi(A_n)T^{it_n}\Phi).$$

This proves relation ($*$). q.e.d.

The vacuum (identity) representation $\hat{\pi}_0$ of $\hat{\mathfrak{F}}$ is irreducible since Ω is the (up to a phase) unique vector in $\mathscr{H}_0$ inducing a ground state of the dynamical system $(\hat{\mathfrak{F}}, \hat{\alpha}_t)$ [25]. In the same way, $\hat{\pi}$ is irreducible since Φ_0 is the (up to a phase) unique vector in $\mathscr{H}$ with external charge configuration q_0 which is a ground state of the dynamical system $(\hat{\mathfrak{A}}, \hat{\alpha}_t)$. In fact, let X be in the commutant $\hat{\pi}(\hat{\mathfrak{F}})'$ of $\hat{\pi}(\hat{\mathfrak{F}})$. Then $X\Phi_0$ has also q_0 as external charge configuration and is therefore a multiple of Φ_0; hence from the cyclicity of Φ_0 for $\hat{\pi}(\hat{\mathfrak{F}})$ one concludes that X is a multiple of the identity. The corresponding subrepresentations $\check{\pi}_0$ and $\check{\pi}$ of $\hat{\pi}_0\restriction\hat{\mathfrak{A}}$ and $\hat{\pi}\restriction\hat{\mathfrak{A}}$ in the respective gauge invariant subspaces are also irreducible. For $\check{\pi}_0$ this follows again from the uniqueness of Ω. The argument for $\check{\pi}$ is somewhat indirect. We know that the subrepresentation $\check{\pi}_{q_0}$ of $\hat{\pi}\restriction\hat{\mathfrak{A}}$ in the subspace with external charge configuration q_0 is irreducible. The irreducibility of $\check{\pi}$ follows then from the unitary equivalence

$$\check{\pi}\simeq\check{\pi}_{q_0}\circ\varrho_0 . \tag{7.2}$$

It is now easy to see that $\hat{\pi}$ and $\hat{\pi}_0$ as well as $\check{\pi}$ and $\check{\pi}_0$ are mutually inequivalent representations. One possibility is to look at the behaviour of the translation operators. Ω induces a translation invariant state on $\hat{\mathfrak{F}}$, but $\mathscr{H}$ cannot contain a vector inducing a translation invariant state on $\hat{\mathfrak{F}}$. Namely, such a vector had to be an eigenvector of $U(\mathbf{x})$, $\mathbf{x}\in\mathbb{Z}^d$; the existence of such an eigenvector, however, is excluded by the weak convergence of $U(\mathbf{x})$ towards zero (Theorem 6.5.).

This argument completes the construction of charged states. The result is formulated in the following theorem.

Theorem 7.2. *For $\beta_g>\beta_g^w$, $\beta_h<\beta_h^w$, there exists an irreducible representation $\check{\pi}$ of $\hat{\mathfrak{A}}$ which is inequivalent to the vacuum representation. $\check{\pi}$ has no external charges, $\check{\pi}(\hat{q}(\mathbf{x}))=1$ for all $\mathbf{x}\in\mathbb{Z}^d$, it is translation covariant, and the time translations are implemented by a positive Hamiltonian H,*

$$e^{itH}\check{\pi}(A)e^{-itH}=\check{\pi}\hat{\alpha}_t(A), \qquad A\in\hat{\mathfrak{A}}.$$

Appendix A.1. Polymer Expansions

Low activity expansions for polymer systems are well known (see e.g. [2]). For the reader's convenience we sketch the essential steps. As a byproduct, our estimates seem to be (as far as we know) slightly better than the published ones.

Let $\mathscr{G}^c$ be a finite set whose elements are called polymers together with a relation "compatible," $\gamma\sim\gamma'$, such that $\gamma\nsim\gamma$ for all $\gamma\in\mathscr{G}^c$. A subset Γ of $\mathscr{G}^c$ is called admissible if $\gamma\sim\gamma'$ for all $\gamma,\gamma'\in\Gamma$, $\gamma\neq\gamma'$. Let $\mathscr{G}$ denote the set of admissible subsets of $\mathscr{G}^c$. One may visualize $\mathscr{G}$ as the set of polymer configurations or as a set of graphs whose connected components are the one element subsets of $\mathscr{G}^c$. We need further the notations $\Gamma\sim\Gamma'$ if $\gamma\sim\gamma'$ for all $\gamma\in\Gamma$, $\gamma'\in\Gamma'$, $\Gamma_{\Gamma'}=\{\gamma\in\Gamma'|\gamma\nsim\Gamma'\}$ and $\mathrm{Conn}(\Gamma)=\{\Gamma'\in\mathscr{G}|\gamma'\nsim\Gamma$ for all $\gamma'\in\Gamma'\}$. We shall often identify $\Gamma\in\mathscr{G}^c$ with its characteristic function. Other functions $\Gamma:\mathscr{G}^c\to\mathbb{Z}_+$ will also occur; we consider them as multi-indices of power series in variables indexed by the elements of $\mathscr{G}^c$.

A polymer model on $\mathscr{G}$ is defined by assigning to each polymer $\gamma\in\mathscr{G}^c$ an activity $\mu(\gamma)\in\mathbb{C}$. Consider the partition function

$$Z=\sum_{\Gamma\in\mathscr{G}}\mu^\Gamma, \qquad \mu^\Gamma=\prod_\gamma\mu(\gamma)^{\Gamma(\gamma)}, \tag{A.1}$$

and the correlation functions

$$\varrho(\Gamma)=Z^{-1}\sum_{\substack{\Gamma'\in\mathscr{G}\\ \Gamma'\sim\Gamma}}\mu^{\Gamma'}, \qquad \Gamma\in\mathscr{G}. \tag{A.2}$$

The identity

$$\begin{aligned}\sum_{\Gamma'\sim\Gamma}\mu^{\Gamma'}&=\sum_{\Gamma'}(1-1)^{|\Gamma'_\Gamma|}\mu^{\Gamma'}=\sum_{\Gamma'}\sum_{\Gamma''\subset\Gamma'_\Gamma}(-1)^{|\Gamma''|}\mu^{\Gamma'}\\ &=\sum_{\Gamma''\in\mathrm{Conn}(\Gamma)}(-\mu)^{\Gamma''}\sum_{\Gamma'\sim\Gamma''}\mu^{\Gamma'}\end{aligned} \tag{A.3}$$

leads to the equation

$$\varrho(\Gamma)=\sum_{\Gamma'\in\mathrm{Conn}(\Gamma)}(-\mu)^{\Gamma'}\varrho(\Gamma'). \tag{A.4}$$

This equation of the Kirkwood-Salsburg type together with the normalization condition $\varrho(\emptyset)=1$ has the unique solution (in the sense of formal power series in the activities)

$$\varrho(\Gamma_0)=\lim_{n\to\infty}\sum_{\substack{\Gamma_1,\dots,\Gamma_n\in\mathscr{G}\\ \Gamma_i\in\mathrm{Conn}(\Gamma_{i-1}),\, i=1,\dots,n}}(-\mu)^{\Gamma_1+\dots+\Gamma_n}. \tag{A.5}$$

Equation (A.5) leads immediately to graph theoretical expressions for the coefficients c_Γ, $\Gamma:\mathscr{G}^c\to\mathbb{Z}_+$, which appear in the power series expansion of $\ln Z$.

For an investigation of the convergence of (A.5) we associate a length $|\gamma|\in\mathbb{N}$ to each $\gamma\in\mathscr{G}^c$ and let $\|\Gamma\|=\sum\Gamma(\gamma)|\gamma|$. We assume that there is a convex, differentiable, monotonically decreasing function $F_0:(b_0,\infty)\to\mathbb{R}_+$, $b_0\in\mathbb{R}$, such that for each $\Gamma\in\mathscr{G}$, $b>b_0$

$$\sum_{\gamma \nsim \Gamma} e^{-b|\gamma|}\leqq F_0(b)\,\|\Gamma\|. \tag{A.6}$$

This implies

$$\sum_{\Gamma'\in\mathrm{Conn}(\Gamma)} e^{-b\|\Gamma'\|}\leqq e^{F_0(b)\|\Gamma\|}. \tag{A.7}$$

Sometimes (A.7) can be improved by taking partially into account that only admissible sets Γ' occur,

$$\sum_{\Gamma'\in\mathrm{Conn}(\Gamma)} e^{-b\|\Gamma'\|}\leqq\left(1+\frac{1}{n}F_0(b)\right)^{n\|\Gamma\|} \tag{A.7'}$$

for some $n\in\mathbb{N}$. Inequality (A.7) implies convergence of (A.5) if there exists some $a>b_0$ such that

$$\|\mu\|=\sup_\gamma|\mu(\gamma)|^{1/|\gamma|}=e^{-(a+F_0(a))}. \tag{A.8}$$

Let a_c be the smallest solution of

$$F_0'(a_c)=-1, \tag{A.9}$$

provided it exists, and $a_c=b_0$ otherwise. Then the convergence condition can be written in the more explicit form

$$\|\mu\|\leqq e^{-(a_c+F_0(a_c))}. \tag{A.10}$$

For the correlation function one obtains the bound

$$|\varrho(\Gamma)|\leqq e^{F_1(-\ln\|\mu\|)\|\Gamma\|}, \tag{A.11}$$

where $F_1:(a_c+F_0(a_c),\infty)\to\mathbb{R}_+$ is defined by

$$F_1(a+F_0(a))=F_0(a). \tag{A.12}$$

To derive the estimate (4.15) we use the identity ($\Gamma\in\mathscr{G}$)

$$\begin{aligned}\sum_{\substack{\Gamma':\mathscr{G}^c\to\mathbb{Z}_+\\ \Gamma'\nsim\Gamma}} c_{\Gamma'}\mu^{\Gamma'}&=\ln\varrho(\Gamma)=\ln Z(0)-\ln Z(1)\\ &=-\int_0^1 d\lambda\sum_{\substack{\gamma\in\mathscr{G}^c\\ \gamma\nsim\Gamma}}\mu(\gamma)\varrho_\lambda(\{\gamma\}),\end{aligned} \tag{A.13}$$

where $Z(\lambda)$ and ϱ_λ are partition function and correlation functions corresponding to the activities $\mu_\lambda(\gamma)=\lambda\mu(\gamma)$ for $\gamma \not\sim \Gamma$ and $\mu_\lambda(\gamma)=\mu(\gamma)$ otherwise, $\gamma\in\mathscr{G}^c$. Since $\|\mu_\lambda\| \leqq \|\mu\|$, $|\varrho_\lambda(\{\gamma\})|\leqq e^{F_1(-\ln\|\mu\|)|\gamma|}$, and from the definition of F_1 and from (A.6)

$$\sum_{\Gamma' \not\sim \Gamma} |c_{\Gamma'}|\,|\mu^{\Gamma'}| \leqq F_1(-\ln\|\mu\|)\,\|\Gamma\|\,. \tag{A.14}$$

Inequality (4.17) is obtained by comparison with the critical activity $\mu_c(\gamma) = e^{-(a_c+F_0(a_c))|\gamma|}$

$$\sum_{\substack{\Gamma' \not\sim \Gamma \\ \|\Gamma'\|\geqq n}} |c_{\Gamma'}|\,|\mu^{\Gamma'}| \leqq \left(\frac{\|\mu\|}{\|\mu_c\|}\right)^n \sum_{\Gamma' \not\sim \Gamma} |c_{\Gamma'}|\, e^{-(a_c+F_0(a_c))\|\Gamma'\|}$$
$$\leqq \left(\frac{\|\mu\|}{\|\mu_c\|}\right)^n \|\Gamma\|\, F_0(a_c)\,. \tag{A.15}$$

The functions F_0 for the models discussed in this paper are given in Appendix A.3. In the standard case with the combinatorial estimate

$$|\{\gamma\in\mathscr{G}^c, \gamma \not\sim \Gamma, |\gamma|=n\}| \leqq \|\Gamma\|\, c^n\,,$$

one obtains $F_0(b)=ce^{-b}(1-ce^{-b})^{-1}$ and

$$\|\mu_c\| = c^{-1}\tfrac{1}{2}(3-\sqrt{5})e^{-1/2(\sqrt{5}-1)} = c^{-1}\cdot 0.2059\,\ldots,$$
$$F_0(a_c)=\tfrac{1}{2}(\sqrt{5}-1)=0.6180\,\ldots.$$

Appendix A.2. Perimeter Law of the Wilson Loop

The fact that the Wilson loop has a perimeter law in the low temperature region of the $\mathbb{Z}_2$ gauge theory has been observed already in the classical paper of Wegner [6] in 1971. A somewhat sketchy proof appeared in [7], and a complete proof has been carried through by Göpfert [8]. For the convenience of the reader we review the proof together with some new estimates.

The proof is a simple application of the polymer expansion method outlined in Appendix A.1. Here, polymers are coconnected coclosed sets of plaquettes P (vortices), and two polymers are compatible if they are not coconnected. Let $M\subset B(\mathbb{Z}^{d+1})$ be closed. Then

$$\langle \chi_M\rangle^0 = \exp\left\{-\sum_{(P,M)=-1} e^{-2\beta_g|P|}\int_{-1}^{1} d\lambda\,\varrho_\lambda(P)\right\}, \tag{A.16}$$

where ϱ_λ denotes the correlation function corresponding to activities μ_λ, $\mu_\lambda(P) = \lambda e^{-2\beta_g|P|}$ for $(P,M)=-1$, $\mu_\lambda(P)=e^{-2\beta_g|P|}$ for $(P,M)=1$. Using (A.11) and the formulas (A.27) and (A.29) we obtain

$$\langle \chi_M\rangle^0 \geqq e^{-\alpha(\beta_g)|M|}\,, \tag{A.17}$$

where $\alpha(\beta_g)$ is given by the formula

$$\alpha(\beta_g)=F^{(12)}(a)\,,\qquad \beta_g=\tfrac{1}{2}(a+F^{(22)}(a))\,,\qquad a\geqq a_c\,, \tag{A.18}$$

$F^{(12)}$, $F^{(22)}$ being defined in Appendix A.3, and a_c being the solution of $F^{(22)\prime}(a) = -1$. The leading term for β_g large is $e^{-4(D-1)\beta_g}$. The bounds in Appendix A.3 imply

$$\alpha(\beta_g) = e^{-4(D-1)\beta_g} c(\beta_g) \tag{A.19}$$

with $c(\beta_g) \leqq c(\beta_g^w)$, $\beta_g^w = \frac{1}{2}(a_c + F^{(22)}(a_c))$. We obtain $\beta_g^w = 0.9215$, $c(\beta_g^w) = 39.76$ in $D=3$ and $\beta_g^w = 0.8953$, $c(\beta_g^w) = 340.9$ in $D=4$ dimensions.

Appendix A.3. Combinatorial Estimates

In this appendix we collect some combinatorial estimates which were needed for the convergence proofs of the expansions used in this paper. From now on we denote $D = d+1$.

Proposition A.3.1. *Let $Q \subset \mathbb{Z}^D$ be a finite set with $|Q| = 2q$, $q \in \mathbb{N}$. Then the number $\mathcal{N}(m)$ of all subsets $M \subset B(\mathbb{Z}^D)$, with $|M| = m$ and $\partial M = Q$ such that each connected component of M has a nonvoid boundary is bounded by ($D \geqq 3$)*

$$\mathcal{N}(m) \leqq \frac{q}{2}\left(\frac{m-1}{2}\right)^{q-1}(2D-1)^m .$$

Proof. The most general set M with the properties mentioned above can be found by decomposing Q in a set of pairs and joining each pair by a path (a corollary to the "Königsberger Brückenproblem"). There are $(2q)!/2^q q!$ pairings of Q, $\binom{m-1}{q-1}$ partitions of m in a sum $m = m_1 + \ldots + m_q$, $m_k \geqq 1$ and at most $2D(2D-1)^{m_k - 2}$ paths of length m_k joining a given pair of points. Thus $\mathcal{N}(m)$ is bounded by

$$\begin{aligned}\mathcal{N}(m) &\leqq (2D-1)^{m-2q} \frac{(2q)!}{2^q q!}\binom{m-1}{q-1}(2D)^q \\ &\leqq (m-1)^{q-1}(2D-1)^m q 2^{-q} . \quad \text{q.e.d.}\end{aligned}$$

In the lattice $\mathbb{Z}^D$, two i-cells are connected if their boundaries contain a common $(i-1)$-cell. Let $\mathcal{N}_i(n)$ denote the number of closed connected sets C_i of i-cells with $|C_i| = n$ which contain a given i-cell, and let

$$f_i(b) = \sum_{n=0}^{\infty} \mathcal{N}_i(n) e^{-bn} \tag{A.20}$$

denote the associated generating function.

Proposition A.3.2. *Let M be a closed set of i-cells, and let $\mathscr{C}_i(M)$ denote the set of closed connected sets of i-cells which are connected with M. Then*

$$\sum_{C_i \in \mathscr{C}_i(M)} e^{-b|C_i|} \leqq (i(D-i) + \tfrac{1}{2}) f_i(b)|M| .$$

Proof. There are at most $2i(D-i)|M|$ i-cells which are connected with M but not contained in M. Thus there are at most $(2i(D-i)+1)|M|$ possible starting points for a closed connected set of i-cells. The statement follows from the fact that each $C_i \in \mathscr{C}_i(M)$ contains at least two of them. q.e.d.

Proposition A.3.3. *Let P be a coclosed set of $(i+1)$-cells, and let $\mathscr{C}_i(P)$ denote the set of closed connected sets of i-cells with an odd winding number with respect to P. Then*

$$\sum_{C_i \in \mathscr{C}i(P)} e^{-b|C_i|} \leqq \frac{i+1}{8i(D-i+1)} |f_i'(b)|\,|P|\,.$$

Proof. Let c denote the i-cell which is spanned by the points 0 and $x^{(j)}, j=1, \ldots, i$, $x_k^{(j)} = \delta_{jk}$, let $\mathscr{C}_i(n,c)$ denote the set of closed connected sets C_i of i-cells containing c with $|C_i| = n$, and let $\mathscr{C}_i(n,P)$ denote the set of $C_i \in \mathscr{C}_i(P)$ with $|C_i| = n$. Each $C_i \in \mathscr{C}_i(n,c)$ is the boundary ∂C_{i+1} of a certain set of $(i+1)$-cells C_{i+1}, and from Proposition A.5.2, C_{i+1} can be chosen such that

$$|C_{i+1}| \leqq \frac{D-i}{8i(D-i+1)} |C_i|^2\,.$$

Consider the semidirect product G of the group of lattice translations and the group of permutations of coordinates. If $n_{j_1, \ldots, j_{i+1}}$ and $p_{j_1, \ldots, j_{i+1}}$ are the numbers of $(i+1)$-cells in C_{i+1} and P, respectively, which are parallel to the $(j_1, \ldots, j_{i+1})$-plane, then, for each permutation π there are at most

$$\sum_{j_1 < \ldots < j_{i+1}} n_{\pi(j_1) \ldots \pi(j_{i+1})} p_{j_1, \ldots, j_{i+1}}$$

possibilities to shift the transformed set πC_{i+1} such that it intersects with P. Summing over all permutations leads to at most $(i+1)!(D-i-1)!\,|C_{i+1}|\,|P|$ transformations in G which map C_i into $\mathscr{C}_i(n,P)$. Since each $C_i' \in \mathscr{C}_i(n,P)$ can be reached by $|C_i'|\,i!(D-i)!$ transformations we obtain

$$(i+1)!(D-i-1)!\frac{D-i}{8i(D-i+1)} n^2 |P|\,|\mathscr{C}_i(n,c)| \geqq i!(D-i)!\,n|\mathscr{C}_i(n,P)|\,.$$

The statement follows then from (A.20). q.e.d.

$\mathscr{N}_i(n)$ can be estimated by an application of the solution of the "Königsberger Brückenproblem." There the i-cells and their faces are considered as islands, and each i-cell is connected with its faces by bridges. The most general set C_i can be found by choosing a path which meets every bridge once. This leads to the bound

$$\mathscr{N}_i(n) \leqq 2[(2i-1)(2D-2i+1)]^{in-(i+2)}\,, \tag{A.21}$$

where we used the fact that the first step is arbitrary and that a different choice of the last $2i+3$ steps can lead at most to one other closed set since the smallest closed set of i-cells has $2(i+1)$ elements.

For the generating function one gets

$$f_i(b) \leqq c_i'(c_i e^{-b})^{2(i+1)}(1 - c_i^2 e^{-2b})^{-1} \tag{A.22}$$

with $c_i' = 2[(2i-1)(2D-2i+1)]^{-(i+2)}$ and $c_i = [(2i-1)(2D-2i+1)]^i$.

If $i > 1$ it is more effective to estimate first the number $\mathscr{N}_i^{(0)}(n)$ of elementary closed sets of i-cells containing a given one, where we call a closed set of i-cells elementary if it cannot be written as a disjoint union of nonvoid closed sets of i-cells. Let

$$f_i^{(0)}(b) = \sum_{n=0}^{\infty} \mathscr{N}_i^{(0)}(n) e^{-bn} \tag{A.23}$$

denote the associated generating function. Note that Proposition A.3.3. holds also for elementary sets if one replaces f_i by $f_i^{(0)}$.

Proposition A.3.4.

$$f_i(a+i(D-i)f_i^{(0)}(a)) \leqq f_i^{(0)}(a).$$

Proof. Each closed connected set of i-cells C_i containing the i-cell c can be decomposed into elementary ones

$$C_i = E_0 \cup E_1^{(1)} \cup \ldots \cup E_1^{(k_1)} \cup \ldots \cup E_n^{(1)} \cup \ldots \cup E_n^{(k_n)},$$

such that $c \in E_0$ and $E_j^{(l)}$ is connected with $E_{j-1}^{(l')}$ for some l' but does not intersect with $E_{j-1}^{(l')}$, $j=1, \ldots, n$. Then

$$\sum_{k_n=0}^{\infty} \frac{1}{k_n!} \sum_{E_n^{(1)}, \ldots, E_n^{(k_n)}} e^{-(a+i(D-i)f_i^{(0)}(a))(|E_n^{(1)}|+\ldots+|E_n^{(k_n)}|)} \leqq e^{i(D-i)f_i^{(0)}(a)(|E_{n-1}^{(1)}|+\ldots+|E_{n-1}^{(k_{n-1})}|)},$$

where we used the fact that $f_i^{(0)}$ is monotonically decreasing. A repeated application of this estimate leads to

$$\sum_{C_i \ni c} e^{-(a+i(D-i)f_i^{(0)}(a))|C_i|} \leqq \sum_{E_0 \ni c} e^{-a|E_0|},$$

which proves the statement. q.e.d.

$\mathcal{N}_i^{(0)}(n)$ can be estimated similarly to Ruelle's estimate on the number of elementary closed sets of $(D-1)$-cells (Peierls contours) [22] (compare also [24]). This leads to the bounds

$$\mathcal{N}_i^{(0)}(n) \leqq 2(2D-2i+1)^{n-(i+2)}, \tag{A.24}$$

and

$$f_i^{(0)}(b) \leqq c_i^{(0)\prime}(c_i^{(0)}e^{-b})^{2(i+1)}(1-c_i^{(0)2}e^{-2b})^{-1}, \tag{A.25}$$

with

$$c_i^{(0)\prime} = 2(2D-2i+1)^{-(i+2)}, \qquad c_i^{(0)} = 2D-2i+1. \tag{A.26}$$

After these general considerations we have the means to estimate the generating function for the Marra-Miracle-Solé expansion of the gauge invariant Ising model which was reviewed in Sect. 4. It is natural to associate to each $\gamma \in \mathcal{G}^c$ a pair of lengths $|\gamma| = (|N_\gamma|, |P_\gamma|)$. We shall equip $\mathbb{R}^2$ with the usual scalar product $xy = x_1y_1 + x_2y_2$. From Proposition A.3.3 and A.3.4 we have for $N \subset B(\mathbb{Z}^D)$, $\partial N = \emptyset$

$$\sum_{\substack{P \subset P(\mathbb{Z}^D) \\ (P,N)=-1 \\ P \text{ coconnected}}} e^{-b|P|} \leqq F^{(12)}(b)|N|,$$

$$F^{(12)}(b) = \frac{D-1}{24(D-2)} |f_{D-2}^{(0)\prime}(a)|, \tag{A.27}$$

$$b = a + 2(D-2)f_{D-2}^{(0)}(a),$$

and for $P \subset P(\mathbb{Z}^D)$, $\partial^* P = \emptyset$,

$$\sum_{\substack{N \subset B(\mathbb{Z}^D) \\ (P,N) = -1 \\ N \text{ connected}}} e^{-b|N|} \leqq F^{(21)}(b)|P|, \qquad F^{(21)}(b) = \frac{1}{4D}|f_1'(b)|. \tag{A.28}$$

If $D=3$ one may use $F^{(12)} = F^{(21)}$ instead of (A.27). Defining

$$F^{(11)}(b) = (D - \tfrac{1}{2})f_1(b), \qquad F^{(22)}(b) = (2(D-2) + \tfrac{1}{2})f_{D-2}(b), \tag{A.29}$$

we arrive at the estimate ($b = (b_1, b_2) \in \mathbb{R}^2$, $\gamma' = (N_{\gamma'}, P_{\gamma'}) \in \mathcal{G}$)

$$\sum_{\substack{\gamma \in \mathcal{G}^c \\ \gamma \not\perp \gamma'}} e^{-b|\gamma|} \leqq F_0(b)|\gamma'|, \tag{A.30}$$

with

$$\begin{aligned} &F_0(a + F(a)) = F(a) + \hat{F}(a), \quad a \in \{x \in \mathbb{R}^2, \det(1 + F'(x)) \geqq 0\}, \\ &F(a_1, a_2) = (F^{(12)}(a_2), F^{(21)}(a_1)), \quad \hat{F}(a_1, a_2) = (F^{(11)}(a_1), F^{(22)}(a_2)). \end{aligned} \tag{A.31}$$

The region $\mathcal{R}_0 = \{b = a + F(a), \det(1 + F'(a)) \geqq 0\}$ in which the left hand side of (A.30) is bounded is somewhat smaller than the one given in [10]; we could not reproduce this stronger result.

The function F_0 occuring in Sect. 4 (denoted here by $F_0^{(4)}$) is related to the function F_0 in (A.30) by

$$F_0^{(4)}(b) = \max_{i=1,2} F_0(b,b)_i.$$

Appendix A.4. Convergence Region of the Marra-Miracle-Solé Expansion

The estimates in Appendix A.3 relied on the estimate of the number $\mathcal{N}_1$ of paths in $\mathbb{Z}^D$ starting from a given bond

$$\ln \mathcal{N}_1/\text{length} \to 2D - 1, \tag{A.32}$$

and on the number $\mathcal{N}_{D-2}^{(0)}$ of elementary coclosed sets of plaquettes starting from a given one

$$\ln \mathcal{N}_{D-2}^{(0)}/\text{length} \to 5, \tag{A.33}$$

if the length tends to infinity. Therefore the convergence region which we shall establish cannot exceed the region $\text{th}\beta_h < (2D-1)^{-1}$ and $e^{-2\beta_g} < 1/5$. We shall come fairly close to this borderline; so a further improvement of the bound on the convergence region should rely on an improvement of these basic estimates.

According to Appendix A.1, the polymer expansion of Marra and Miracle-Solé will converge if

$$(\text{th}\beta_h, e^{-2\beta_g}) = (e^{-b_1}, e^{-b_2}), \tag{A.34}$$

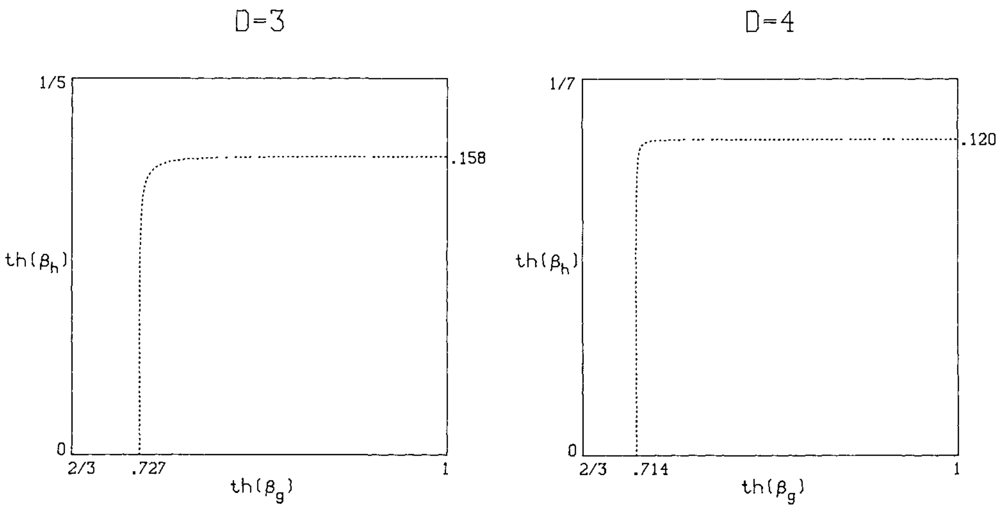

Fig. A.1a and b. The bounds on the convergence region, as estimated in Appendix A.4

where $(b_1, b_2) \in \mathscr{R} = \{a + F_0(a), a \in \mathscr{R}_0\}$, F_0 and $\mathscr{R}_0$ being defined in Appendix A.3. Using Eqs. (A.27)–(A.31) it follows that

$$\mathscr{R} = \{b = a + H(a)\}, \qquad H(a) = \hat{F}(a) + 2F(a). \tag{A.35}$$

H is monotonically decreasing and H' is monotonically increasing with respect to the partial ordering

$$x \geqq y \Leftrightarrow x_1 \geqq y_1 \quad \text{and} \quad x_2 \geqq y_2. \tag{A.36}$$

Hence $\mathscr{R} = \{b \in \mathbb{R}^2, b \geqq b_c$ for some $b_c \in B_c\}$ with

$$B_c = \{a_c + H(a_c), a_c \in A_c\}, \tag{A.37}$$

A_c being the set of maximal solutions of

$$\det(1 + H'(a)) = 0. \tag{A.38}$$

The convergence region determined by this method is shown in Fig. A.1a ($D=3$) and Fig. A.1b ($D=4$). In particular it contains the convergence regions of the pure gauge theory and of the Ising model.

Appendix A.5

Proposition A.5.1. *Let M_r be a square with side length $2r$ contained in a lattice plane. For each set of plaquettes P*

$$|P| \geqq \frac{r}{2}\{|M_r| - |\partial P \triangle M_r|\} = \frac{r}{2}\{2|\partial P \cap M_r| - |\partial P|\}.$$

Proof. Let P_r be the minimal set of plaquettes with $\partial P_r = M_r$ (i.e. $|P_r| = 4r^2$), and P' the projection of P on the plane containing the square. Then $|P'| \leqq |P|$ and

$|\partial P' \triangle M_r| \leqq |\partial P \triangle M_r|$. Since a plane surface with boundary of length l has at most an area of $\left(\frac{l}{4}\right)^2$, $|P' \triangle P_r| \leqq \frac{|\partial P' \triangle M_r|^2}{16}$. Now,

$$|P'| = |P'| + |P' \triangle P_r| - |P' \triangle P_r| \geqq |P_r| - |P' \triangle P_r| = 4r^2 - |P' \triangle P_r|$$
$$\geqq 4r^2 - \frac{|\partial P' \triangle M_r|^2}{16} \geqq 2r\left(2r - \frac{|\partial P' \triangle M_r|}{4}\right) \geqq \frac{r}{2}(8r - |\partial P \triangle M_r|). \quad \text{q.e.d.}$$

Proposition A.5.2. *Let C_i be a finite closed set of i-cells in $\mathbb{Z}^D$. Then there exists a set of $(i+1)$-cells C_{i+1}, such that $\partial C_{i+1} = C_i$ and*

$$|C_{i+1}| \leqq \frac{D-i}{8i(D-i+1)} |C_i|^2 .$$

Proof. We shall only consider the case $i=1$. In this case the above bound is optimal, as may be seen by chosing C_1 as the path of length $2D$ connecting successively the points $0, e_1, e_1 + e_2, \ldots, e_1 + \ldots + e_D, e_2 + \ldots + e_D, e_3 + \ldots + e_D, \ldots, 0$ (e_i denotes the i^{th} unit vector) the minimal area of which is $D(D-1)/2$. The proof for $i > 1$ is completely analogous, but it is obvious that the bound obtained is not optimal.

It is sufficient to consider connected sets C_1. For $D=1$ the inequality is trivially satisfied. Let $D \geqq 2$ and assume that the statement of the proposition is true for $D-1$. We may chose a hyperplane and a set of vertical plaquettes C_2' such that

$$\partial C_2' = C_1 \triangle C_1' ,$$

where C_1' is the projection of C_1 onto the distinguished hyperplane. Let $x|C_1|$ denote the number of horizontal bonds in C_1, $0 \leqq x \leqq 1$. Then $|C_1'| \leqq x|C_1|$ and for a suitable level of the hyperplane

$$|C_2'| \leqq \tfrac{1}{4} x(1-x) |C_1|^2 .$$

By the induction hypothesis, there exists a set of plaquettes C_2'' in the hyperplane with $\partial C_2'' = C_1'$ and

$$|C_2''| \leqq |C_1|^2 \frac{D-2}{8(D-1)} x^2 .$$

Then $C_2 = C_2' \triangle C_2''$ has the properties required in the proposition. q.e.d.

Acknowledgements. The possibility that the low temperature region of $\mathbb{Z}_2$ gauge theories might provide an example for nonlocalizable particles has been pointed out to one of us (K.F.) by Gerhard Mack; this was actually the starting point for the present investigation and is gratefully acknowledged. We also profitted from discussions with Gernot Münster, Hildegard Meyer and Charles Pfister.

References

1. Wilson, K.: Confinement of quarks. Phys. Rev. D**10**, 2445 (1974)
2. Seiler, E.: Gauge theories as a problem of constructive quantum field theory and statistical mechanics. In: Lecture Notes in Physics, Vol. 159. Berlin, Heidelberg, New York: Springer 1982

3. Mack, G., Meyer, H.: A disorder parameter that tests for confinement in gauge theories with quark fields. Nucl. Phys. **200**B, 249 (1982)
4. Bricmont, J., Fröhlich, J.: An order parameter distinguishing between different phases of Lattice gauge theories with matter fields. Phys. Lett. **122**B, 73–77 (1983)
5. Fredenhagen, K.: Structural aspects of gauge theories in the algebraic framework of quantum field theories, (Talk presented at the Colloquium in honour of Prof. R. Haag at the occasion of his 60th birthday, Hamburg, November 15, 1982), University Freiburg, THEP 82/9
6. Wegner, F.: Duality in generalized Ising models and phase transitions without local order parameters. J. Math. Phys. **12**, 2259 (1971)
7. Gallavotti, G., Guerra, F., Miracle-Solé, S.: Mathematical problems in theoretical physics, In: Lecture Notes in Physics, Vol. 80. dell'Antonio, G., Doplicher, S., Jona-Lasinio, G., (eds.). Berlin, Heidelberg, New York: Springer 1978
8. Göpfert, M.: Das Peierls-Argument für $Z(2)$-Gitter-Eichtheorien, Diplomarbeit Hamburg (1979) (unpublished)
9. Fradkin, E., Shenker, S.: Phase diagrams of Lattice gauge theories with Higgs fields. Phys. Rev. D**19**, 3682 (1979)
10. Marra, R., Miracle-Solé, S.: On the statistical mechanics of the gauge invariant Ising model. Commun. Math. Phys. **67**, 233 (1979)
11. Buchholz, D., Fredenhagen, K.: Locality and the structure of particle states. Commun. Math. Phys. **84**, 1–54 (1982)
12. 't Hooft, G.: A property of electric and magnetic flux in nonabelian gauge theories. Nucl. Phys. B**153**, 141 (1979)
 't Hooft, G.: On the phase transition towards permanent quark confinement. Nucl. Phys. B**138**, 1 (1978)
13. Fredenhagen, K., Marcu, M.: Order parameters for lattice gauge theories with matter fields (to be published)
14. Osterwalder, K., Seiler, E.: Gauge field theories on a lattice. Ann. Phys. **110**, 440 (1978)
15. Lüscher, M.: Construction of a selfadjoint, strictly positive transfer matrix for Euclidean lattice gauge theories. Commun. Math. Phys. **54**, 283 (1977)
16. Bratteli, O., Robinson, D.: Operator algebras and statistical mechanics. I, II. Berlin, Heidelberg, New York: Spriner 1981
17. Ciorănescu, I., Zsidó, L.: Analytic generators for one-parameter groups, Tôhoku Math. J. **28**, 327–362 (1976)
18. Reed, M., Simon, B.: Methods of modern mathematical physics. New York, San Francisco, London: Academic Press 1972
19. Glimm, J., Jaffe, A.: Quantum physics. A functional integral point of view, Berlin, Heidelberg, New York: Springer 1981
20. Fredenhagen, K.: On the existence of dynamics in Euclidean lattice gauge theories (to be published)
21. Gantmacher, F.R.: Matrizenrechnung. Berlin: VEB Deutscher Verlag der Wissenschaften 1959
22. Ruelle, D.: Statistical mechanics. Reading, MA: Benjamin 1969
23. Bricmont, J., Lebowitz, J.L., Pfister, C.-E.: Non-translation invariant Gibbs states with coexisting phases. III. Analyticity properties. Commun. Math. Phys. **69**, 267–291 (1979)
24. Balian, R., Drouffe, J.M., Itzykson, C.: Gauge fields on a lattice. II. Gauge invariant Ising model. Phys. Rev. D**11**, 2098 (1975)
25. Streater, R.F., Wightman, A.S.: PCT, spin and statistics and all that. New York: Benjamin 1964

Communicated by R. Haag

Received July 20, 1983